*J. Fluid Mech.* (2002), *vol.* 458, *pp.* 75–101.  
DOI: 10.1017/S0022112002007747 Printed in the United Kingdom

# An explicit potential-vorticity-conserving approach to modelling nonlinear internal gravity waves

By ÁLVARO VIÚDEZ AND DAVID G. DRITSCHEL

Mathematical Institute, University of St Andrews, St Andrews, UK

(Received 15 June 2000 and in revised form 27 September 2001)

This paper discusses a potential-vorticity-conserving approach to modelling nonlinear internal gravity waves in a rotating Boussinesq fluid. The focus of the work is on the pseudo-plane motion (motion in the $x, z$-plane), for which we present a broad range of numerical results. In this case there are two material coordinates, the density and the $y$-component of the velocity in the inertial frame of reference, which are related to the $x$ and $z$ displacements of fluid particles relative to a reference configuration. The amount of potential vorticity within a fluid region bounded by isosurfaces of these material coordinates is proportional to the area within this region, and is therefore conserved as well. Two new potentials, defined in terms of the displacements and combining the vorticity and density fields, are introduced as new dependent variables. These potentials entirely govern the dynamics of internal gravity waves for the linearized system when the basic state has uniform potential vorticity. The final system of equations consists of three prognostic equations (for the potential vorticity and the Laplacians of the two potentials) and one diagnostic equation, of Monge–Ampère type, for a third potential. This diagnostic equation arises from the nonlinear definition of potential vorticity. The ellipticity of the Monge–Ampère equation implies both inertial and static stability. In three dimensions, the three potentials form a vector, whose (three-dimensional) Laplacian is equal to the vorticity plus the gradient of the perturbation density.

Numerical simulations are carried out using a novel algorithm which directly evolves the potential vorticity, in a Lagrangian manner (following fluid particles), without diffusion. We present results which emphasize the way in which potential vorticity anomalies modify the characteristics of internal gravity waves, e.g. the propagation of internal wave packets, including reflection, refraction, and amplification. We also show how potential vorticity anomalies may generate internal gravity waves, along with the subsequent 'geostrophic adjustment' of the flow to a 'balanced' wave-less state. These examples, and the straightforward extension of the theoretical and numerical approach to three dimensions, point to a direct and accurate means to elucidate the role of potential vorticity in internal gravity wave interactions. As such, this approach may help a better understanding of the observed characteristics of internal gravity waves in the oceans.

## 1. Introduction

Internal gravity waves (IGWs) are ubiquitous in the oceans and atmosphere and have been the subject of extensive research. Since Thomson (1879), this research has

therefore encompassed many observational, numerical and theoretical studies of the kinematics, dynamics and thermodynamics of IGWs. This has included the generation, propagation (including reflection and refraction), interaction, instability, breaking, and dissipation of IGWs in both rotating and non-rotating (inertial) reference frames.

As research on IGWs has progressed, an increasing number of complex nonlinear effects have been discovered. That nonlinearity is important has been revealed by many works. For instance, Holloway (1980, 1982) showed that there are serious limitations to weak-interaction theory in describing nonlinear interactions between IGWs. Subsequently, Broutman (1984), Broutman & Young (1986), and Broutman & Grimshaw (1988) examined the nonlinear interaction between an IGW packet and an inertial current, and Müller *et al.* (1986) examined the nonlinear interaction of several IGWs. More recently, Thorpe (1998) studied the nonlinear reflection of internal waves from a density discontinuity at the base of the oceanic mixed layer. The two-dimensional nonlinear instability of IGW packets near critical levels (where the wave speed equals the local current speed) in the absence of rotation was investigated by Winters & D'Asaro (1989), and the generation of potential vorticity (PV) by viscosity and diffusion associated with three-dimensional convective instability of IGWs, also in the absence of rotation, was studied by Winters & D'Asaro (1994). Using the two-dimensional Boussinesq equations (again without rotation), Sutherland (1996, 1999) found that nonlinear effects may significantly modify the transmission of IGW packets incident upon a reflecting level, where the Doppler-shifted frequency of the waves is comparable to the background Brunt–Väisälä frequency. These are just a few of the studies that emphasize the role of nonlinearity in IGW interactions.

Much of the progress in IGW research has stemmed from the fact that simple plane waves, of any amplitude and wave vector, are solutions of both the linear and nonlinear incompressible Boussinesq equations, since the nonlinear terms vanish identically. Furthermore, it turns out that PV, a quantity with a nonlinear dependence on the velocity field, is not only materially conserved but is also homogeneous in the case of plane waves. This PV homogeneity, however, does not generally follow for superposed IGWs, which may or may not have PV anomalies. It may be useful in fact to distinguish IGWs with and without PV anomalies, as argued by Müller *et al.* (1986), since PV anomalies are not wave-like but rather materially conserved quantities. This leads to an alternative way of studying nonlinear IGW processes in a rotating stratified fluid; in particular, it permits one to see directly the often hidden effects of PV anomalies. In other words, the analysis of fluid motion in terms of PV anomalies can be used in an effective way to study the complexity of nonlinear IGW processes. The objective of the present work is to introduce a novel, natural approach for this purpose.

The theoretical basis of our approach is developed in § 2. Here, commencing from momentum balance and mass conservation, we derive the vorticity equation and introduce two special material variables in terms of which a material displacement vector is defined and the material vorticity is explicitly computed. The displacement variables are used to define two new potentials, useful for the numerical method, which replace the material variables as dependent fields. The PV balance is then introduced in terms of these potentials, and the system of equations (three prognostic and one diagnostic) is completed. Finally, the elliptic-hyperbolic nature of the diagnostic equation and its relation to inertial and static stability is discussed.

The numerical procedure to solve the equations is described in § 3. The numerical code is a pseudospectral algorithm similar to that developed for the two-dimensional shallow-water equations (Dritschel, Polvani & Mohebalhojeh 1999). The material

conservation of potential vorticity is dealt with explicitly using 'contour advection' (Dritschel & Ambaum 1997), and the nonlinear diagnostic equation is solved iteratively.

A wide range of numerical results are presented in § 4. These include illustrations of how PV anomalies modify plane IGW fields, of the interaction between PV anomalies and IGW packets, and of the generation of IGWs by PV anomalies. The theoretical basis for the generalization of this approach to three dimensions is described in § 5. Finally, our conclusions are given in § 6.

## 2. Theoretical development

### 2.1. *Basic equations*

In the following, we will use the subindices p and n to denote the vector component parallel and normal to the $(x, z)$-plane, respectively. Cartesian coordinates $(x, y, z)$ are used throughout the paper. We consider isochoric motion such that, in a reference frame rotating with constant angular velocity $f/2$, the velocity field is $\boldsymbol{u} = \boldsymbol{u}_{\rm p} + \boldsymbol{u}_{\rm n}$, where $\boldsymbol{u}_{\rm p}(x, z, t) = u(x, z, t)\boldsymbol{\imath} + w(x, z, t)\boldsymbol{k}$, and $\boldsymbol{u}_{\rm n}(x, z, t) = v(x, z, t)\boldsymbol{\jmath}$. This motion is referred to as pseudo-plane motion of the second kind (Truesdell & Toupin 1960, p. 329) because velocities do not depend on the coordinate $y$ yet $v \neq 0$. If $v = 0$ the motion is referred to as pseudo-plane motion of the first kind (see Marris 1998). We denote pressure by $p = p(x, z, t)$, and density by $\rho = \rho(x, z, t)$. In this case the non-hydrostatic Boussinesq momentum, mass conservation, and the incompressibility equations are

$$\dot{u} - fv = -\alpha_0 p_x, \tag{2.1a}$$

$$\dot{v} + fu = 0, \tag{2.1b}$$

$$\dot{w} = -\alpha_0 p_z - \alpha_0 g\rho, \tag{2.1c}$$

$$\dot{\rho} + \rho \operatorname{div} \boldsymbol{u} = 0, \tag{2.1d}$$

$$\operatorname{div} \boldsymbol{u} = 0, \tag{2.1e}$$

where the initial unknowns are the three-dimensional velocity field $(u, v, w)$, the pressure $p$, and the density $\rho$. Subindices $(x, y, z)$ denote partial derivatives; $\dot{(\,)} = \mathrm{d}(\,)/\mathrm{d}t = (\,)_t + \boldsymbol{u} \cdot \operatorname{grad}(\,)$ denotes the material derivative and the rest of the symbols have the usual meaning (see table 1). These equations are transformed using the two-dimensional streamfunction $\psi$ as a new independent field, in terms of which $\boldsymbol{u}_{\rm p} = -\boldsymbol{\jmath} \times \nabla\psi$.

### 2.2. *Vorticity*

The absolute vorticity $\boldsymbol{\omega}^a \equiv \operatorname{curl} \boldsymbol{u} + f\boldsymbol{k} = (\xi, \eta, \zeta + f)$ can be decomposed into two components: $\boldsymbol{\omega}^a_{\rm p} \equiv -v_z\boldsymbol{\imath} + (v_x + f)\boldsymbol{k}$ (parallel to the $x, z$-plane and hereinafter referred to as plane vorticity), and $\boldsymbol{\omega}_{\rm n} \equiv (u_z - w_x)\boldsymbol{\jmath} = \eta\boldsymbol{\jmath} = -\nabla^2\psi\boldsymbol{\jmath}$ (normal to the $x, z$-plane). Note that the absolute plane vorticity depends only on the spatial gradients of $v$ and on the planetary vorticity $f$. The curl of the absolute acceleration $\boldsymbol{a}^a = \dot{\boldsymbol{u}} + f\boldsymbol{k} \times \boldsymbol{u} + \nabla\phi_c$ (where $\phi_c$ is the potential for the centripetal acceleration)

$$\operatorname{curl} \boldsymbol{a}^a = \operatorname{curl} \boldsymbol{a}^a_{\rm p} = \alpha_0 g \rho_x \boldsymbol{\jmath} = (\operatorname{curl} \boldsymbol{a}^a)_{\rm n}, \tag{2.2}$$

is generally non-zero, the motion is not circulation-preserving; however, since its plane component is zero ($(\operatorname{curl} \boldsymbol{a}^a)_{\rm p} = \boldsymbol{0}$), some of the distinctive features of circulation-preserving motions can be applied to the $(x, z)$-plane. With respect to a non-rotating

| Symbol | Description |
|---|---|
| $\boldsymbol{u}_{\rm p}(x,z,t)=(u,w)$ | two-dimensional velocity |
| $\rho(x,z,t)$ | mass density field |
| $\rho_0$ | averaged mass density |
| $\alpha_0\equiv\rho_0^{-1}$ | specific volume (constant) |
| $\rho'(z)=\rho'_z z+\rho_0$ | ambient density (linear) |
| $N^2\equiv-\alpha_0\, g\rho'_z$ | square of the background Brunt–Väisälä frequency (constant) |
| $p$ | pressure |
| $g$ | acceleration due to gravity |
| $f$ | Coriolis parameter (constant) |
| $\sigma\equiv-\alpha_0\, g(\rho-\rho')$ | buoyancy |
| $X\equiv f^{-1}v+x$ | $x$-coordinate of the $y$-momentum isosurface in the reference configuration |
| $Z\equiv N^{-2}\sigma+z$ | isopycnal height in the reference configuration |
| $\mathscr{X}=(\mathscr{X},\mathscr{Y},\mathscr{Z})\equiv\boldsymbol{X}-\boldsymbol{x}$ | displacement vector (dual) |
| $\nabla=\boldsymbol{\imath}\partial/\partial x+\boldsymbol{k}\partial/\partial z$ | two-dimensional gradient operator |
| grad, div, curl | three-dimensional operators |
| Grad | three-dimensional gradient in the material description |
| $\nabla\mathscr{X}_{\rm p}$ | two-dimensional displacement gradient (dual) |
| $\boldsymbol{\varphi}_{\rm p}\equiv(\varphi,\phi)$ | |
| $\mathscr{A}_{\rm p}=(\mathscr{A},\mathscr{C})\equiv\nabla^2\boldsymbol{\varphi}_{\rm p}$ | |
| $\boldsymbol{T}\equiv\boldsymbol{u}_{\rm p}(\boldsymbol{\jmath}\times\boldsymbol{\varphi}_{\rm p})$ | |
| $\boldsymbol{\omega}\equiv\operatorname{curl}\boldsymbol{u}=(\xi,\eta,\zeta)$ | relative vorticity |

TABLE 1. Description of symbols used.

frame ($f=0$), the plane and perpendicular components of the vorticity equation ($\dot{\boldsymbol{\omega}}-\boldsymbol{\omega}\cdot\operatorname{grad}\boldsymbol{u}=\operatorname{curl}\boldsymbol{a}$) satisfy the relations

$$(\dot{\boldsymbol{\omega}})_{\rm p}=\boldsymbol{\omega}_{\rm p}\cdot\operatorname{grad}\boldsymbol{u},\quad(\dot{\boldsymbol{\omega}})_{\rm n}=\alpha_0 g\rho_x\boldsymbol{\jmath}. \tag{2.3a,b}$$

Equation (2.3$a$) states the conservation of the plane material vorticity following fluid particles (as we will explicitly prove below). In the rotating reference frame, (2.3$b$) can be written in terms of the $y$-component of the relative vorticity $\eta=-\nabla^2\psi$ as

$$\dot{\eta}=fv_z-\sigma_x. \tag{2.4}$$

### 2.3. *Material variables*

The velocity component $v$ and the buoyancy $\sigma$ evolve according to

$$\mathrm{d}(v+fx)/\mathrm{d}t=0,\quad\mathrm{d}(\sigma+N^2z)/\mathrm{d}t=0, \tag{2.5a,b}$$

respectively. The first equation expresses conservation of the $y$-component of the absolute velocity (i.e. the component along the $y$-axis rotating with frequency $f$ with respect to an inertial frame) when the total force in the $y$-direction is zero (the centrifugal acceleration being omitted). The second equation is just an equivalent way of rewriting the conservation of mass density since $\sigma(x,z)+N^2z=-\alpha_0 g[\rho(x,z)-\rho_0]$. Note that the background Brunt–Väisälä frequency $N$ is constant.

