J. Fluid Mech. (1995), *vol.* 303, *pp.* 169–201

On improving a one-layer ocean model with thermodynamics

By P. RIPA

Centro de Investigación Científica y de Educación Superior de Ensenada,
Km. 107 Carretera Tijuana-Ensenada, (22800) Ensenada, B.C., México

(Received 7 February 1995 and in revised form 6 July 1995)

A popular method used to incorporate thermodynamic processes in a shallow water model (e.g. one used to study the upper layer of the ocean) is to allow for density variations in time and horizontal position, but keep all dynamical fields as depth independent. This is achieved by replacing the horizontal pressure gradient by its vertical average. These models have limitations, for instance they cannot represent the 'thermal wind' balance (between the horizontal density gradient and the vertical shear of the velocity) which dominates at low frequencies. A new model is now proposed which uses velocity and density fields varying linearly with depth, with coefficients that are functions of horizontal position and time. This model can explicitly represent the thermal wind balance, but its use is not restricted to low-frequency dynamics.

Volume, mass, buoyancy variance, energy and momentum are conserved in the new model. Furthermore, these integrals of motion have the same dependence on the dynamical fields as the exact (continuously stratified) case. The evolution of the three components of the absolute vorticity field are correctly represented. Conservation of density–potential vorticity is not fulfilled, though, owing to artificial removal of the vertical curvature of the velocity field.

The integrals of motion are used to construct a 'free energy' $\mathscr{E}_f$, which is quadratic to the lowest order in the deviation from a steady state with (at most) a uniform velocity field. $\mathscr{E}_f$ is positive definite, and therefore the free evolution of the system cannot lead to an 'explosion' of the dynamical fields. (This is not the case if the velocity shear and/or the density vertical gradient is excluded in the model, which results in a non-negative definite free energy.)

In a model with one active layer, linear waves on top of a steady state with no currents are, to a very good approximation, those of the first two vertical modes of the continuously stratified model. These are the familiar geophysical gravity and vortical waves (e.g. Poincaré, Rossby, and coastal Kelvin waves at mid-latitudes, equatorial waves, etc.).

Finally, baroclinic instability is well represented in the new model. For long perturbations (wavelengths of the order of the deformation radius of the first mode) the agreement with more precise calculations is excellent. On the other hand, the comparison with the eigenvalues of Eady's problem (which corresponds to wavelengths of the order of the deformation radius of the second mode) shows differences of the order of 40%. Nevertheless, the new model does have a high-wavenumber cutoff, even though it is constrained to linear profiles in depth and therefore cannot reproduce the exponential trapping of Eady's problem eigensolutions.

In sum, the integrals of motion, vorticity dynamics, free waves and baroclinic

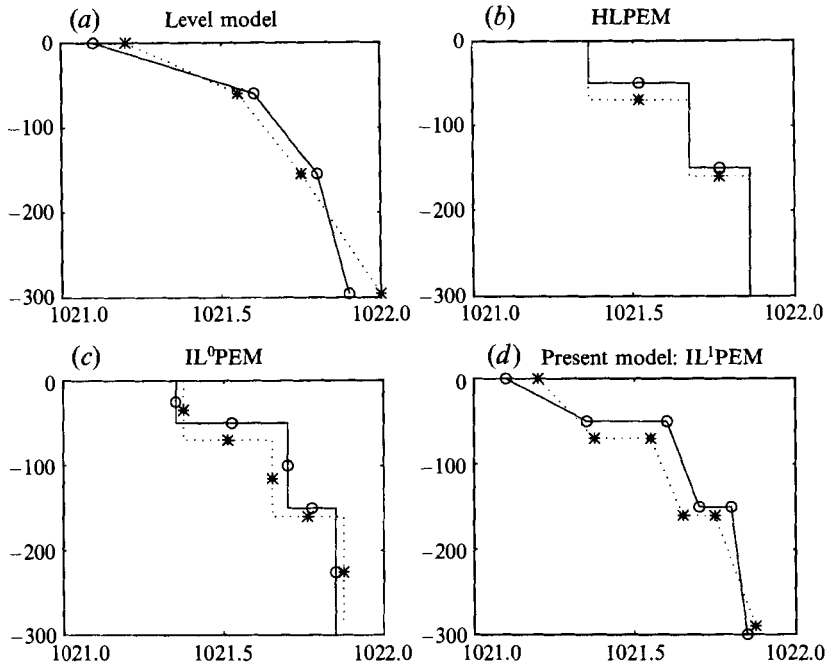

FIGURE 1. Vertical density profile, as represented by different types of models. The solid and dotted lines show the structure at two different horizontal positions and/or times, whereas the corresponding symbols (circles or asteriks) indicate the amount of information actually used in each model. In this paper only the cases with one – active – layer are presented; however, the generalization to more layers is straightforward.

instability results all give confidence in the new model. Its main novelty, however, lies in the ability to incorporate thermodynamic processes.