At this point, it is convenient to introduce the following fields:

$$X(x,z,t) \equiv x + \mathscr{X}(x,z,t), \quad Z(x,z,t) \equiv z + \mathscr{Z}(x,z,t), \tag{2.6a,b}$$

where

$$\mathscr{X} \equiv f^{-1}v, \quad \mathscr{Z} \equiv N^{-2}\sigma. \tag{2.7a,b}$$

The field $X(x,z,t)$ assigns to the isoline of absolute $y$-velocity located at $(x,z,t)$ a number that corresponds to the $x$-distance (from the origin $x=0$) that that isoline has in the reference configuration where $v=0$. Likewise, the field $Z(x,z,t)$ assigns to the isopycnal located at $(x,z,t)$ a number that corresponds to the $z$-distance (from the origin $z=0$) that that isoline has in a reference configuration where $\sigma=0$. These field transformations do not result in any approximation of the original equations (2.1*b*, *c*, *d*,); instead, they are just a way of considering the conserved absolute $y$-velocity and density (note that $Z=(\rho-\rho_0)/\rho_z'$) of every fluid particle in terms of its location (distance) from the 'unperturbed' reference configuration, in which $\mathscr{X}=0$ ($v=0$) and $\mathscr{Z}=0$ ($\sigma=0$). This reference configuration need not be the configuration of the system at any specific time, and we will see that for some initial conditions, it is never an actual system configuration.

Only for completeness we introduce $Y(x,y,z,t) = y + \mathscr{Y}(x,z,t)$, the $y$-coordinate in the reference configuration of the fluid particle located on $(x,y,z)$ at time $t$. Thus, by definition, $\dot{Y}=0$, and $\dot{\mathscr{Y}} = -\dot{y} = -v$. The fields $-\mathscr{X}(x,z,t) \equiv x - X(x,z,t)$, $-\mathscr{Y}(x,y,z,t) \equiv y - Y(x,z,t)$, and $-\mathscr{Z}(x,z,t) \equiv z - Z(x,z,t)$, represent the $x$-, $y$- and $z$-distance of the respective isoline at $(x,y,z)$ (of absolute $y$-velocity and density) from its position in the reference configuration. The vector $-\boldsymbol{\mathscr{X}} \equiv -(\mathscr{X},\mathscr{Y},\mathscr{Z})$ is called the displacement vector, and $\boldsymbol{\mathscr{X}}$ its dual (Truesdell & Toupin 1960, § 19). Its components may be thought of as displacements from the reference configuration, keeping in mind that the reference configuration need not be realizable. In terms of $X$, and $Z$, the two conservation equations for $v$ and $\rho$ are simply

$$\dot{X}=0, \quad \dot{Z}=0. \tag{2.8}$$

Thus the fields $\boldsymbol{X} \equiv (X,Y,Z)$ may be considered as material variables of the fluid. In the reference configuration, isosurfaces of these variables are planes. The $X$-field becomes, furthermore, the streamfunction of the dimensionless absolute vorticity in the $(x,z)$-plane, $f^{-1}\boldsymbol{\omega}_{\mathrm{p}}^{a}$,

$$f^{-1}\boldsymbol{\omega}_{\mathrm{p}}^{a} = -\boldsymbol{\jmath} \times \nabla X = (\operatorname{grad} X \times \operatorname{grad} Y)_{\mathrm{p}}. \tag{2.9}$$

$X$-isolines are therefore the vector lines of the absolute plane vorticity $\boldsymbol{\omega}_{\mathrm{p}}^{a}$, or $\boldsymbol{\omega}_{\mathrm{p}}^{a}$-lines. The above representation of the plane absolute vorticity as the plane component of the cross-product of the gradient of two conserved quantities is shown below to be relevant to potential vorticity conservation.

### 2.4. *Material vorticity*

Let us define the deformation gradient $\boldsymbol{F} \equiv \operatorname{Grad} \boldsymbol{x}$; then

$$\boldsymbol{F}^{-1} \equiv \operatorname{grad} \boldsymbol{X} = (\mathscr{X}_x+1)\boldsymbol{\imath I} + \mathscr{Y}_x \boldsymbol{\imath J} + \mathscr{Z}_x \boldsymbol{\imath K} + \boldsymbol{\jmath J} + \mathscr{X}_z \boldsymbol{kI} + \mathscr{Y}_z \boldsymbol{kJ} + (\mathscr{Z}_z+1)\boldsymbol{kK}, \tag{2.10}$$

and the Jacobian of the deformation gradient is

$$\jmath \equiv |\operatorname{grad} \boldsymbol{X}| = J_{xz}(X,Z) = 1 + \mathscr{X}_x + \mathscr{Z}_z + \mathscr{X}_x\mathscr{Z}_z - \mathscr{X}_z\mathscr{Z}_x. \tag{2.11}$$

Recall that $J \equiv |\operatorname{Grad} \boldsymbol{x}| = \jmath^{-1}$. In the $(x,z)$-plane,

$$\boldsymbol{F}_{\mathrm{p}}^{-1} \equiv \nabla \boldsymbol{X} = (\mathscr{X}_x+1)\boldsymbol{\imath I} + \mathscr{Z}_x \boldsymbol{\imath K} + \mathscr{X}_z \boldsymbol{kI} + (\mathscr{Z}_z+1)\boldsymbol{kK}, \tag{2.12}$$

and $J_\mathrm{p} \equiv |\nabla \boldsymbol{X}| = J$. Beltrami's three-dimensional material vorticity (see e.g. Casey & Naghdi 1991) is defined by $\boldsymbol{\omega}^\circ \equiv J\boldsymbol{F}^{-\mathsf{T}} \cdot \boldsymbol{\omega}$ (where $\boldsymbol{F}^{-\mathsf{T}} \equiv (\boldsymbol{F}^{\mathsf{T}})^{-1}$), and therefore

$$\boldsymbol{\omega}^{a\circ} = J\boldsymbol{F}^{-\mathsf{T}} \cdot \boldsymbol{\omega}^a = [\eta + fJ_{xz}(X, Y)]J\boldsymbol{J} + f\boldsymbol{K} = \boldsymbol{\omega}_\mathrm{n}^{a\circ} + \boldsymbol{\omega}_\mathrm{p}^{a\circ}, \tag{2.13}$$

which is not constant in time since the motion is not circulation-preserving. The plane material vorticity $\boldsymbol{\omega}_\mathrm{p}^{a\circ} \equiv J_\mathrm{p}\boldsymbol{F}_\mathrm{p}^{-\mathsf{T}} \cdot \boldsymbol{\omega}_\mathrm{p}^a = f\boldsymbol{K}$ is, however, constant in time. This is essentially equivalent to conservation of potential vorticity, because of the equivalence between Rossby–Ertel's PV and Beltrami's vorticity (Viúdez 2001). This is a special case of Cauchy's vorticity formula for circulation-preserving motion: $\boldsymbol{\omega}^\circ = \boldsymbol{\omega}^\circ(t = 0)$.

### 2.5. *Potentials*

In order to solve the system of equations it is convenient to replace $\mathscr{X}$ and $\mathscr{Z}$ (or $v$ and $\sigma$) by two new potentials $\boldsymbol{\varphi}_\mathrm{p} \equiv (\varphi, \phi)$ defined by the relations

$$[\mathrm{tr}\,(-\boldsymbol{J} \times \nabla\boldsymbol{\varphi}_\mathrm{p}), \mathrm{tr}\,(\nabla\boldsymbol{\varphi}_\mathrm{p})] = (-\varphi_z + \phi_x, \varphi_x + \phi_z) = (\mathscr{X}, \mathscr{Z}) = \boldsymbol{\mathscr{X}}_\mathrm{p} \tag{2.14}$$

which imply

$$\begin{aligned}\boldsymbol{\mathscr{A}}_\mathrm{p} &\equiv (\mathscr{A}, \mathscr{C}) \equiv \nabla^2\boldsymbol{\varphi}_\mathrm{p} = (\nabla^2\varphi, \nabla^2\phi) \\ &= (-\mathscr{X}_z + \mathscr{Z}_x, \mathscr{X}_x + \mathscr{Z}_z) = [\mathrm{tr}\,(-\boldsymbol{J} \times \nabla\boldsymbol{\mathscr{X}}_\mathrm{p}), \mathrm{tr}\,(\nabla\boldsymbol{\mathscr{X}}_\mathrm{p})].\end{aligned} \tag{2.15}$$

Note that $\varphi$ and $\phi$ have dimensions of $\mathsf{L}^2$, and therefore $\mathscr{C}$ and $\mathscr{A}$ are dimensionless, while the displacements $\mathscr{X}$ and $\mathscr{Z}$ have dimensions of $\mathsf{L}$. In order to obtain dynamical equations for $\mathscr{A}$ and $\mathscr{C}$ we need their material time derivatives which, with the help of the relations

$$\begin{aligned}\overline{\mathscr{X}}_x^{\cdot} &= -J_{xz}(w, \mathscr{X} + x), \qquad \overline{\mathscr{X}}_z^{\cdot} = J_{xz}(u, \mathscr{X} + x), \\ \overline{\mathscr{Z}}_x^{\cdot} &= -J_{xz}(w, \mathscr{Z} + z), \qquad \overline{\mathscr{Z}}_z^{\cdot} = J_{xz}(u, \mathscr{Z} + z)\end{aligned}$$

(where $\overline{\mathscr{X}}_x^{\cdot} = \mathrm{d}\mathscr{X}_x/\mathrm{d}t$, etc.) are

$$\dot{\mathscr{A}} = \eta - J_{xz}(u, \mathscr{X}) - J_{xz}(w, \mathscr{Z}), \quad \dot{\mathscr{C}} = J_{xz}(u, \mathscr{Z}) - J_{xz}(w, \mathscr{X}). \tag{2.16a, b}$$

Note that $J_{xz}(u, \mathscr{X}) + J_{xz}(w, \mathscr{Z}) = \nabla \cdot (\boldsymbol{u}_{\mathrm{p}x}\mathscr{Z} - \boldsymbol{u}_{\mathrm{p}z}\mathscr{X})$ and therefore the local rate of change of $\mathscr{A}$ can be written as $\mathscr{A}_t = \eta - \nabla \cdot (\nabla \cdot \boldsymbol{T})$, where the tensor $\boldsymbol{T} \equiv \boldsymbol{u}_\mathrm{p}(\boldsymbol{J} \times \boldsymbol{\varphi}_\mathrm{p})$.

### 2.6. *Potential vorticity*

In order to introduce a framework where PV conservation can be explicitly considered, it is necessary to adopt the rate of change of PV as an independent equation in place of one of the other independent rates of change ($\dot{\mathscr{A}}$, $\dot{\eta}$ or $\dot{\mathscr{C}}$). First let us define the dimensionless potential vorticity $\Pi$ and PV anomaly $\varpi$ densities as

$$\Pi \equiv f^{-1}\boldsymbol{\omega}^a \cdot \mathrm{grad}\, Z, \quad \varpi \equiv \Pi - 1, \tag{2.17a, b}$$

which, because of (2.9) and (2.15), can be written as

$$\Pi = (\mathrm{grad}\, X \times \mathrm{grad}\, Y) \cdot \mathrm{grad}\, Z = \mathscr{J}(X, Y, Z) = J_{xz}(X, Z) = J, \tag{2.18a}$$

$$\varpi = \nabla^2\phi + J_{xz}(\mathscr{X}, \mathscr{Z}). \tag{2.18b}$$

The PV density can be written as the three-dimensional Jacobian of three conserved quantities (or equivalently, as the two-dimensional Jacobian of two conserved quantities) because of the representation of the plane absolute vorticity (2.9) and the independence of $Z$ from $y$. When these quantities are adopted as material coordinates the demonstration of PV conservation is immediate since it becomes the Eulerian

continuity equation (which is an identity) $\mathrm{d}\jmath/\mathrm{d}t + \jmath \operatorname{div} \boldsymbol{u} = 0$ (see e.g. Truesdell 1954, § 24). In this case the amount of PV in a volume bounded by isosurfaces of $X$, $Y$ and $Z$, can be interpreted as the volume *in the material space* (which is always constant since particles are fixed) bounded by these isosurfaces. Since we are considering isochoric motion the volume in the spatial representation of that fluid element cannot change ($\operatorname{div} \boldsymbol{u} = 0$), and thus the PV density in this case obeys the same type of conservation law ($\dot{\varpi} = 0$) as for specific PV ($\mathrm{d}(\alpha\varpi)/\mathrm{d}t = 0$, where $\alpha \equiv \rho^{-1}$ is the specific volume). Volume conservation is however not required for the conservation of PV. An easy way of deriving the PV density using the material coordinates $\boldsymbol{X}$ is to employ the material expression of PV density (see e.g. Viúdez 1999) which in this case is $\Pi = f^{-1}\jmath\,\boldsymbol{\omega}^{\circ} \cdot \operatorname{Grad} Z = f^{-1}\jmath\,\boldsymbol{\omega}^{\circ} \cdot \boldsymbol{K} = \jmath$. In terms of PV the reference configuration is characterized by having zero PV density anomaly $\varpi = 0$ (or constant PV density $\Pi = 1$). The amount of PV (an extensive quantity) within a fluid element bounded by the isosurfaces $X = X_1, X_2$, $Y = Y_1, Y_2$, and $Z = Z_1, Z_2$, is simply

$$Q \equiv \int_v \Pi \,\mathrm{d}v = \int_v \jmath \,\mathrm{d}v = \int_V \mathrm{d}V = (X_2 - X_1)(Y_2 - Y_1)(Z_2 - Z_1). \tag{2.19}$$

Thus spatial variations in the PV field show up as differences in the area (in the $x, z$-plane) enclosed by contours (with a constant contour interval) of $X$ and $Z$. We provide a simple example of this in § 4.

### 2.7. *The final set of equations*

Given the relationship $\varpi = \mathscr{C} + J_{xz}(\mathscr{X}, \mathscr{Z})$ in (2.18*b*), a natural choice is to use the rate of change of $\varpi$ instead of the rate of change of $\mathscr{C}$. Thus, the variables ($u$, $v$, $w$, and $\sigma$) are replaced by the potentials $\varphi$ and $\phi$, streamfunction $\psi$, and dimensionless potential vorticity anomaly $\varpi$, satisfying three prognostic and one diagnostic equations:

$$\eta_t = (f^2 - N^2)\phi_{xz} - N^2\varphi_{xx} - f^2\varphi_{zz} - \boldsymbol{\nabla} \cdot (\boldsymbol{u}\eta), \tag{2.20a}$$

$$\mathscr{A}_t = \eta - \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \cdot \boldsymbol{T}), \tag{2.20b}$$

$$\dot{\varpi} = 0, \tag{2.20c}$$

$$\varpi = \nabla^2\phi + J_{xz}(\mathscr{X}, \mathscr{Z}), \tag{2.20d}$$

with $\boldsymbol{T}$, $\boldsymbol{u}$, $\eta$, $\mathscr{A}$, $\mathscr{X}$, and $\mathscr{Z}$, written in terms of $\varphi$, $\psi$, and $\phi$. The elliptic-hyperbolic nature of (2.20*d*) and its relation to inertial and convective stability are described below.