1. Introduction

Level ocean models work with a variable density at fixed depths (see figure 1*a*), whereas homogeneous layer primitive equation models (HLPEM) have fixed densities but variable layer thicknesses (figure 1*b*). The former can easily incorporate vertical heat and salt fluxes (through the surface, by diffusion, etc.), something that the HLPEM cannot do because of the requirement of a constant density in each layer. On the other hand, the HLPEM have generally 'more physics' than level ones (for the same amount of vertical degrees of freedom) because the equations employed in layered models are exactly like those corresponding, in the hydrostatic approximation, to a step-wise density profile, whereas level models are but a finite difference approximation. A case in point is a one-layer model, often used to study the barotropic variability (the whole water column on top of a rigid bottom) or baroclinic motions (one active layer on top of a motionless layer of heavier water). This is the simplest vertical structure possible, which is modelled correctly by a one-layer model but not by a model with one or two fixed levels.

In order to be able to represent thermodynamic processes and at the same time keep their simplicity in the vertical structure, layered models must allow for lateral density inhomogeneities; the simplest possibility corresponds to vertically averaging,

in each layer, the velocity, pressure gradient and buoyancy fields (figure 1*c*). The inhomogeneous layers primitive equations models (ILPEM) were introduced more than two decades ago by Lavoie (1972) in a study of the effects of lakes on the lower atmosphere circulation, and later, by Schopf & Cane (1983) for the study of equatorial ocean dynamics. Since then there has been an explosion of works which use this simple idea (e.g. Anderson, 1984; DeSzoeke & Richman, 1984; Anderson & McCreary 1985; McCreary & Kundu 1988; McCreary, Lee & Enfield 1989; Cherniawsky *et al.* 1990; Cherniawsky & Holloway 1991; McCreary, Fukamachi & Kundu 1991; McCreary & Yu 1992; Ripa 1993*a*; McCreary & Lu 1994; Balmaseda, Anderson & Davey 1994; Fukamachi, McCreary & Proehl 1995; and Ripa 1995*a*). Moreover, Darby & Willmott (1993), Young (1994), and Ripa (1995b) developed simplified versions of these models, valid for low-frequency dynamics. The ILPEM were generalized in Ripa (1993*a*), allowing for an arbitrary number of layers and a lower boundary condition of either reduced gravity type or that corresponding to a rigid bottom, which may include topography. The conservation laws (or lack thereof, as is the case of potential vorticity) are also fully discussed in Ripa (1993*a*).

There has not been a careful discussion about the validity of these models, even though they clearly have limitations, for instance they cannot explicitly represent the 'thermal wind' balance (between the horizontal density gradient and the vertical shear of the velocity) which dominates at low frequencies. It is important to realize that the ILPEM are indeed an approximation, because an inhomogeneous density field implies (through the hydrostatic balance) a vertical variation of the horizontal pressure gradient ∇p, which is neglected. Dempsey & Rotunno (1988) justified the model of Lavoie (1972) by postulating a Reynold stress $\langle \boldsymbol{u}'w' \rangle$ whose divergence exactly cancels the vertical variation of ∇p; however, this is no more than an *ad hoc* hypothesis (which has also been invoked in some of the publications mentioned above). An interesting possibility consists in comparing the results of these models with those of more accurate ones (as it is done, for the particular problem of long-perturbation baroclinic instability, by Fukamachi *et al.* 1995). In order to be able to do this comparison for other problems (in particular, fully nonlinear ones), it is necessary to have the 'next' model in a series of approximations with an increasing number of degrees of freedom in their vertical structure.

With this idea in mind, a new model is developed in §3, which uses velocity and density fields varying linearly with depth (figure 1*d*). The previous models and the new one will be denoted by IL^0PEM and IL^1PEM, respectively, because they are the first and second truncations of an 'exact' one, indicated by IL^∞PEM, where the index v in IL^v indicates the amount of vertical variation allowed, in the sense of the degree of polynomials in depth. The IL^∞PEM is exact within the realm of the hydrostatic approximation, and is developed in σ-coordinates in §2, to be used as a comparison with the results of those approximations.