### 2.8. *Stability and the Monge–Ampère equation* (2.20*d*).

For given $\varphi$ and $\varpi$ equation (2.20*d*) can be written as

$$(\phi_{xx}\phi_{zz} - \phi_{xz}^2) + (1 + \varphi_{xz})\phi_{xx} + (1 - \varphi_{xz})\phi_{zz} + (\varphi_{zz} - \varphi_{xx})\phi_{xz} + (\varphi_{xx}\varphi_{zz} - \varphi_{xz}^2 - \varpi) = 0, \tag{2.21}$$

which is a particular case of the Monge–Ampère equation $E(\phi_{xx}\phi_{zz} - \phi_{xz}^2) + A\phi_{xx} + 2B\phi_{xz} + C\phi_{zz} + D = 0$, where the coefficients $A$, $B$, $C$, $D$, and $E$, are continuous functions of $x$ and $z$, and in our case

$$E = 1, \quad A = 1 + \varphi_{xz}, \quad B = (\varphi_{zz} - \varphi_{xx})/2, \tag{2.22a,b,c}$$

$$C = 1 - \varphi_{xz}, \quad D = \varphi_{xx}\varphi_{zz} - \varphi_{xz}^2 - \varpi. \tag{2.22d,e}$$

Rellich's Theorem (see e.g. Courant & Hilbert 1962, p. 324) states that if $R \equiv AC - B^2 - DE > 0$ there exists at most two solutions to the Monge–Ampère equation

which assume the same boundary values. In the case (2.21) we have

$$R = 1 + \varpi - (\mathscr{A}/2)^2. \tag{2.23}$$

Furthermore, since $R$ can also be written as

$$R = (\phi_{xx} + 1 - \varphi_{xz})(\phi_{zz} + 1 + \varphi_{xz}) - [\phi_{xz} - (\varphi_{zz} - \varphi_{xx})/2]^2, \tag{2.24}$$

one of the two possible solutions (with $R > 0$) must satisfy

$$\phi_{xx} + 1 - \varphi_{xz} > 0 \Rightarrow \mathscr{X}_x + 1 > 0 \quad (fv_x + f^2 > 0),$$
$$\phi_{zz} + 1 + \varphi_{xz} > 0 \Rightarrow \mathscr{Z}_z + 1 > 0 \quad (\sigma_z + N^2 > 0),$$

which is both inertially and statically stable. The other solution satisfies $\phi_{xx}+1-\varphi_{xz} < 0$ and $\phi_{zz} + 1 + \varphi_{xz} < 0$, and therefore is both inertially and statically unstable. For plane internal waves the amplitudes of $\eta$ and $\mathscr{A}$ are related by $\mathscr{A}_0 = \eta_0/\omega$, where $\omega$ is the intrinsic frequency of internal waves given by $\omega^2 = (f^2m^2 + N^2k^2)/(k^2 + m^2)$, and therefore $R > 0$ implies $\eta_0 < 2|\omega|$.

The numerical procedure to solve this Monge–Ampère equation and the other equations in (2.20) is described in the next section.

## 3. The numerical algorithm

The numerical code closely parallels that developed for the two-dimensional shallow-water equations (Dritschel *et al.*, hereafter referred to as DPM). The material conservation of potential vorticity (2.20*c*) is dealt with explicitly using 'contour advection' (Dritschel & Ambaum 1997, hereafter referred to as DA), in which the fluid particles $\boldsymbol{X}_\varpi$ representing contours (isolevels) of $\varpi$ are simply advected using the local fluid velocity $\boldsymbol{u}_\mathrm{p}(\boldsymbol{x}_\mathrm{p}, t)$, where $\boldsymbol{x}_\mathrm{p} = (x, z)$. That is, one solves $\partial \boldsymbol{x}_\mathrm{p}(\boldsymbol{X}_\varpi, t)/\partial t = \boldsymbol{u}_\mathrm{p}(\boldsymbol{x}_\mathrm{p}, t)$ for all the points $\boldsymbol{X}_\varpi$ on all the $\varpi$-contours. Note that $\varpi$-contours $\boldsymbol{X}_\varpi$ for a given PV jump and the contours of $\boldsymbol{X}$ are both material lines. There is no explicit reference to the grid here, and as a result, there are no numerical stability constraints for the time integration scheme (here, we use the third-order Adams–Bashforth scheme described in DPM). The time step only affects the accuracy of advection. High accuracy is achieved as long as the product of the local strain $|\mathrm{grad}\,\boldsymbol{u}_\mathrm{p}|$ and the time step $\Delta t$ is small compared to unity.

The velocity at any point $\boldsymbol{x}_\mathrm{p}$ is found by bilinear interpolation of the gridded velocity field $\boldsymbol{u}_\mathrm{p}$, as in DPM. The high accuracy of this interpolation is demonstrated in DA. In particular, there are no diffusive effects.

For general continuous distributions of $\varpi$, a finite number of contours are used. Between each contour, $\varpi$ is uniform, and so the contours correspond to discontinuities in $\varpi$. This discrete representation, even for a modest number of contours, can closely approximate a continuous distribution, as shown in Legras & Dritschel (1993) and in DPM. In turn, each contour is represented by a finite number of points, or nodes, whose number changes in time according to the complexity of the contour.

Contour complexity is limited by 'surgery' (Dritschel 1989), which joins together contours or breaks them apart when contours, or parts of the same contour, become closer than a prescribed distance, $\delta$. DA recommend setting $\delta$ to be a tenth of the basic grid scale. Here we choose $\delta$ to be a factor $0.2/m_g$ times the geometric mean grid scale, $(\Delta x\,\Delta z)^{1/2}$, since the grid is typically anisotropic ($\Delta z \ll \Delta x$). The parameter $m_g$ is the fine/coarse grid ratio, usually 2 or 4 (the latter being used in this work). The fine grid is used in the PV contour-to-grid conversion, then successive 1-2-1

averages are performed to generate the PV on the coarse grid used to represent the other dynamic variables. Note that, in the current application to stable wave motions, PV contours tend to oscillate relative to their initial configuration without exhibiting severe deformation, so surgery does not act. Surgery may act however when the waves become unstable (as reported in § 4.4). Other details of the point distribution and redistribution are the same as given in DA and DPM.

The system's evolution also depends on the two variables $\eta$ and $\mathscr{A}$. Their evolution is computed using a standard pseudospectral technique, wherein nonlinear products are carried out on the grid (in physical space) while derivatives and other linear operations are carried out in spectral space, using fast Fourier transforms to convert the fields from one representation to the other. The time integration is carried out using an explicit leap-frog scheme with a time step strongly satisfying the CFL constraint on numerical stability. To avoid the desynchronization of the even and odd time levels, we use the Robert–Asselin time filter, with coefficient 0.05, as in DPM.

The potential vorticity $\varpi$ couples with the evolution of $\eta$ and $\mathscr{A}$ only through the term $(N^2 - f^2)\phi_{xz}$ on the right-hand side of (2.20*a*). This term forces one to solve a nonlinear diagnostic equation (2.20*b*) for $\phi$. This is done as follows. The field of $\varpi$ is first recovered on the grid by the fast contour-to-grid conversion scheme described in DA. In fact, it is found on a grid four times finer in each direction than the basic grid, then averaged back to the original grid as described in DA (the factor of four was also used in DPM). Then, the gridded field of $\varpi$ is transformed to spectral space, where it is used in the following iteration (the gridded fields of $\varphi_x$ and $\varphi_z$ are also found in advance of the iteration). Starting from a time-extrapolated guess for $\phi$, each step of the iteration consists of (1) computing $\phi_x$ and $\phi_z$, (2) adding these to $-\varphi_z$ and $\varphi_x$, (3) computing the $x$- and $z$-derivatives of these sums, (4) computing the Jacobian term and its spectral transform, and finally (5) finding the next guess for $\phi$ from the (spectral) inversion of $\nabla^2$ on $\pi - J(\mathscr{X}, \mathscr{Z})$. This iteration converges rapidly (one or two iterations) except when contours of $X$ or $Z$ are close to overturning. Even then, the scheme may converge, but usually the flow becomes highly complex soon thereafter and convergence becomes impossible. This is further discussed in the context of specific examples in the following section.

While developing this numerical method, it became clear that some form of high-wavenumber filtering is required for the convergence of the diagnostic equation for $\phi$ and, consequently, for the numerically stable evolution of $\eta$ and $\mathscr{A}$. Moreover, it is necessary to use $x$ and $z$ grid scales consistent with the values of $f$ and $N$, specifically: $\Delta z/\Delta x = O(f/N)$. We have found that choosing $\Delta z/\Delta x = 1$ leads to the growth of grid-scale noise, principally in the iteration scheme used to find $\phi$. Filtering slows this growth but cannot stop it. Remarkably, even explicit viscous diffusion cannot stop it. The only practical solution is to choose $\Delta z/\Delta x = O(f/N)$, and apply filtering to prevent the generation of two-grid interval noise.

We have applied the filter used by Broutman *et al.* (1997), which has the form $F(k) = \exp(-Ck^r)$, where $k$ is the $x$- or $z$-wavenumber, $r = 10$, and $C$ is chosen so that $F = 10^{-14}$, or machine precision, at the maximum wavenumber $k_{\max}$. This filter multiplies all $x$- and $z$-wavenumbers used in computing spectral derivatives. Laplace's operator however uses the unfiltered wavenumbers, since only its inverse is used (e.g. in the iteration to find $\phi$ and in computing the streamfunction $\psi$ from $\eta$).

## 4. Numerical results

In this section we illustrate the numerical solutions of equations (2.20) for a range of cases involving different initial fields of vorticity, potential vorticity and $\mathscr{A}$. First, as

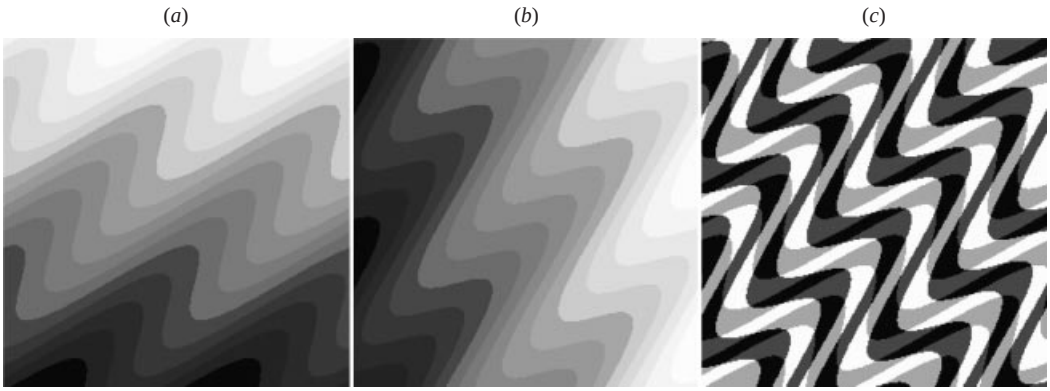


FIGURE 1. Initial contours of $Z$ ($a$), $X$ ($b$), and $Z$ and $X$ superimposed ($c$, with $Z$ alternated light grey and white, and $X$ alternated dark grey and white), for the plane wave case $k = 2$, $m = -2$. The graphical display of the contours has been discretized to show the $256 \times 256$ grid resolution.

a simple example, we consider plane waves with no PV anomaly. Next, we introduce a PV anomaly into the plane wave field and describe how this induces changes in the density and vorticity of the wave motion, specifically through the generation and propagation of vorticity wave packets and small-scale disturbances. Afterwards, we study the interaction between a vorticity wave packet and a PV anomaly and describe various refraction and reflection phenomena. Finally, we consider the generation of internal waves by PV anomalies.

In the following, all quantities are made dimensionless by choosing the Brunt–Väisälä period to be the unit of time (so that $N = 2\pi$), and the $x$-width of the domain to be $L_x = 2\pi$.

### 4.1. *Plane wave solutions*

As an illustrative example of the concepts developed in the previous section, and as a test for the numerical solution method as well, it is worth considering plane wave solutions. In this case the nonlinear terms in the equations vanish and the solutions take the form $\tilde{g} = \tilde{g}_0 \, \mathrm{e}^{\mathrm{i}(kx+mz-\omega t)}$ which, in terms of the complex vertical displacement $-\tilde{\mathscr{Z}}$, are

$$\left.\begin{aligned} &\tilde{\sigma} = N^2\tilde{\mathscr{Z}}, \quad \tilde{u} = -\mathrm{i}\omega m k^{-1}\tilde{\mathscr{Z}}, \quad \tilde{v} = -fmk^{-1}\tilde{\mathscr{Z}}, \\ &\tilde{w} = \mathrm{i}\omega\tilde{\mathscr{Z}}, \quad \tilde{\eta} = -\omega(k^2+m^2)k^{-1}\tilde{\mathscr{Z}}, \quad \tilde{\mathscr{X}} = -mk^{-1}\tilde{\mathscr{Z}}, \end{aligned}\right\} \tag{4.1}$$

where the intrinsic frequency $\omega$ satisfies

$$\omega^2 = (f^2m^2 + N^2k^2)/(k^2+m^2). \tag{4.2}$$

In this case the PV density $\Pi = 1$ everywhere ($\varpi = 0$) and the initial conditions for $\eta$ and $\mathscr{A}$ are the real parts of $\tilde{\eta}$ and $\tilde{\mathscr{A}} = -\mathrm{i}\tilde{\eta}/\omega$. We choose a plane wave with high convective and inertial instability by setting $k = 2$, $m = -2$, frequency ratio $N/f = 50$ (so that $\omega = 4.4$), and vorticity amplitude $\eta_0 = 12$, in order to produce large horizontal and vertical displacements ($\mathscr{X}_0 = \mathscr{Z}_0 = 0.68$). The time step used is $\Delta t = 0.001$ in a $256 \times 256$ grid (in this case of equal $x$ and $z$ dimensions).

The initial contours of $X$ and $Z$ are shown in figure 1. Both $X$- and $Z$-contours move without deformation ($X$-contours to the right and $Z$-contours downwards), while the intersecting points are the plane projection (in the $x,z$-plane) of the positions of fluid particles. The area in physical space bounded by constant $X$- and $Z$-contours

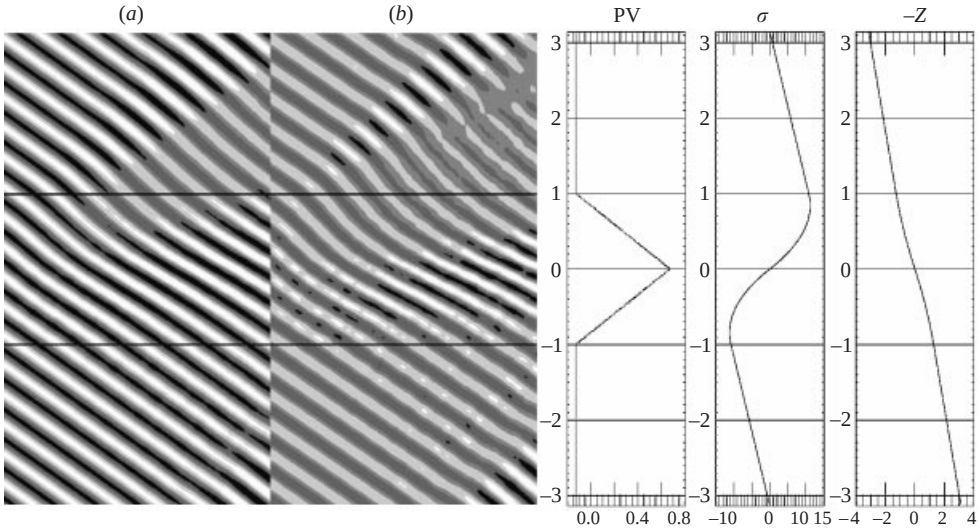


FIGURE 2. Relative vorticity $\eta(z,t)$, as a function of $z$ and time $t$ (at $x=0$), for an initial plane wave in the presence of a tent PV profile (ribbon width $h=2$) with $\Delta\varpi=0.4$ ($a$, $(\eta_{\min},\eta_{\max})=(-0.14,0.14)$) and $\Delta\varpi=0.8$ ($b$, $(\eta_{\min},\eta_{\max})=(-0.21,0.20)$). Solid fill interval $\Delta=(\eta_{\max}-\eta_{\min})/5$, with the darkest colour being the lowest value. The initial vorticity amplitude $\eta_0=0.1$, $k=1$, $m=-16$, $N/f=50$, $\Delta t=0.01$ on a $16\times256$ grid. The vertical axis is $z\in[-\pi,\pi]$, and the horizontal axis is time $t\in[0,100]$ (2 inertial periods). The initial tent PV profile $\varpi(z)$, density disturbance $\sigma(z)$, and $-Z(z)=\alpha_0 gN^{-2}[\rho(z)-\rho_0]$, for the case $\Delta\varpi=0.8$, are also shown.

(represented by the colour changes in the greyscale image) is always equal to $\Delta X\,\Delta Z$ (where $\Delta X\equiv X_2-X_1$ and $\Delta Z\equiv Z_2-Z_1$ are the contour intervals), that is, the area only depends on the difference between contour lines and not on the actual value of the contours. A fluid element may be defined by the four material points $(X_1,Z_1)$, $(X_2,Z_1)$, $(X_1,Z_2)$, and $(X_2,Z_2)$. The stretching of this fluid element in one direction (say, $x$) is exactly balanced by its compression in the other direction (say, $z$), in such a way that the area $\Delta X\,\Delta Z$ remains constant in time, which is an expression of the conservation of PV. This example is illustrative because exactly the same happens in more complicated fluid configurations, the only difference being that, when the PV anomaly $\varpi$ is different from zero, the area associated with, say, pairs of contours $(X_1,X_1+\Delta X_0)$ and $(Z_2,Z_2+\Delta Z_0)$ depends on the contours $X_1$ and $Z_1$ (that is, is not *spatially* constant), though the area of every fluid element remains constant in time.

For plane waves with zero PV anomaly, the potential $\phi=0$. Numerically, however, (2.20$d$) has to be solved, and $\phi$ becomes, though smaller than the convergence criterion, different from zero. The highly unstable flow described above was a numerical solution for at least 10 buoyancy periods ($10^4$ time steps). During that time the maximum and minimum deviations (in the entire domain) of the numerical solutions with respect to the theoretical exact values ($\sigma_0$, $u_0$, $v_0$, $w_0$, $\eta_0$) increased linearly at a rate of 0.06% per buoyancy period.

### 4.2. *Plane wave field in the presence of PV anomalies.*

In order to simulate the behaviour of internal waves in the presence of PV anomalies, an initial vorticity plane wave field $\tilde{\eta}=\tilde{\eta}_0 e^{i(kx+mz)}$ is superposed on a horizontal ribbon of anomalous PV. Forty PV layers are used in the ribbon for the examples in

this subsection. Note that, neglecting the internal wave contributions, the vertical PV anomaly profile is related to the buoyancy by $\varpi(z) = N^{-2}\sigma_z$.

The PV anomaly is prescribed by a tent profile (cf. figure 2) with a positive PV jump $\Delta\varpi \equiv \varpi_e - \varpi_b = 0.8$, where $\varpi_b$ is the background PV anomaly (outside the ribbon), and $\varpi_e$ is the PV anomaly extremum (maximum or minimum) within the PV profile. In order for $\varpi(z)$ to have a zero vertical mean, the PV anomaly outside the ribbon must be $\varpi_b = -h\Delta\varpi/(2L_z) = -0.8/(2\pi) \approx -0.13$, where $h = 2$ is the width of the ribbon and $L_z = 2\pi$ is the $z$-width, or height, of the domain. Since $\varpi_b$ is negative $\sigma(z)$ decreases with $z$ outside the ribbon. This $\varpi$ profile therefore weakens the vertical stratification outside the ribbon when $\Delta\varpi$ is positive, and strengthens it when $\Delta\varpi$ is negative. Since the stratification vanishes when $N^2\sigma_z + 1 = 0$ the initial density profile outside the ribbon is neutral when $\varpi_b = N^2\sigma_z = -1$, which implies an upper limit for the PV jump, $\Delta\varpi = 2L_z/h = 2\pi$, within the tent profile.

The evolution of the flow (figure 2) is noticeably different from the previous example for uniform PV. First, the vorticity amplitude varies spatially, and moreover, the speed of wave propagation is altered. In the case of weak to moderate amplitude PV disturbances, $\Delta\varpi = 0.4$ (figure 2*a*), wave crests initially entering the PV ribbon from the top (such waves have a negative phase speed) accelerate downwards as they encounter a region of larger vertical stratification (their vertical phase speed decreases since $\partial\eta/\partial z$ is larger in the upper part of the ribbon, $z \in [0, 1]$). Waves arriving in the lower part of the ribbon, on the contrary, decelerate as the vertical stratification weakens (their negative vertical phase speed increases since contours of equal phase become more horizontal). Consequently, wave ridges and troughs tend to spread apart at the upper end of the PV ribbon (where they are faster), and concentrate at the lower end (where they are slower). This can be verified by counting the crests at the beginning and ending time within the ribbon; there are 5 initially and 6 at the end in the case $\Delta\varpi = 0.8$.

The generation and upward propagation of wave packets having a larger vorticity amplitude take place in the ridges of the contours of equal phase, meaning that the vorticity amplitude increases when the vertical phase speed increases with time in the wave. Similarly, the vorticity amplitude decreases when the vertical phase speed decreases. The maximum vorticity amplitude $\max(|\eta(x, z, t)|) = 0.14$, is 1.4 times the initial vorticity of the plane wave. While the changes in phase speed and wavenumber are transmitted upwards in the form of a wave packet leaving the PV ribbon, small wave disturbances, related to the oscillations of the discrete PV contours, are transmitted downwards along lines of constant phase. Qualitatively similar effects, but with larger amplitudes, result for larger PV anomalies, such as shown for $\varpi = 0.8$ in figure 2(*b*). In this case, the maximum vorticity amplitude $\max(|\eta(x, z, t)|) = 0.32$ is more than three times the initial vorticity of the plane wave.

These effects may be more clearly observed by examining the difference in the vorticity fields between the cases with and without the PV ribbon (figure 4). A wave packet is generated over the PV ribbon and has a slightly different orientation (or propagation direction) than the original plane wave. This causes destructive or constructive interference, and both occur in each case (notice the darker troughs to the left of the main region of weakened amplitudes in figure 2) but particularly in case (*b*) for a stronger anomaly. The same happens in figure 3, and the observed differences between the two figures is a result of the difference in the sign of the relative angle of propagation of the wave packet. In figure 2, the wave packet moves a little to the right of the original plane wave, while in figure 3 it moves a little to the left of it.

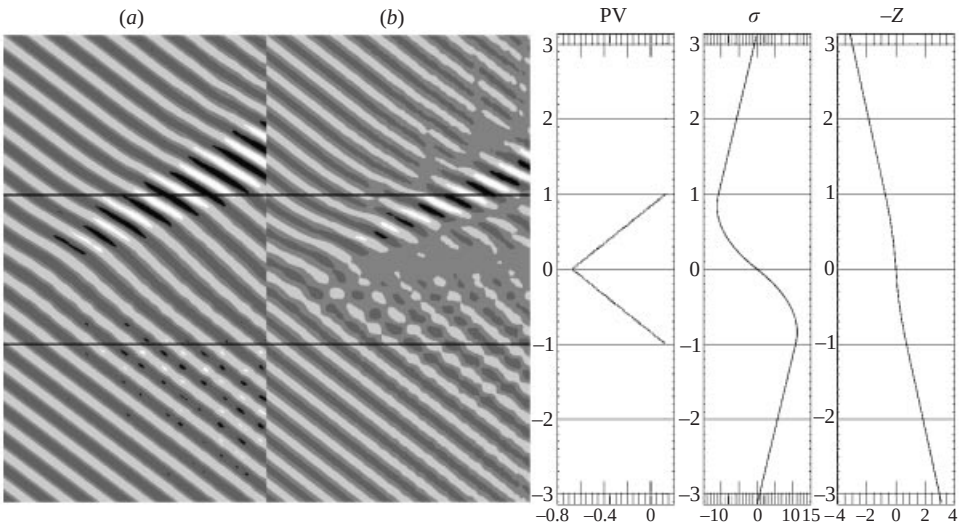


FIGURE 3. As in figure 2 but for a negative tent PV profile: $\Delta\varpi = -0.4$ ($a$, $(\eta_{\min}, \eta_{\max}) = (-0.18, 0.19)$) and $\Delta\varpi = -0.8$ ($b$, $(\eta_{\min}, \eta_{\max}) = (-0.28, 0.29)$).

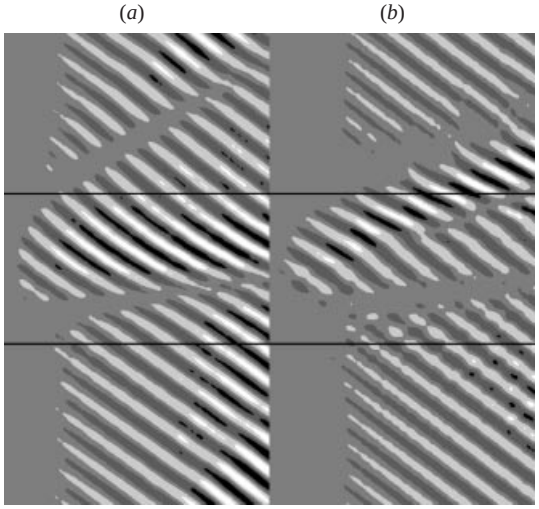


FIGURE 4. The difference of relative vorticity $\Delta\eta(z,t)$ (at $x = 0$) between the cases with and without the PV ribbon. ($a$) $(\Delta\eta_{\min}, \Delta\eta_{\max}) = (-0.25, 0.26)$ corresponds to the case $\varpi = 0.8$ (figure 2$b$) and ($b$) $(\Delta\eta_{\min}, \Delta\eta_{\max}) = (-0.33, 0.35)$ to the case $\varpi = -0.8$ (figure 3$b$). Solid fill interval $\Delta = (\Delta\eta_{\max} - \Delta\eta_{\min})/5$.

Since the vertical phase speed of the initial wave field is negative, qualitative differences may be expected when the sign of the PV profile is changed. When $\varpi < 0$, $\sigma$ decreases with height within the ribbon (figure 3) and therefore the vertical stratification is weaker there. In the case of weak to moderate amplitude PV anomalies, e.g. $\Delta\varpi = -0.4$ (figure 3$b$), the vertical phase speed of the waves entering the PV ribbon from the top decreases in magnitude while the vertical phase speed of waves arriving in the lower part of the ribbon increases in magnitude (contours of equal phase become more vertically oriented). As a consequence wave ridges and troughs tend to concentrate in the upper half of the PV ribbon and spread apart in the lower

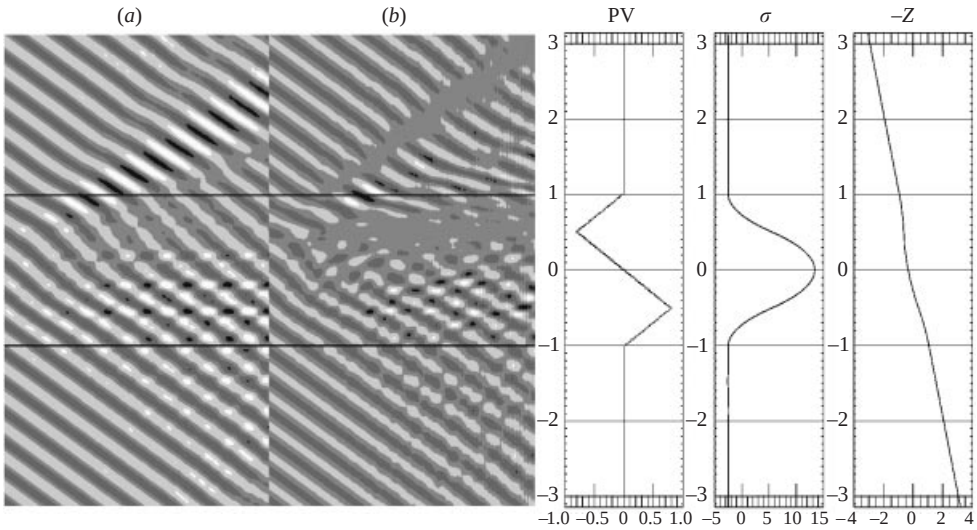


FIGURE 5. As in figure 2 but for a PV zigzag profile with $\Delta\varpi = 0.4$ ($a$, $(\eta_{\min}, \eta_{\max}) = (-0.23, 0.21)$); and $\Delta\varpi = 0.8$ ($b$, $(\eta_{\min}, \eta_{\max}) = (-0.32, 0.32)$).

half. The maximum vorticity amplitude $\max(|\eta(x,z,t)|) = 0.21$ is two times the initial vorticity of the plane wave (a larger change in amplitude than was found for a positive PV anomaly of the same magnitude). As in the case of positive PV anomalies, wave packets emerge from the upper and lower edges of the ribbon, and larger amplitude PV anomalies accentuate these effects, cf. the case $\Delta\varpi = -0.8$ shown on figure 3($b$). In this case, the maximum vorticity amplitude ($\max(|\eta(x,z,t)|) = 0.41$) becomes four times larger than the initial vorticity of the plane wave.