The IL^1PEM is able to explicitly represent the 'thermal wind' balance; however, its use is not restricted to long time scales. A low-frequency approximation of the new model is presented in Ripa, (1995*c*) and may be called IL^1QGM (and therefore, that in Ripa 1995*b*, is the IL^0QGM). All these models are classified in table 1. (The symbols ϑ, h, $\boldsymbol{u}$, and ψ denote buoyancy, layer thickness, velocity and streamfunction; an overbar indicates depth-average.) The model of Young (1994) has a free parameter $f\tau$, where f is the Coriolis parameter and τ is the momentum mixing time scale. For $f\tau = 0$ Young's model coincides with the IL^0QGM (Ripa 1995*b*). For $f\tau \to \infty$, on the other hand, Young's model has an implicit representation of the velocity shear (through the thermal wind balance) and therefore it is not quite the same as the IL^1QGM.

	Independent variables	Dependent variables PE	QG	Free energy	Comments
HL	x, y, t	$h, \bar{\boldsymbol{u}}$	$\bar{\psi}$	$\mathscr{E}_f > 0$	Exact
IL^0	x, y, t	$\bar{\vartheta}, h, \bar{\boldsymbol{u}}$	$\bar{\psi}, \bar{\vartheta}$	$\mathscr{E}_f \geqslant 0$	Young ($f\tau \to 0$)
				$\mathscr{E}_f \leqslant 0$	Young ($f\tau \to \infty$)
IL^1	x, y, t	$\bar{\vartheta}, \vartheta_\sigma, h, \bar{\boldsymbol{u}}, \boldsymbol{u}_\sigma$	$\bar{\psi}, \psi_\sigma, \vartheta_\sigma$	$\mathscr{E}_f > 0$	
↓					
IL^∞	x, y, σ, t	$\vartheta, h, \boldsymbol{u}$		$\mathscr{E}_f > 0$	Exact

TABLE 1. Classification of 1–layer or $1\frac{1}{2}$–layer models

For simplicity, only one active layer is considered in the theoretical studies of each type of model. However, in realistic applications it is common to use a stack of layers of the same type (as indicated in figure 1) or even of different types (e.g. an IL^1 on top of an IL^0 to model the upper layers of the ocean).

One important test of the validity of the new model has to do with its conservation laws, more precisely on how many integrals of motion of the exact system are also conserved by the new one and how well they are represented. Energy and momentum conservation, as well as the integrals of motion related to functions of the density, are discussed in §4. The vorticity theorems have a special treatment, in §4.2, where a distinction is made between the 'σ–potential vorticity', which is conserved in the HLPEM but not in models with variable density (within each layer), and the 'density-potential vorticity', which is conserved in the exact model.

The exact model, as well as the HLPEM and the IL^0PEM have a family of integrals of motion (called 'Casimirs') which are related to the vorticity and density fields. These integrals of motion are useful in the derivation of stability theorems (e.g. McIntyre & Shepherd 1987; Shepherd 1990; Ripa 1991, 1992*a*) or instability theorems (Ripa 1992*b*). In order to find the Casimir integrals of motion it is useful to know the Hamiltonian structure of the problem; this is done for the new model in §5. Readers not interested in Hamiltonian dynamics can safely skip §5, except perhaps for table 3, where the Casimir integrals of motion of the exact model (IL^∞PEM), the new model (IL^1PEM), the HLPEM and the IL^0PEM are compared.

The integrals of motion are used in §6 to build a conserved *free energy* $\mathscr{E}_f$, quadratic to the lowest order in the deviation from a certain *reference state*. (The potential energy part of this integral of motion is the so-called *available potential energy*.) I would like to stress the importance of $\mathscr{E}_f$ being quadratic (to the lowest order) in some perturbation field.

(*a*) First, if the linear term vanishes then the quadratic one is independent of the choice of representation of the perturbation field. Consider, as an analogy, $y = f(x)$ and also $y = F(X)$, through some change of variable $x = g(X)$. It is easy to see that the second differentials of y are, in general, different, $\frac{1}{2}f''(x_0)(x-x_0)^2 \neq \frac{1}{2}F''(X_0)(X-X_0)^2$, unless the first differential vanishes identically, $f'(x_0) = 0 = F'(X_0)$. Consequently, if the perturbation is known to $O(\varepsilon)$, then $\mathscr{E}_f$ can be calculated to its lowest order, $O(\varepsilon^2)$, something which is not true for integrals whose first variation does not vanish (McIntyre & Shepherd 1987). This explains why the 'wave energy' of the Kelvin–Helmholtz instability problem can be positive definite in depth coordinates and sign indefinite in density coordinates (Ripa 1990): the 'mean flow energy' is linear in the deviation from the basic state, but it has an $O(\varepsilon^2)$ contribution whose sign is also coordinate-dependent (see Ripa 1993*c*).