Other PV profiles considered produce similar results. For instance, see figures 5 and 6 for the behaviour of plane waves in the presence of a zigzag, or double-tent PV profile. The zigzag profile has zero mean PV inside the ribbon, and therefore $\varpi_{\rm b} = 0$. Though the $\sigma(z)$-field in the zigzag profile looks simpler than in the PV tent profile, its effects on the wave field are more complicated. This is because the density anomaly both increases and decreases inside the ribbon. One still sees the generation and upward propagation of wave packets along with the downward propagation of smaller-scale disturbances. And, again, these phenomena are more pronounced with increasing PV anomaly; for instance in the case $\varpi = \pm 0.8$ (shown on figures 5$b$ and 6$b$), the maximum vorticity amplitudes $\max(|\eta(x,z,t)|)$ are 0.32 and 0.39, respectively, which are about three and four times the initial vorticity of the plane wave.

### 4.3. *Wave packet and PV interaction*

The evolution of a vertically confined wave packet is simulated by imposing the initial field

$$\eta = \eta_0 \sin(kx + mz) \exp(-z^2/L^2),$$

and the $\mathscr{A}$ field whose spectral components satisfy the plane wave relations (4.1) for each spectral component of $\eta$. Here $L$ is the vertical length scale of the wave packet. A vertically and horizontally confined wave packet is simulated in the same way except $\eta$ above is multiplied by an additional factor of $\exp(-x^2/L^2)$. The initial wave packet does not contain any PV.

In the case of no PV anomaly (figure 7$a$) the wave packet moves upward (the

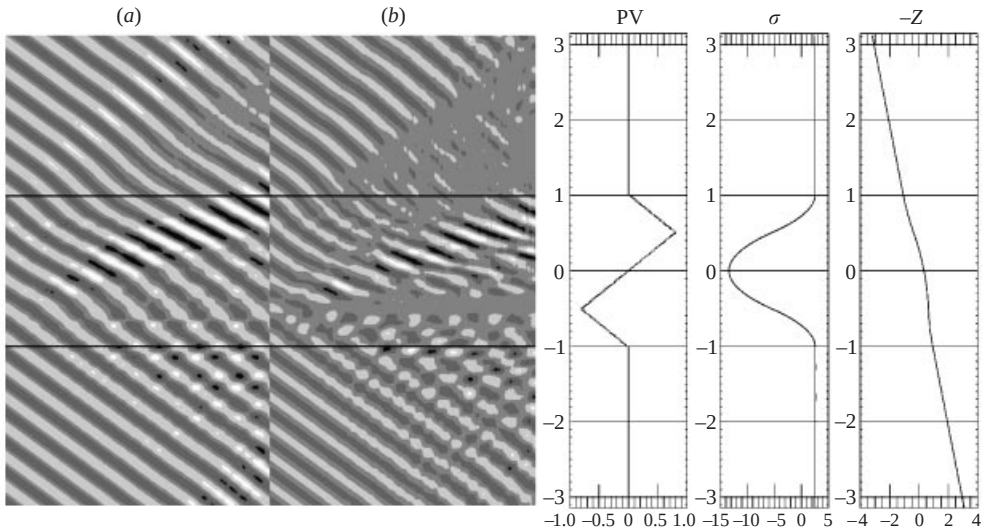


FIGURE 6. As in figure 5 but for an inverse zigzag PV profile: (*a*) $(\eta_{\min}, \eta_{\max}) = (-0.21, 0.21)$; and (*b*) $(\eta_{\min}, \eta_{\max}) = (-0.35, 0.39)$.

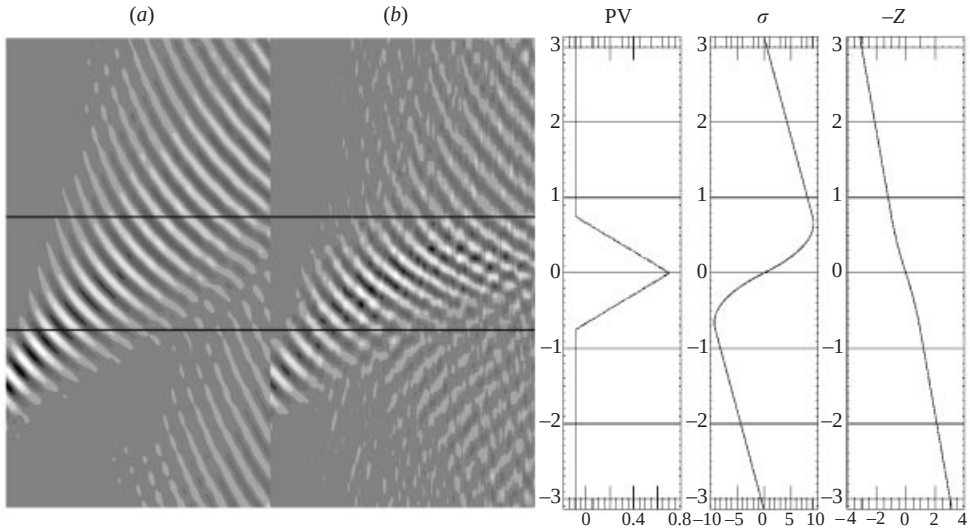


FIGURE 7. Relative vorticity $\eta(z,t)$, showing the evolution of a wave packet (initial amplitude $\eta = 0.02$, length scale $L = 0.4$, $k = 1$, $m = -16$ and $N/f = 50$) with no PV anomaly (*a*, $(\eta_{\min}, \eta_{\max}) = (-0.021, 0.021)$), and in the presence of a tent PV ribbon (*b*, $(\eta_{\min}, \eta_{\max}) = (-0.030, 0.031)$), ribbon width $h = 1.5$, $\Delta\varpi = 0.8$). Note that the breaking amplitude, as deduced from (4.1), is here $\eta_c = 2.72$. Solid fill interval $\Delta = (\eta_{\max} - \eta_{\min})/9$. Integration time is 150 buoyancy periods (3 inertial periods). The grid size is $16 \times 256$. The initial profiles of $\varpi(z)$, $\sigma(z)$, and $-Z(z)$ are shown on the right.

phases move downward) without varying its direction, and at the same time it spreads, decreasing its amplitude and increasing its length. The encounter of the wave packet with a region of anomalous PV can however substantially modify this simple spreading motion. An upward propagating wave packet encountering a positive PV anomaly ribbon (figure 7*b*) exhibits both reflection and refraction. In the positive

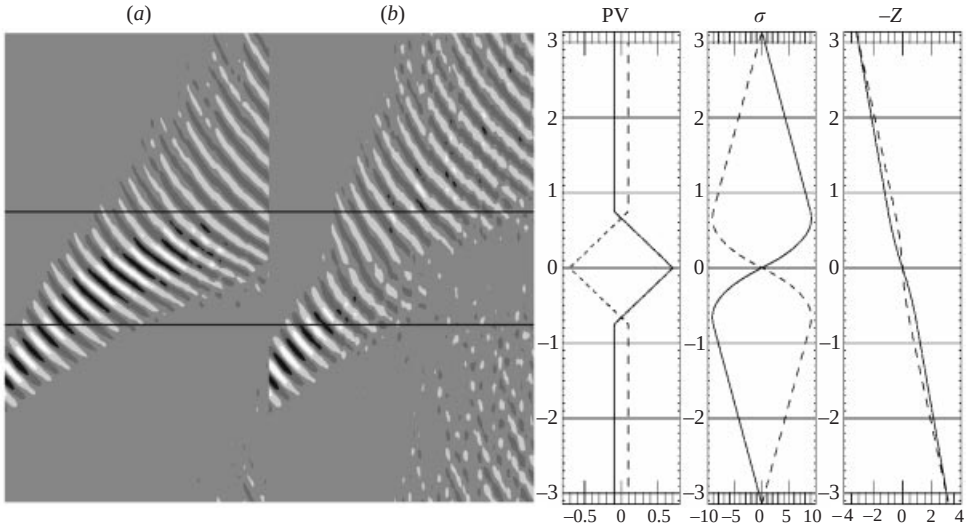


FIGURE 8. As in figure 7 but for a wave packet with initial amplitude $\eta = 0.4$ interacting with a positive PV ribbon with $\Delta\varpi = 0.8$ ($a$, $(\eta_{\min}, \eta_{\max}) = (-0.43, 0.44)$), and a negative PV ribbon with $\Delta\varpi = -0.8$ ($b$, $(\eta_{\min}, \eta_{\max}) = (-0.44, 0.44)$). The initial profiles of $\varpi(z)$, $\sigma(z)$, and $-Z(z)$ for the positive (solid line) and negative (dashed line) PV ribbons are shown on the right.

PV case the wave packet increases its magnitude and is decelerated inside the PV ribbon, where the vertical phase speed decreases (contours become more horizontally oriented) and the vertical wavenumber increases. When the wave packet encounters the PV ribbon, a vorticity wave packet is generated and propagates downwards (reflection).

While deceleration of the wave packet also occurs for larger vorticity amplitudes ($\eta = 0.4$) and large positive PV disturbances ($\varpi = 1.2$), acceleration of the wave packet occurs for negative PV anomalies. Figure 8 allows a comparison of these two phenomena. Positive PV (and therefore larger stratification in the middle of the ribbon) reduces the vertical phase speed, and negative PV (and therefore smaller stratification) increases the vertical phase speed inside the ribbon. In the positive PV case the wave amplitude is larger (compared to the evolution of the PV-free wave packet in figure 7) and the wave packet decelerates. In the negative PV case the opposite happens, the packet amplitude decreases, and the wave packet accelerates. The wave packet leaves the PV ribbon sooner when the PV anomaly is negative than when it is positive. The patchy contour field at the bottom in the negative PV case (figure 8) is a consequence of the interaction between the downward reflected wave packet and the accelerated upward propagating wave packet coming from the top boundary. This interaction is not seen in the positive PV case because the wave packet has been decelerated, and therefore it is delayed. In the case of larger vorticity amplitudes there is also reflection, but it is more than an order of magnitude smaller than the amplitude of the primary wave packet and cannot be seen in figure 8. In the weak vorticity case (figure 7) the PV ribbon acts as a larger obstacle to the incident wave than in the large vorticity case (figure 8). In the second case, the wave appears to tunnel through with relatively minor effects.

A quantitative measure of this reflection, and also the transmission, of the wave packet is provided by the average reflection and transmission coefficient ratios $\bar{r}_\eta^R$ and

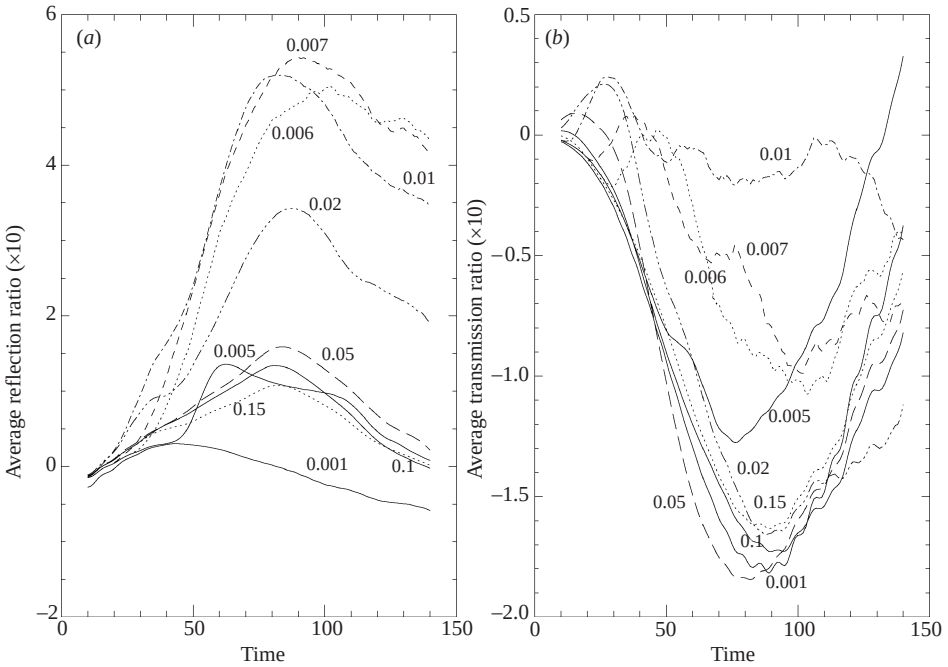


FIGURE 9. (*a*) The average reflection ratio $\bar{r}_\eta^R \times 10$, and (*b*) the average transmission ratio $\bar{r}_\eta^T \times 10$, as defined by (4.3), for the wave packet and PV ribbon described in figure 7 (with $\Delta\varpi = 0.8$) and for the initial vorticity amplitudes $\eta$ printed on the plots. The window for the running time average $w_t = 21$.

$\bar{r}_\eta^T$. The average (normalized) reflection coefficient ratio is defined by

$$\bar{r}_\eta^R(t) = \frac{1}{\bar{\eta}_0} \sum_{t'=-w_t/2}^{w_t/2} \langle\langle |\eta_\varpi(x,z,t+t')| - |\eta_0(x,z,t+t')| \rangle\rangle_{z_0}, \tag{4.3}$$

where $\eta_\varpi$ is the vorticity field for the case of a propagating wave packet (with initial maximum amplitude $\eta$) in the presence of a PV ribbon with amplitude $\Delta\varpi$, and $\eta_0$ is the vorticity field for the case of the corresponding freely propagating wave packet ($\Delta\varpi = 0$). The double average $\langle\langle\ \rangle\rangle_{z_0}$ is taken over the complete $x$-domain, and over a $z$-interval centred at the initial depth of the wave packet ($z_0 \approx -1.3$) with a vertical width of 1.2 (that is, just below the PV ribbon), while $w_t$ is the window for the running time average, and $\bar{\eta}_0 \equiv \langle\langle \eta_{\varpi(x,z,0)} \rangle\rangle_{z_0}$ is the initial average vorticity amplitude (over the same box size) of the incident wave packet, so that $\bar{r}_\eta^R(t \sim 0) \sim 1$. The average transmission coefficient ratio $\bar{r}_\eta^T$ is defined in the same way but the averaging domain is taken just above the PV ribbon ($z_0 \approx 1.3$).

The coefficients $\bar{r}_\eta^R(t)$ and $\bar{r}_\eta^T(t)$ for different values of $\eta$ are shown in figure 9. Relative reflection is small for vorticity amplitudes smaller than 0.005, it increases for increasing vorticity amplitudes, reaching a maximum $\bar{r}_\eta^R \approx 0.6$ for $\eta = 0.007$, and decreases for higher amplitudes, consistent with the images displayed in figures 7 and 8 (for $\eta = 0.02$ and $\eta = 0.4$, respectively). The time dependence of $\bar{r}_\eta^R$ shows the formation and downward propagation of the reflected wave packet, which leaves the average region after $t = 90$. For a positive PV anomaly ($\Delta\varpi = 0.8$) the time dependence of the relative transmission ratio $\bar{r}_\eta^T$ (decreasing then increasing with time)

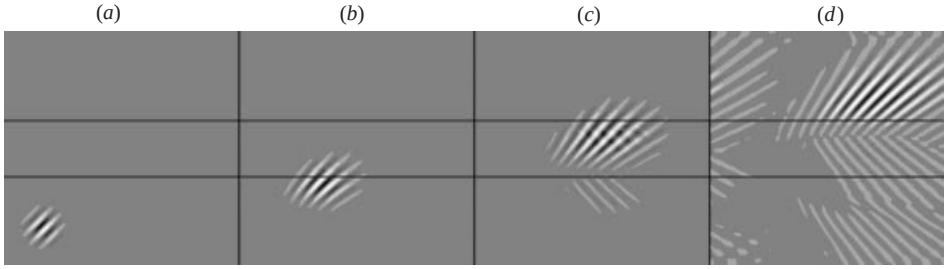


FIGURE 10. Time evolution ((*a–d*) $t = 0$, 9, 18, and 27 buoyancy periods) of the relative vorticity $\eta(x, z)$, for a horizontally and vertically confined wave packet (initial amplitude $\eta = 0.02$, $k = -l = 16$, wave packet length scale $L = 0.4$ in both directions) in the presence of a zigzag PV ribbon (ribbon width $h = 1.5$, $\Delta\varpi = 0.1$, and number of contours $n_c = 80$). Extreme values are $(\eta_{\min}, \eta_{\max}) = (\mp 0.019, \mp 0.013, \mp 0.012, \mp 0.009)$. Solid fill interval $\Delta = (\eta_{\max} - \eta_{\min})/9$, and grid size is $128 \times 128$.

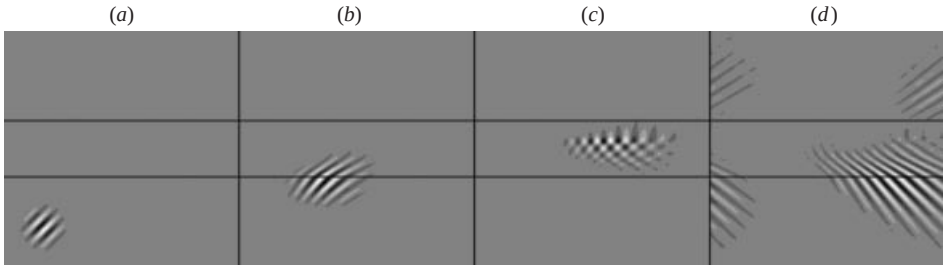


FIGURE 11. As in figure 10 but with an initial wave amplitude $\eta = 0.6$, $\Delta\varpi = 0.8$, and number of contours $n_c = 4$. Extreme values are $(\eta_{\min}, \eta_{\max}) = (\mp 0.6, \mp 0.4, \mp 0.5, \mp 0.2)$.

shows clearly the retardation of the wave packet relative to the freely propagating wave packet case. The wave packets having smaller relative reflection coefficients (those with initial maximum vorticity amplitudes $\eta \in (0.001, 0.005, 0.05, 0.1, 0.15)$) are the ones having larger (in absolute terms) relative transmission coefficients, while the wave packets with larger relative reflection coefficients (those with $\eta \in (0.006, 0.007, 0.01, 0.02)$) have the smallest relative transmission coefficients.

The evolution of a vertically and horizontally confined IGW packet as it encounters a PV ribbon is illustrated in figure 10 (see Sutherland 2001 for a recent discussion of the evolution of compact wave packets in a non-rotating fluid). In this case the vertical component of the group velocity is positive (upwards) and the vertical component of the phase velocity is negative (downwards). Two small-amplitude wave packets are reflected as the primary wave packet enters and leaves the PV ribbon. The amplitude of the refracted and reflected wave packet depends on the initial wave packet amplitude as well as the amplitude and shape of the PV jump. For example (see figure 11), for a larger initial wave packet amplitude and a larger zigzag PV jump with a small number of PV contours (this implies sharp vertical density gradients) the amplitude of the second reflected wave packet generated when the incident wave packet leaves the PV ribbon is much larger than the amplitude of the first reflected (not visible in the figure) and refracted wave packets. Indeed, the wave packet is nearly completely reflected.

Finally, we remark that the $\phi$-field may be a useful means of detecting wave packets in general circumstances. Figure 12 shows the evolution of a wave packet similar to that shown in the previous two figures except in the absence of a PV anomaly. Here,

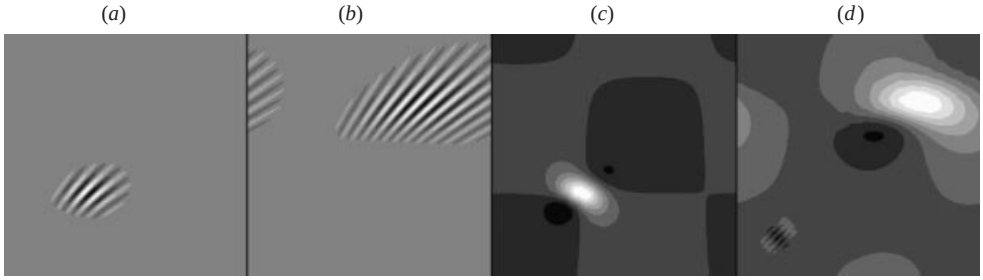


FIGURE 12. Snapshots at $t = 9$ and 27 buoyancy periods of the relative vorticity $\eta(x, z)$ ($a$ and $b$) and $\phi(x, z)$ ($c$ and $d$) for a horizontally and vertically confined wave packet having 25 times the initial vorticity amplitude ($\eta = 0.5$) as that in figure 10 but with no PV anomaly. Extreme values are $(\eta_{\min}, \eta_{\max}) = (\mp 0.32, \mp 0.15)$ and $(\phi_{\min}, \phi_{\max}) = [(-3.8, 12.5), (-1.1, 2.1)] \times 10^{-7}$. Solid fill intervals $\Delta\eta = (\eta_{\max} - \eta_{\min})/9$, and $\Delta\phi = (\phi_{\max} - \phi_{\min})/9$.

we also display the $\phi$-field in figure 12($c, d$). Note that the short-wavelength patterns in $\eta$ are not present in $\phi$, and moreover $\phi$ exhibits a coherent maximum moving with the centre of the wave packet.

### 4.4. *IGW generation*

Internal waves may also be spontaneously generated from many initial PV distributions, so long as they are not perfectly horizontal. Here we choose for simplicity a simple horizontal PV tent ribbon with a sinusoidal perturbation in $x$. Initially, the PV contours are displaced vertically by $\delta_z(x) = \delta_{z0} \sin(Kx)$. The initial relative vorticity $\eta$ and $\mathscr{A}$ fields are set to zero. In the evolution of the PV distribution ($\delta_{z0} = 0.25$, $K = 1$, $h = 0.5$, figure 13) the upper PV contours compress at both ridges ($x \in [-3\pi/4, -\pi/4]$) and troughs ($x \in [\pi/4, 3\pi/4]$) but spread apart in between, while the lower contours are modified in the opposite way. This ultimately results in a quasi-steady state (where the only unsteadiness is due to internal waves arriving from the boundaries, an artifact of the periodic conditions there). This PV-contour (or Rossby) wave motion is related to the generation of upward and downward propagating internal wave packets in both the vorticity and buoyancy fields (figure 14). The time evolution of the velocity (figure 15) shows internal wave packets in the three velocity components; here only $v$ reaches (or approaches) a steady geostrophically adjusted state in the PV ribbon. An initial PV jump half the size ($\Delta\varpi = 0.2$), and a ribbon width twice as large ($h = 1$) results in smaller wave packet amplitudes (figure 16), though the initial maximum buoyancy remains similar ($\sim 2$). In the case $\Delta\varpi = 0.5$ the minimum Rellich's parameter $R_{\min} = \min[R(x, z)] < 0$ throughout the simulation, and this appears to be associated with the development of convective instabilities at much later times (145 buoyancy periods or about 3 inertial periods, see figure 17). For smaller PV amplitudes (e.g. $\Delta\varpi = 0.2$), $R_{\min}$ remains positive, and the flow remains stable for at least 250 buoyancy periods (5 inertial periods), the length of the simulation. However, we are not aware of any theoretical reason why $R < 0$ may be associated with flow instability.

### 4.5. *Linear dynamics*

In this subsection we discuss the linear dynamics and compare nonlinear and linear results for the case of the upward propagating wave packets interacting with a PV ribbon studied in the previous sections. Linear dynamics is usually derived from the linearization of the equations of motion when the dependent variables are written in

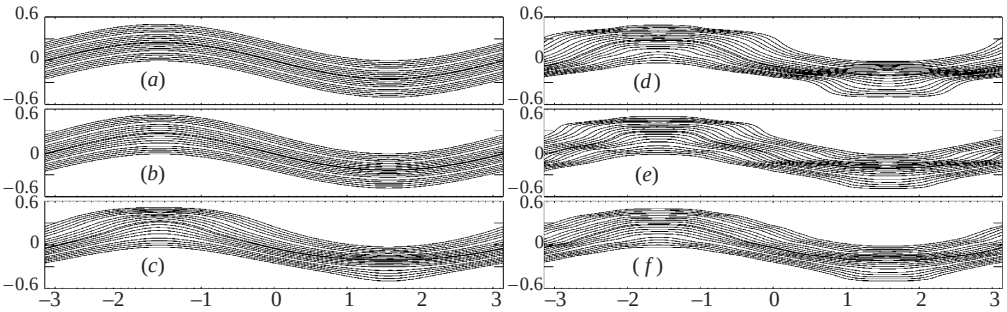


FIGURE 13. PV contours at time (*a*) $t = 0$, (*b*) 1, (*c*) 2, (*d*) 5, (*e*) 10 and (*f*) 20 buoyancy periods. Only every other contour is plotted.

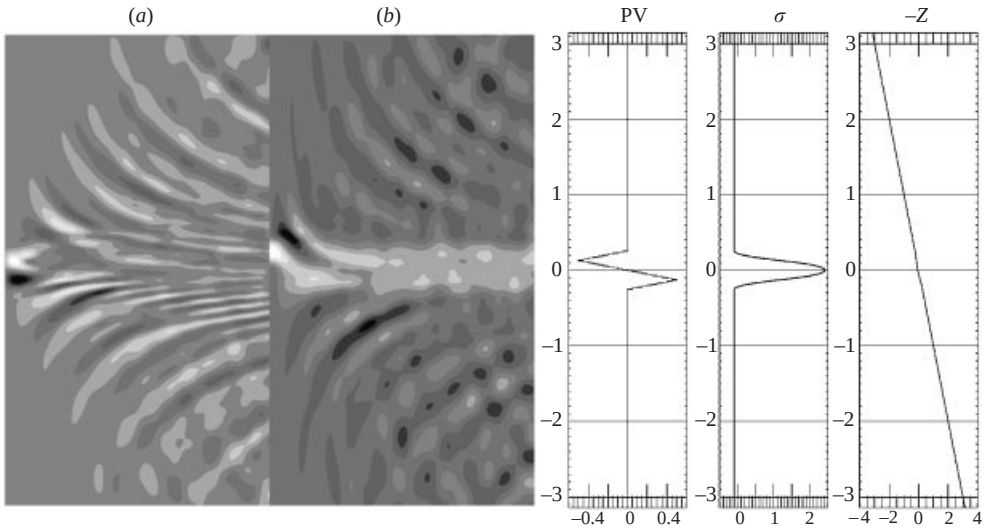


FIGURE 14. Contours showing the generation of wave packets of relative vorticity $\eta(x = -\pi, z, t)$ (*a*, $(\eta_{\min}, \eta_{\max}) = (-3.08, 3.08)$) and $\sigma(x = -\pi/2, z, t)$ (*b*, $(\sigma_{\min}, \sigma_{\max}) = (-1.49, 2.42)$). Ribbon width $h = 0.5$, $\Delta\varpi = 0.5$, $K = 1$, and $\delta_{z0} = 0.25$. The integration time is 25 buoyancy periods (0.5 inertial periods, $25 \times 10^3$ time steps with $\Delta t = 0.001$) and the grid size is $16 \times 256$. The initial profiles of $\varpi(z)$, $\sigma(z)$, and $-Z(z)$ are shown on the right.