	Case A Rigid bottom	Case B Reduced gravity
Layer def.	$h_0(x) \leqslant z \leqslant h_0(x) + h(x,t)$	$-h(x,t) \leqslant z \leqslant 0$
$\vartheta(x,z,t)$	$g\left[\rho(x,z,t) - \rho_{up}\right]/\rho_0$	$g\left[\rho_{down} - \rho(x,z,t)\right]/\rho_0$
BC $\sigma = -1$	$p = 0,\ w = \mathrm{D}(h_0 + h)/\mathrm{D}t$ at $z = h_0 + h$,	$p = 0,\ w = -\mathrm{D}h/\mathrm{D}t$ at $z = -h$
BC $\sigma = 1$	$w = \boldsymbol{u} \cdot \nabla h_0$ at $z = h_0$	$w = 0$ at $z = 0$

TABLE 2.

(*b*) Second, whether $\mathscr{E}_f$ is positive definite, non-negative definite or sign indefinite (see table 1) is important for the evolution of the system, as shown in §7 in a comparison of different types of models. A system with quadratic nonlinearity which has a sign-indefinite integral of motion $\mathscr{E}_f$ may be subject to the phenomenon of 'explosive resonant triad': even if the linearized problem is stable (the normal modes have real eigen-frequencies) the interaction between modes with different sign of $\mathscr{E}_f$ leads to an 'explosive' nonlinear instability (Morrison & Kotschenreuther 1990; Keuny & Morrison 1994; and Vanneste 1995).

Another test of the validity of the new model has to do with the waves it supports by linearizing the evolution equations in the deviation from a steady state. This is done in §8, where the free waves are compared with those of the exact model, and in §9, where a similar comparison is done of the baroclinic instability problem. Main conclusions are presented in §10 and some mathematical details are left for the Appendices.

2. The full model: IL$^\infty$PEM

Consider one layer of active fluid, with thickness $h(x,t)$, where x is a horizontal coordinate and t is time. The horizontal surface is either a sphere or a plane (there is no need to be more specific until the zonal momentum is considered); the vertical coordinate z is then either radial or perpendicular to the plane. There are two possibilities for the vertical structure of the model: either this active layer is on top of a rigid bottom, with topography denoted by $z = h_0(x)$, or below a horizontal 'rigid lid' at $z = 0$ (see cases A and B in table 2); throughout this section, these two cases correspond to the upper and lower signs respectively, but the notation in the bulk in this paper is chosen so that it is the same in both cases. The other boundary is a soft interface with a passive (infinitely deep) layer of constant density ρ_{up} or ρ_{down} respectively (in particular, ρ_{up} might vanish). The density $\rho(x,z,t)$ is used to define a field ϑ in such a way that it is positive in both cases (see table 2; ρ_0 is a reference density used in the Boussinesq approximation); this field will be called the 'buoyancy' even though it is such only in case B.

In addition to $\vartheta(x,z,t)$, the dynamical variables are the horizontal $\boldsymbol{u}(x,z,t)$ and vertical $w(x,z,t)$ velocity fields, and the kinematic pressure $p(x,z,t)$. The equations of motion in the *primitive equations model* (i.e. hydrostatic and with the horizontal

Coriolis force) without any forcing or dissipation are

$$\left.\begin{aligned} \mathrm{D}\vartheta/\mathrm{D}t &= 0, \\ \nabla\cdot\boldsymbol{u} + \partial_z w &= 0, \\ \mathrm{D}\boldsymbol{u}/\mathrm{D}t + f\hat{\boldsymbol{z}}\times\boldsymbol{u} + \nabla p &= 0, \\ \partial_z p &= \mp\vartheta, \end{aligned}\right\} \tag{2.1}$$

where $\mathrm{D}(.)/\mathrm{D}t := \partial_t(.) + \boldsymbol{u}\cdot\nabla(.) + w\partial_z(.)$ and f is the Coriolis parameter (∇ is the horizontal nabla operator and $\mp$ refers to cases A and B). The boundary conditions are presented in the last two rows of table 2.

For the purposes of this paper, it is convenient to use a coordinate σ, linear with z, and defined such that $\sigma = -1$ ($\sigma = 1$) corresponds to the soft (rigid) boundary, namely

$$\pm z = v(\boldsymbol{x},\sigma,t) := h_0 + \frac{1-\sigma}{2}h. \tag{2.2}$$

In Case B (lower sign), usually $h_0 = 0$ (although one might imagine a laboratory experiment with a non-horizontal top lid). The reader interested only in the ocean's upper layer may ignore all occurrences of h_0, and take the σ variable as equal to 1 at the surface of the ocean and equal to -1 at the base of the active layer.

In order to consider the dynamical fields as functions of $(\boldsymbol{x},\sigma,t)$ – instead of $(\boldsymbol{x},z,t)$ – the first step is to write down the transformation laws for the differential operators, from $(\partial_t,\nabla,\partial_z)$ to $(\tilde{\partial}_t,\tilde{\nabla},\partial_\sigma)$, where the tilde indicates derivation at constant σ instead of at constant v (i.e. constant z). When operating on quantities that are not a function of either σ or z, like h or $\overline{\phi}$ (for any ϕ; see (2.10) below), the tilde will be dropped. The transformation laws are

$$\left.\begin{aligned} \partial_t &= \tilde{\partial}_t + 2h^{-1}(\tilde{\partial}_t v)\partial_\sigma, \\ \nabla &= \tilde{\nabla} + 2h^{-1}(\tilde{\nabla} v)\partial_\sigma, \\ \partial_z &= \mp 2h^{-1}\partial_\sigma, \end{aligned}\right\} \tag{2.3}$$

since $\partial_\sigma v = -h/2$; notice that $\partial_t v = \nabla v = 0$, whereas $\partial_z v = \pm 1$, as it should. The horizontal Jacobian $[a,b]^z := \hat{\boldsymbol{z}}\cdot\nabla a\times\nabla b$ transforms as

$$[a,b]^z = -2h^{-1}\left(\partial_\sigma v\,[a,b]^\sigma + \partial_\sigma b\,[v,a]^\sigma + \partial_\sigma a\,[b,v]^\sigma\right) \tag{2.4}$$

where $[a,b]^\sigma := \hat{\boldsymbol{z}}\cdot\tilde{\nabla}a\times\tilde{\nabla}b$. For instance, $[\vartheta,\sigma]^z = 2h^{-1}[\vartheta,v]^\sigma$, a relationship that is needed in Appendix A. Finally, the substantial derivative is written as

$$\mathrm{D}(.)/\mathrm{D}t = \tilde{\partial}_t(.) + \boldsymbol{u}\cdot\tilde{\nabla}(.) + \mu\partial_\sigma(.) \tag{2.5}$$

where

$$\mu := \frac{\mathrm{D}\sigma}{\mathrm{D}t} = \frac{2}{h}\left(\frac{\mathrm{D}h_0}{\mathrm{D}t} + \frac{1-\sigma}{2}\frac{\mathrm{D}h}{\mathrm{D}t} \mp w\right). \tag{2.6}$$

Recall that h_0 and h are σ (or z) independent and therefore $\mathrm{D}(h_0,h)/\mathrm{D}t = (\partial_t + \boldsymbol{u}\cdot\nabla)(h_0,h)$.

The evolution equations in the σ-system, obtained from the set (2.1) using these transformation laws, are found to be

$$\mathrm{IL}^\infty\mathrm{PEM}: \begin{cases} \mathrm{D}\vartheta/\mathrm{D}t = 0, \\ \tilde{\partial}_t h + \tilde{\nabla}\cdot(h\boldsymbol{u}) + h\partial_\sigma\mu = 0, \\ \mathrm{D}\boldsymbol{u}/\mathrm{D}t + f\hat{\boldsymbol{z}}\times\boldsymbol{u} + \tilde{\nabla}p + \vartheta\tilde{\nabla}v = 0, \\ \partial_\sigma p = \frac{1}{2}h\vartheta, \end{cases} \tag{2.7}$$

where D/Dt must be interpreted in the sense of (2.5). The boundary conditions at the two last rows of table 2 translate into

$$p = 0 \ @ \ \sigma = -1, \quad \mu = 0 \ @ \ \sigma = \pm 1. \tag{2.8}$$

Notice that in σ-variables Cases A and B from table 1 are equivalent, i.e. there are no more double signs in (2.7).

It is better to split the incompressibility approximation (2.7, second equation) into two parts by performing an integration in σ, from -1 to 1, and using the boundary condition (2.8), which results in

$$\left.\begin{aligned} \tilde{\nabla} \cdot [h(\boldsymbol{u} - \bar{\boldsymbol{u}})] + h\partial_\sigma \mu = 0, \\ \tilde{\partial}_t h + \nabla \cdot (h\bar{\boldsymbol{u}}) = 0, \end{aligned}\right\} \tag{2.9}$$

where, throughout this paper, an overbar denotes a vertical average within the active layer $\overline{(\cdots)} = h^{-1} \int (\cdots) \mathrm{d}z$ or, in σ-coordinates,

$$\overline{(\cdots)} := \frac{1}{2} \int_{-1}^{1} (\cdots) \, \mathrm{d}\sigma. \tag{2.10}$$