terms of a mean quantity and a perturbation. Thus, for example, the linear version of the conservation of mass density $\dot{Z} = 0$ (recall that $Z(x, z, t) = z + \mathscr{Z}(x, z, t)$) is not $Z_t = 0$ which would imply $\mathscr{Z}_t = 0$, but $\mathscr{Z}_t + w = 0$ since $\boldsymbol{u} \cdot \nabla(z + \mathscr{Z})$ is replaced by $w$. The solution to the linear equations therefore provides a good approximation to the solution of the full nonlinear equations when the perturbations relative to a given mean distribution remain small. The linear equations for the perturbed primitive quantities are therefore

$$\mathscr{Z}_t = -w, \quad v_t = -fu, \tag{4.4a,b}$$

$$\eta_t = fv_z - N^2\mathscr{Z}_x, \quad u_x + w_z = 0. \tag{4.4c,d}$$

Now we deal with the linear version of the equations solved in this work and its relation to the above system. The approach described in this work is based on the conservation of PV ($\dot{\varpi} = 0$), a *nonlinear* function of the absolute vorticity and the

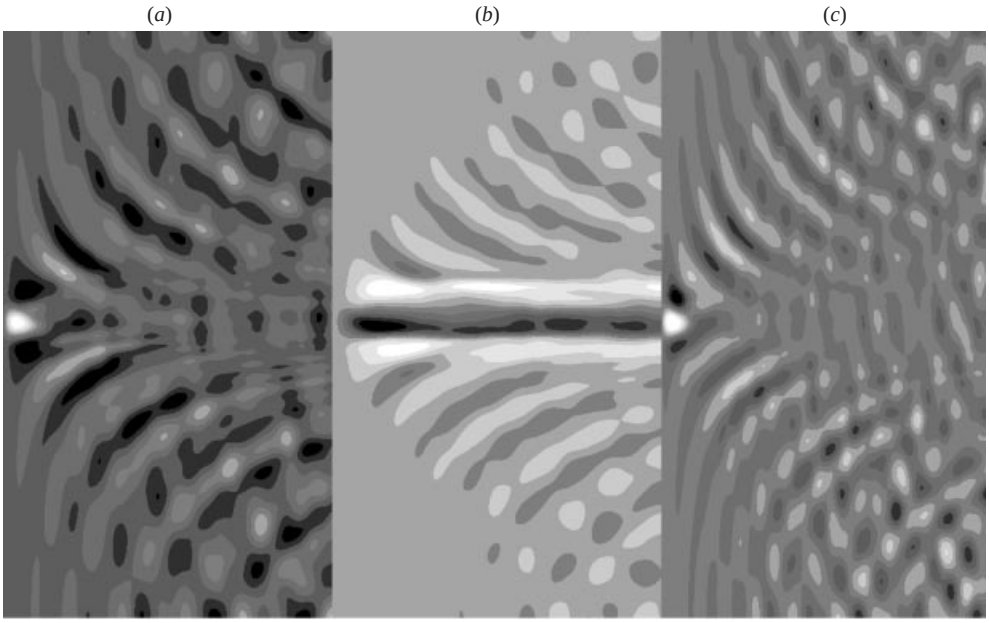


FIGURE 15. As in figure 14 but for $u$ ($a$, $(u_{\min}, u_{\max}) = (-0.27, 0.27)$), $v$ ($b$, $(v_{\min}, v_{\max}) = (-0.065, 0.065)$), and $w$ ($c$, $(w_{\min}, w_{\max}) = (-0.051, 0.065)$).

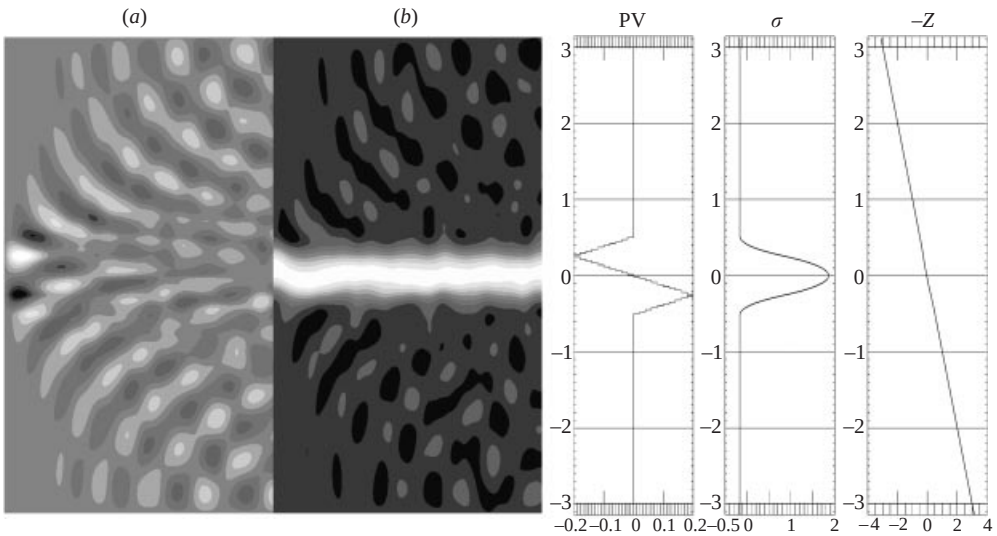


FIGURE 16. As in figure 14 but for a PV ribbon width $h = 1$, $\Delta\varpi = 0.2$, and $\delta_{z0} = 0.1$. Vorticity ($a$, $(\eta_{\min}, \eta_{\max}) = (-0.62, 0.62)$) and $\sigma$ ($b$, $(\sigma_{\min}, \sigma_{\max}) = (-0.53, 1.94)$).

density gradient that is materially conserved if the flow obeys the full nonlinear equations. However, in general and strictly speaking, $\dot{\varpi} \neq 0$ when the flow $\boldsymbol{u}$ obeys the linear equations (4.4). Note that in the linear approach a flow satisfying the linear equations (4.4) does not need to satisfy the nonlinear equations (i.e. the advective terms need not necessarily be zero) and solutions different from plane waves are therefore also possible in the linear dynamics described by (4.4). Since our approach

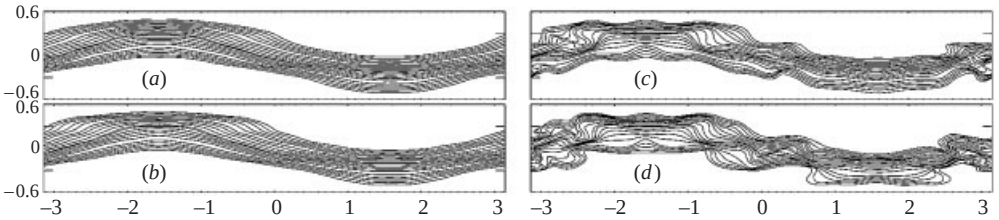


FIGURE 17. As in figure 13 but showing the PV contours at (*a*) $t = 50$, (*b*) 100, (*c*) 140 and (*d*) 145 buoyancy periods.

basically uses the conservation of PV as an independent equation it seems that it is not possible to obtain a set of linear equations fully consistent with (4.4) including the linear conservation of PV ($\varpi_t = 0$). There are several alternative solutions to circumvent this problem.

A possible solution is to define a new PV perturbation as a linear function of the displacements

$$\varpi^l \equiv \mathscr{X}_x + \mathscr{Z}_z = \mathscr{C} = \nabla^2\phi, \tag{4.5}$$

such that the total PV is $\Pi = 1 + \varpi = 1 + \varpi^l + J_{xz}(\mathscr{X}, \mathscr{Z})$. In the linear dynamics given by (4.4) the equations equivalent to (2.20) are

$$\eta_t = (f^2 - N^2)\phi_{xz} - N^2\varphi_{xx} - f^2\varphi_{zz}, \tag{4.6a}$$

$$\mathscr{A}_t = \eta, \quad \varpi^l_t = 0, \quad \varpi^l = \nabla^2\phi. \tag{4.6b,c,d}$$

Thus, the linear PV $\varpi^l$ is steady and PV contour advection is therefore not needed. Since $\phi$ is also steady, and if initially $\phi_x = 0$ (that is, if $\varpi^l_x = 0$), the equations (4.6) reduce to

$$\eta_t = -N^2\varphi_{xx} - f^2\varphi_{zz}, \quad \mathscr{A}_t = \eta, \tag{4.7a,b}$$

which imply

$$\nabla^2\varphi_{tt} + N^2\varphi_{xx} + f^2\varphi_{zz} = 0. \tag{4.8}$$

A different possibility, though obviously not fully consistent with (4.4), is to decompose the PV anomaly

$$\varpi(x,z,t) = \varpi^\circ(z) + \varpi'(x,z,t) \tag{4.9}$$

into a $z$-dependent steady vertical profile $\varpi^\circ(z)$ corresponding to the initial PV distribution $\varpi(z,0)$ (this can be done in those cases where the initial PV distribution is independent of $x$) and a perturbation PV $\varpi'(x,z,t)$ such that $\varpi'(x,z,0) = 0$. Thus the equation for the rate of change of PV may be linearized, in a similar way to the previous linearization of $\dot\rho = 0$ in (4.4) so that (4.6*c*,*d*) are replaced by

$$\varpi'_t = -w\varpi^\circ_z, \quad \varpi' = \nabla^2\phi. \tag{4.10a,b}$$

Another option related to the above equations, and the one we have adopted in our comparison between linear and nonlinear dynamics because it fits better in our approach of using a materially conserved quantity, is to linearize the final equations (2.20) but keeping the nonlinear conservation of the linear PV, so that the final (quasi-)linear equations are

$$\eta_t = (f^2 - N^2)\phi_{xz} - N^2\varphi_{xx} - f^2\varphi_{zz}, \tag{4.11a}$$

$$\mathscr{A}_t = \eta, \quad \dot\varpi^l = 0, \quad \varpi^l = \nabla^2\phi. \tag{4.11b,c,d}$$

A comparison of the linear and nonlinear dynamics is presented in figure 18 for the case of an upward propagating wave packet interacting with a horizontal PV ribbon. In the case of no PV anomaly the free propagation and spreading of the wave packet is very similar for both the linear dynamics (figure 18*a*) and nonlinear dynamics (figure 18*d*). The maximum differences between both fields (figure 18*g*) are very small (0.005) and the largest difference values seem to be related to differences due to the initialization. For $\varpi \neq 0$ the results are different however. In the case $\varpi > 0$, the linear behaviour of the wave packet (figure 18*b*) exhibits the typical retardation effect inside the PV ribbon also observed in the previous subsections for the nonlinear case; however the nonlinear dynamics significatively increases this retardation (figure 18*e*) so that the wave packet tends to remain longer inside the PV ribbon. The difference between both fields (figure 18*h*) shows this increased retardation effect, with maximum values of 0.6 along a horizontal layer inside the PV ribbon. In the case $\varpi < 0$ the nonlinear dynamics tends instead to accelerate the wave packet upwards more rapidly than the linear dynamics. The linear behaviour of the wave packet (figure 18*c*) also exhibits the typical acceleration effect inside the PV ribbon already observed in the previous subsections for the nonlinear case; however, the nonlinear dynamics significatively increases this acceleration (figure 18*f*) so that the wave packet crosses the PV ribbon faster. The difference between both fields (figure 18*i*) shows again this increased acceleration effect, with maximum values of 0.5 in a layer more vertically oriented than in the linear case located above the PV ribbon. Thus, we may conclude that though the linear dynamics reproduces qualitatively both the retardation and acceleration of a wave packet through a PV ribbon, it fails to capture these effects quantitatively, both in terms of wave phases and amplitudes.

## 5. The three-dimensional generalization

In this section we introduce the three-dimensional generalization of the two-dimensional theory given in § 2. Let $\nabla$, in this section only, denote the three-dimensional gradient operator. To begin, note that the vector $\mathscr{A}_{\mathrm{p}} \equiv \nabla_h^2 \boldsymbol{\varphi}_{\mathrm{p}}$ is just the plane component of the three-dimensional vector

$$\boldsymbol{\mathscr{A}} = (\mathscr{A}, \mathscr{B}, \mathscr{C}) \equiv \nabla^2 \boldsymbol{\varphi} = (\nabla^2\varphi, \nabla^2\psi, \nabla^2\phi) = \boldsymbol{\omega} f^{-1} + \nabla \mathscr{Z}. \tag{5.1}$$

This vector therefore combines the kinematic and thermodynamic properties of the fluid. Taking the divergence of (5.1), and using the identity $\nabla^2\boldsymbol{\varphi} = \nabla(\nabla\cdot\boldsymbol{\varphi}) - \nabla\times\nabla\times\boldsymbol{\varphi}$, we obtain

$$\mathscr{Z} = \nabla \cdot \boldsymbol{\varphi}, \quad \boldsymbol{u} = -f\nabla \times \boldsymbol{\varphi}, \tag{5.2a, b}$$

so that $-f\boldsymbol{\varphi}$ is the velocity potential.