3. Setting up the new model: IL^{1}PEM

The HLPEM is obtained from (2.7) and (2.8) by specifying an initial state such that ϑ = constant and $\partial_\sigma \boldsymbol{u} = 0$ (i.e. $\boldsymbol{u} = \bar{\boldsymbol{u}}$). It is easy to see that these conditions are preserved by the dynamics, and then $\mu = 0$ follows from (2.7) (i.e. σ is constant following particles), $p = (1 + \sigma)\vartheta h/2$, and the second and third equations in (2.7) reduce to the classical shallow water equations

$$\text{HLPEM} \quad \begin{cases} \mathrm{D}h/\mathrm{D}t + h\nabla \cdot \bar{\boldsymbol{u}} = 0, \\ \mathrm{D}\bar{\boldsymbol{u}}/\mathrm{D}t + f\hat{\boldsymbol{z}} \times \bar{\boldsymbol{u}} + \bar{\vartheta}\nabla h + \bar{\vartheta}\nabla h_0 = 0, \end{cases} \tag{3.1}$$

where $\mathrm{D}/\mathrm{D}t = \partial_t + \bar{\boldsymbol{u}} \cdot \nabla$. If ϑ is *not* initially uniform, then a vertical shear will in general develop, $\partial_\sigma \boldsymbol{u} \neq 0$, and a fully three-dimensional problem will have to be solved, unless some closure hypothesis is made in order to limit the amount of vertical structure allowed. The IL^{0}PEM corresponds to simply σ-averaging the equations (2.7): if ϑ is σ-independent then $p = (1 + \sigma)\bar{\vartheta}h/2$; calculating the vertical average of ∇p, it follows that

$$\text{IL}^0\text{PEM} \quad \begin{cases} \mathrm{D}\bar{\vartheta}/\mathrm{D}t = 0, \\ \mathrm{D}h/\mathrm{D}t + h\nabla \cdot \bar{\boldsymbol{u}} = 0, \\ \mathrm{D}\bar{\boldsymbol{u}}/\mathrm{D}t + f\hat{\boldsymbol{z}} \times \bar{\boldsymbol{u}} + \bar{\vartheta}\nabla h + \frac{1}{2}h\nabla\bar{\vartheta} + \bar{\vartheta}\nabla h_0 = 0 \end{cases} \tag{3.2}$$

(see for instance Ripa 1993*a*). In this paper, I will go one step further, by considering the *ansatz*

$$\left.\begin{aligned} \boldsymbol{u}(x, \sigma, t) = \bar{\boldsymbol{u}}(\boldsymbol{x}, t) + \sigma \ \boldsymbol{u}_\sigma(\boldsymbol{x}, t), \\ \vartheta(\mathbf{x}, \sigma, t) = \bar{\vartheta}(\boldsymbol{x}, t) + \sigma \ \vartheta_\sigma(\boldsymbol{x}, t). \end{aligned}\right\} \tag{3.3}$$

Thus, in the reduced gravity case, $\bar{\boldsymbol{u}} + \boldsymbol{u}_\sigma$ is the surface velocity and $\bar{\boldsymbol{u}} - \boldsymbol{u}_\sigma$ is the velocity at the base of the active layer; similarly for the buoyancies $\bar{\vartheta} + \vartheta_\sigma$ and $\bar{\vartheta} - \vartheta_\sigma$. Recall that $|\sigma| \leqslant 1$: a positive buoyancy ϑ requires $\bar{\vartheta} > |\vartheta_\sigma|$. However, ϑ_σ is related to the (instantaneous) Brunt–Väisälä frequency squared, N^2, by

$$\vartheta_\sigma(\boldsymbol{x}, t) = \tfrac{1}{2}N^2(\boldsymbol{x}, t)h(\boldsymbol{x}, t), \tag{3.4}$$

and therefore physically acceptable values correspond to $\bar{\vartheta} > \vartheta_\sigma > 0$.

In order for the linear vertical structure of (3.3) to be preserved, all occurrences of σ^2 in the equations of motion need to be eliminated; the easiest recipe is to replace them by their mean value, i.e. $\sigma^2 \mapsto 1/3$ (here, $\mapsto$ means 'replacement'). For instance, if $a = \bar{a} + \sigma a_\sigma$ and $b = \bar{b} + \sigma b_\sigma$, then $ab \mapsto \overline{ab} + \sigma (ab)_\sigma$, where $\overline{ab} = \bar{a}\bar{b} + a_\sigma b_\sigma/3$ and $(ab)_\sigma = \bar{a}b_\sigma + a_\sigma \bar{b}$. Therefore, using (3.3) in the first equation of (2.9) and the boundary condition (2.8), the σ-velocity μ is found to have a term proportional to σ^2, which is then replaced by its vertical average, i.e.

$$\mu = 2h^{-1}(1-\sigma^2)\nabla\cdot(h\boldsymbol{u}_\sigma) \mapsto \bar{\mu} = \tfrac{1}{3}h^{-1}\nabla\cdot(h\boldsymbol{u}_\sigma). \tag{3.5}$$

Similarly, using the ansatz (3.3) in the mass conservation equation, $[\tilde{\partial}_t + (\bar{\boldsymbol{u}} + \sigma\boldsymbol{u}_\sigma)\cdot\tilde{\nabla} + \bar{\mu}\partial_\sigma](\bar{\vartheta} + \sigma\vartheta_\sigma) = 0$, and separating the mean and linear parts, the resulting approximation is

$$\mathrm{D}\vartheta/\mathrm{D}t = 0 \mapsto \overline{\mathrm{D}\vartheta/\mathrm{D}t} + \sigma\ \left(\mathrm{D}\vartheta/\mathrm{D}t\right)_\sigma = 0. \tag{3.6}$$

Consequently, the equation $\mathrm{D}\vartheta/\mathrm{D}t = 0$, at all $(\boldsymbol{x}, \sigma, t)$, is approximated by $\overline{\mathrm{D}\vartheta/\mathrm{D}t} = 0$ and $\left(\mathrm{D}\vartheta/\mathrm{D}t\right)_\sigma = 0$, at all $(\boldsymbol{x}, t)$, where in general

$$\left.\begin{aligned} \overline{\mathrm{D}A/\mathrm{D}t} &= \partial_t\bar{A} + \bar{\boldsymbol{u}}\cdot\nabla\bar{A} + \tfrac{1}{3}\boldsymbol{u}_\sigma\cdot\nabla A_\sigma + \bar{\mu}A_\sigma \\ &= \partial_t\bar{A} + \bar{\boldsymbol{u}}\cdot\nabla\bar{A} + \tfrac{1}{3}h^{-1}\nabla\cdot(h\boldsymbol{u}_\sigma A_\sigma), \\ \left(\mathrm{D}A/\mathrm{D}t\right)_\sigma &= \partial_t A_\sigma + \bar{\boldsymbol{u}}\cdot\nabla A_\sigma + \boldsymbol{u}_\sigma\cdot\nabla\bar{A}. \end{aligned}\right\} \tag{3.7}$$

Finally, the pressure field is obtained using the ansatz (3.3) in the hydrostatic balance, the fourth equation of (2.7), and the boundary condition (2.8), which yield $p = (1+\sigma)h\bar{\vartheta}/2 - (1-\sigma^2)h\vartheta_\sigma/4$. Calculating ∇p and separating its mean and linear part in σ gives $\nabla p \mapsto \overline{\nabla p} + \sigma\ (\nabla p)_\sigma$, where

$$\left.\begin{aligned} \overline{\nabla p} &= (\bar{\vartheta} - \tfrac{1}{3}\vartheta_\sigma)\nabla h + \tfrac{1}{2}h\nabla(\bar{\vartheta} - \tfrac{1}{3}\vartheta_\sigma) + \bar{\vartheta}\nabla h_0, \\ (\nabla p)_\sigma &= \tfrac{1}{2}\vartheta_\sigma\nabla h + \tfrac{1}{2}h\nabla\bar{\vartheta} + \vartheta_\sigma\nabla h_0. \end{aligned}\right\} \tag{3.8}$$

Putting all this together the evolution equations for the new model are found to be

$$\mathrm{IL}^1\mathrm{PEM}: \begin{cases} \overline{\mathrm{D}\vartheta/\mathrm{D}t} = 0, \\ \left(\mathrm{D}\vartheta/\mathrm{D}t\right)_\sigma = 0, \\ \partial_t h + \nabla\cdot(h\bar{\boldsymbol{u}}) = 0, \\ \overline{\mathrm{D}\boldsymbol{u}/\mathrm{D}t} + f\hat{\boldsymbol{z}}\times\bar{\boldsymbol{u}} + \overline{\nabla p} = 0, \\ \left(\mathrm{D}\boldsymbol{u}/\mathrm{D}t\right)_\sigma + f\hat{\boldsymbol{z}}\times\boldsymbol{u}_\sigma + (\nabla p)_\sigma = 0, \end{cases} \tag{3.9}$$

where both parts of the material derivative and pressure gradient are defined in (3.7) and (3.8). How to introduce forcing in these equations is discussed in §4.1, after the discussion of the integrals of motion, because forcing has to be included in a way which is compatible with the conservation laws.

In the original system (2.7), Newton's equation can also be written as $\tilde{\partial}_t\boldsymbol{u} + \mu\partial_\sigma\boldsymbol{u} + \chi\hat{\boldsymbol{z}}\times\boldsymbol{u} + \tilde{\nabla}(p + \vartheta\nu + \frac{1}{2}\boldsymbol{u}^2) = \nu\tilde{\nabla}\vartheta$, where $\chi = f + \hat{\boldsymbol{z}}\cdot\tilde{\nabla}\times\boldsymbol{u}$, which is more useful than the equations with the term $\boldsymbol{u}\cdot\nabla\boldsymbol{u}$ for the derivation of the vorticity and energy theorems.