The rate of change of $\boldsymbol{\mathscr{A}}$ is

$$\begin{aligned} \dot{\boldsymbol{\mathscr{A}}} &= \dot{\boldsymbol{\omega}} f^{-1} + \overline{\nabla \dot{\mathscr{Z}}} \\ &= (\boldsymbol{\omega}^a f^{-1} - \nabla \mathscr{Z}) \cdot \nabla \boldsymbol{u} + \nabla \dot{\mathscr{Z}} - \nabla \mathscr{Z} \times \boldsymbol{\omega} - f^{-1} N^2 \boldsymbol{k} \times \nabla \mathscr{Z} \\ &= f^{-1} \boldsymbol{\omega} \cdot \nabla \boldsymbol{u} - \nabla \boldsymbol{u} \cdot \nabla \mathscr{Z} - \boldsymbol{k} \times (\boldsymbol{\omega} + f^{-1} N^2 \nabla \mathscr{Z}). \end{aligned} \tag{5.3}$$

Note that $\dot{\mathscr{Z}} = -w$. Using (5.3) the Cartesian components of $\dot{\boldsymbol{\mathscr{A}}}$ can be written as

$$\dot{\mathscr{A}} = \eta + f^{-1} N^2 \mathscr{Z}_y - f^{-1} J_{xz}(u, v) + f^{-1} J_{xy}(u, w) - J_{xz}(w, \mathscr{Z}) - J_{xy}(v, \mathscr{Z}), \tag{5.4a}$$

$$\dot{\mathscr{B}} = -\xi - f^{-1} N^2 \mathscr{Z}_x - f^{-1} J_{yz}(u, v) + f^{-1} J_{xy}(v, w) + J_{xy}(u, \mathscr{Z}) - J_{yz}(w, \mathscr{Z}), \tag{5.4b}$$

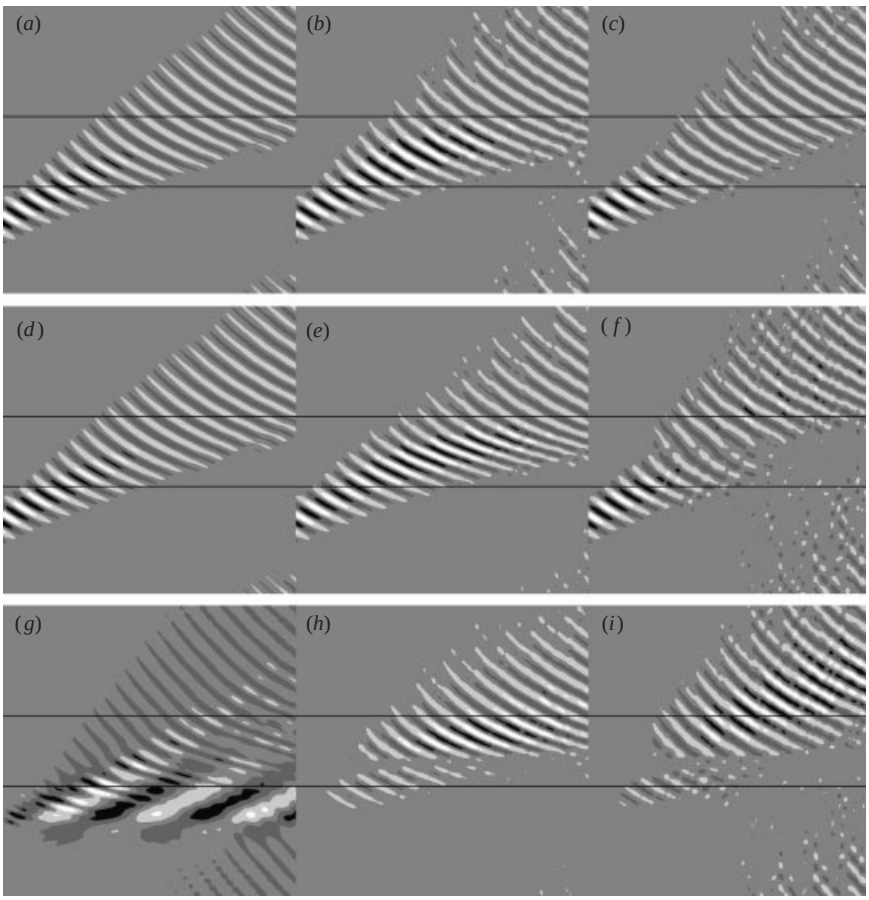


FIGURE 18. Relative vorticity $\eta(z,t)$, showing the evolution of a wave packet (initial amplitude $\eta = 0.4$ and length scale $L = 0.4$) for the linear case ($\eta^l$, parts ($a$), ($b$), and ($c$)), nonlinear case ($\eta$, parts ($d$), ($e$), and ($f$)), and vorticity difference ($\eta - \eta^l$, parts ($g$), ($h$), and ($i$)) in the presence of a tent horizontal PV ribbon of width $h = 1.5$. Cases ($a$), ($d$), and ($g$) show the evolution of a free wave packet (no PV anomaly), while $\Delta\varpi = 0.8$ for cases ($b$), ($e$), and ($h$), and $\Delta\varpi = -0.8$ for cases ($c$), ($f$), and ($i$). Solid fill interval $\Delta = (\eta_{\max} - \eta_{\min})/5$. Minimum and maximum values are $(\eta_{\min}, \eta_{\max}) = [(a),(-0.40, 0.40); (b),(-0.40, 0.41); (c),(-0.40, 0.40); (d),(-0.40, 0.40); (e),(-0.41, 0.43); (f),(-0.40, 0.40); (g),(-0.005, 0.007); (h),(-0.63, 0.57); (i),(-0.52, 0.51)]$. Integration time is 150 buoyancy periods (3 inertial periods). The grid size is $16 \times 256$.

$$\dot{\mathscr{C}} = -f^{-1}J_{xz}(w,v) + f^{-1}J_{yz}(w,u) + J_{yz}(v,\mathscr{Z}) + J_{xz}(u,\mathscr{Z}), \tag{5.4c}$$

and it is easily demonstrated that for plane motion ($\partial/\partial y = 0$) the plane components of $\dot{\mathscr{A}}$ ($\dot{\mathscr{A}}$ and $\dot{\mathscr{C}}$) correspond to (2.16$a$) and (2.16$b$), respectively, while the $y$-component ($\dot{\mathscr{B}}$) becomes the vorticity equation (2.4) (the streamfunction here $\psi$ corresponds to $-\psi f^{-1}$ in the previous sections).

The potential vorticity density (anomaly) is

$$\varpi = \Pi - 1 = \nabla^2\phi + f^{-1}J_{xz}(v,\mathscr{Z}) - f^{-1}J_{yz}(u,\mathscr{Z}) - f^{-1}J_{xy}(w,\mathscr{Z}), \tag{5.5}$$

which in the two-dimensional case reduces to (2.18). Thus, the system of equations consists of the prognostic equations for $\dot{\mathscr{A}}$, $\dot{\mathscr{B}}$, and PV conservation ($\dot{\varpi} = 0$), and a nonlinear diagnostic equation for the potential $\phi$ coming from the definition of PV. This diagnostic equation appears to be the generalization of the Monge–Ampère

equation for $\phi$ found in the two-dimensional case since it can be written similarly:

$$E[\phi_{zz}(\phi_{xx}+\phi_{yy})-\phi_{xz}^2-\phi_{yz}^2]+A(\phi_{xx}+\phi_{yy})+2B\phi_{xz}+2B'\phi_{yz}+C\phi_{zz}+D=0, \quad (5.6)$$

with coefficients

$$A=1+\Theta_z, \quad B=\nabla^2\varphi/2-\Theta_x, \quad B'=\nabla^2\psi/2-\Theta_y, \qquad (5.7a,b,c)$$

$$C=1-\Theta_z, \quad D=\Theta_x\nabla^2\varphi+\Theta_y\nabla^2\psi-|\nabla\Theta|^2-\varpi, \quad E=1, \quad (5.7d,e,f)$$

where $\Theta\equiv\varphi_x+\psi_y$. If we generalize Rellich's parameter as $R\equiv AC-B^2-B'^2-ED$, we have

$$[E(\phi_{xx}+\phi_{yy})+C](E\phi_{zz}+A)-(E\phi_{xz}-B)^2-(E\phi_{yz}-B')^2>0 \qquad (5.8)$$

as the condition for ellipticity. Thus, in this case, $[E(\phi_{xx}+\phi_{yy})+C](E\phi_{zz}+A)>0$, and the positivity of both bracketed terms implies

$$1+\phi_{xx}+\phi_{yy}-\varphi_{xz}-\varphi_{yz}>0\Rightarrow f^{-1}\zeta+1>0,$$
$$1+(\varphi_x+\psi_y+\phi_z)_z>0\Rightarrow\mathscr{Z}_z+1>0,$$

i.e. inertial and static stability. Moreover, note that we can also write

$$R=1+\varpi-(\mathscr{A}^2+\mathscr{B}^2)/4, \qquad (5.9)$$

which appears to be a simple generalization of (2.23). This similarity between the three-dimensional and two-dimensional equations suggests that their solution properties may be similar. There are thus compelling reasons to believe that the two-dimensional solution procedure discussed in this paper can be generalized readily to three dimensions. This work is currently underway.

## 6. Concluding remarks

In this work we have introduced an explicit potential-vorticity-conserving approach to modelling nonlinear internal gravity waves in geophysical fluids. The method integrates three prognostic equations and solves one diagnostic equation for the three potentials $\boldsymbol{\varphi}=(\varphi,\psi,\phi)$ and potential vorticity density $\varpi$. Numerical simulations have been presented for the two-dimensional case (in the $x,z$ plane), within a doubly periodic domain, and using a hybrid contour-advection, pseudo-spectral algorithm. Potential vorticity is held in a contour representation and its conservation equation is integrated by tracking material points along PV contours explicitly. Because conservation of PV is explicit, the approach we have introduced is an appropriate, potentially powerful tool to analyse a great number of complex IGW processes in terms of potential vorticity.

Some illustrative, idealized applications of the theory and the numerical method have been provided. These examples include the simulation of the plane waves in the presence of PV anomalies, reflection and refraction phenomena associated with the interaction between IGW packets and various PV distributions, and the production of IGW packets by PV anomalies and the subsequent geostrophic adjustment of the flow. The theoretical basis of the generalization of this approach to the three-dimensional case has also been described.

The main purpose of this study has been to introduce a new approach to modelling nonlinear IGWs and, in particular, to unravelling the role played by PV. We have

not attempted to be exhaustive here, but what we have shown is that there are many aspects of the interaction between IGWs and PV that deserve careful examination. These include IGW instability processes (which requires a three-dimensional numerical code), wave propagation in shear flows, wave–wave interactions (including inertial waves), the dependence on rotation and stratification (on the ratio of $f/N$ in particular), as well as the dissipation and production of PV in convective IGW processes. From a theoretical point of view, there remains a need to understand the connection between Rellich's parameter $R$ and IGW instabilities. We hope to report on these topics in a future work.

We would like to thank Dr Oliver Bühler for his help with wave packet initialization. Support for this research has come from the UK National Environment Research Council (grant number GR3/11899).

## REFERENCES

Broutman, D. 1984 The focusing of short internal waves by an inertial wave. *Geophys. Astrophys. Fluid Dyn.* **30**, 199–225.

Broutman, D. & Grimshaw, R. 1988 The energetics of the interaction between short small-amplitude internal waves and inertial waves. *J. Fluid Mech.* **196**, 903–106.

Broutman, D., Macaskill, C., McIntyre, M. E. & Rottman, J. 1997 On Doppler-spreading models of internal waves. *Geophys. Res. Lett.* **24**, 2813–2816.

Broutman, D. & Young, W. R. 1986 On the interaction of small-scale oceanic internal waves with near-inertial waves. *J. Fluid Mech.* **166**, 3396–3404.

Casey, J. & Naghdi, P. M. 1991 On the Lagrangian description of vorticity. *Arch. Rat. Mech. Anal.* **115**, 1–14.

Courant, R. & Hilbert, D. 1962 *Methods of Mathematical Physics, Vol. II.* Interscience.

Dritschel, D. G. 1989 Contour dynamics and contour surgery: Numerical algorithms for extended, high-resolution modelling of vortex dynamics in two-dimensional, inviscid, incompressible flows. *Comput. Phys. Rep.* **10**, 77–146.

Dritschel, D. G. & Ambaum, M. H. P. 1997 A contour-advective semi-Lagrangian algorithm for the simulation of fine-scale conservative fields. *Q. J. R. Met. Soc.* **123**, 1097–1130 (referred to herein as DA).

Dritschel, D. G., Polvani, L. M. & Mohebalhojeh, A. R. 1999 The contour-advective semi-Lagrangian algorithm for the shallow water equations. *Mon. Wea. Rev.* **127**, 1551–1565 (referred to herein as DPM).

Holloway, G. 1980 Oceanic internal waves are not weak waves. *J. Phys. Oceanogr.* **10**, 906–914.

Holloway, G. 1982 On interaction time scales of oceanic internal waves. *J. Phys. Oceanogr.* **12**, 293–296.

Legras, B. & Dritschel, D. G. 1993 A comparison of the contour surgery and pseudospectral methods. *J. Comput. Phys.* **104**, 287–302.

Marris, A. W. 1998 On pseudo-plane motions. *Arch. Rat. Mech. Anal.* **143**, 53–75.

Müller, P., Holloway, G., Henyey, F. & Pomphrey, N. 1986 Nonlinear-interactions among internal gravity-waves. *Rev. Geophys.* **24**, 493–536.

Sutherland, B. R. 1996 Internal gravity wave radiation into weakly stratified fluid. *Phys. Fluids* **8**, 430–441.

Sutherland, B. R. 1999 Propagation and reflection of internal waves. *Phys. Fluids* **11**, 1081–1090.

Sutherland, B. R. 2001 Finite-amplitude internal wavepacket dispersion and breaking. *J. Fluid Mech.* **429**, 343–380.

Thomson, W. (Lord Kelvin) 1879 On gravitational oscillations of rotating water. *Proc. R. Soc. Edinburgh* **10**, 92–100. [Reprinted in *Phil. Mag.* **10**, 109–116, 1880; *Mathematical and Physical Papers*, vol. 4, pp. 141–148, 1910.]

Thorpe, S. A. 1998 Nonlinear reflection of internal waves at a density discontinuity at the base of the mixed layer. *J. Phys. Oceanogr.* **28**, 1853–1860.

Truesdell, C. 1954 *The Kinematics of Vorticity.* Indiana University Press.

TRUESDELL, C. & TOUPIN, R. 1960 The classical field theories. In *Handbuch der Physik III/1* (ed. S. Flügge), pp. 226–793. Springer.

VIÚDEZ, A. 1999 On Ertel's potential vorticity theorem. On the impermeability theorem for potential vorticity. *J. Atmos. Sci.* **56**, 507–516.

VIÚDEZ, A. 2001 The relation between Beltrami's material vorticity and Rossby–Ertel's potential vorticity. *J. Atmos. Sci.* **58**, 2509–2517.

WINTERS, K. B. & D'ASARO, E. A. 1989 Two-dimensional instability of finite amplitude internal gravity wave packets near a critical level. *J. Geophys. Res.* **94**, 12709–12719.

WINTERS, K. B. & D'ASARO, E. A. 1994 Three-dimensional wave instability near a critical level. *J. Fluid Mech.* **272**, 255–284.