A similar operation can be done with the new model (3.9). Defining

$$\left.\begin{aligned} \bar{\chi} &:= f + \hat{\boldsymbol{z}} \cdot \nabla \times \bar{\boldsymbol{u}}, \\ \chi_\sigma &:= \hat{\boldsymbol{z}} \cdot \nabla \times \boldsymbol{u}_\sigma, \\ \bar{b} &:= \tfrac{1}{2}\bar{\boldsymbol{u}}^2 + \tfrac{1}{6}\boldsymbol{u}_\sigma^2 + (\bar{\vartheta} - \tfrac{1}{3}\vartheta_\sigma)h + \bar{\vartheta}h_0, \\ b_\sigma &:= \bar{\boldsymbol{u}} \cdot \boldsymbol{u}_\sigma + \vartheta_\sigma(h_0 + \tfrac{1}{2}h), \end{aligned}\right\} \tag{3.10}$$

the last two equations in (3.9) can also be written as

$$\left.\begin{aligned} &\partial_t \bar{\boldsymbol{u}} + \bar{\mu}\boldsymbol{u}_\sigma + \hat{\boldsymbol{z}} \times (\bar{\chi}\bar{\boldsymbol{u}} + \tfrac{1}{3}\chi_\sigma \boldsymbol{u}_\sigma) + \nabla \bar{b} = h_0 \nabla \bar{\vartheta} + \tfrac{1}{2}h\nabla(\bar{\vartheta} - \tfrac{1}{3}\vartheta_\sigma), \\ &\partial_t \boldsymbol{u}_\sigma + \hat{\boldsymbol{z}} \times (\chi_\sigma \bar{\boldsymbol{u}} + \bar{\chi}\boldsymbol{u}_\sigma) + \nabla b_\sigma = (h_0 + \tfrac{1}{2}h)\nabla\vartheta_\sigma - \tfrac{1}{2}h\nabla\bar{\vartheta}. \end{aligned}\right\} \tag{3.11}$$

These expressions show clearly that rotational forces arise because of density inhomogeneities in the layer ($\nabla\bar{\vartheta} \neq 0$, $\nabla\vartheta_\sigma \neq 0$).

These expressions also show that if $\bar{\vartheta}$ and ϑ_σ were independent of position along a closed rigid boundary, then the circulations $\oint \bar{\boldsymbol{u}} \cdot \mathrm{d}\boldsymbol{x}$ and $\oint \boldsymbol{u}_\sigma \cdot \mathrm{d}\boldsymbol{x}$ around that boundary would be constant. However, this constancy need not be expected: from the original system (2.1) it follows that $\oint \boldsymbol{u}|_\rho \cdot \mathrm{d}\boldsymbol{x}$ is constant, where $\boldsymbol{u}|_\rho$ is the velocity in a given isopycnal (this equation makes sense only for non-outcropping isopycnals); $\oint \bar{\boldsymbol{u}} \cdot \mathrm{d}\boldsymbol{x}$ is not conserved in the continuously stratified case.

4. Conservation laws

One important test of the validity of the approximation (3.9) has to do with which conservation laws of the original system (2.1) – or (2.7) – are preserved, and how. The solutions of the original system (2.1) are constrained by the existence of several integrals of motion, e.g. volume, mass and energy $\mathrm{d}(\mathscr{I}_0, \mathscr{I}_1, \mathscr{E})/\mathrm{d}t = 0$, with

$$(\mathscr{I}_0, \mathscr{I}_1, \mathscr{E}) := \int\!\!\int_D (h, h\bar{\vartheta}, E)\, \mathrm{d}a, \tag{4.1}$$

where D is the horizontal domain of the system, $\mathrm{d}a$ is the differential of area, and

$$E := h\overline{(\boldsymbol{u}^2/2 + v\vartheta)}. \tag{4.2}$$

Volume and mass are also conserved for the new system (3.9) since $\partial_t h + \nabla \cdot (h\bar{\boldsymbol{u}}) = 0$ and $\partial_t(h\bar{\vartheta}) + \nabla \cdot (h\bar{\vartheta}\bar{\boldsymbol{u}} + \tfrac{1}{3}h\vartheta_\sigma \boldsymbol{u}_\sigma) = 0$. As for the energy, in terms of the fields (3.3) of the new model the energy density (4.2) is found to be equal to

$$E = \tfrac{1}{2}h\bar{\boldsymbol{u}}^2 + \tfrac{1}{6}h\boldsymbol{u}_\sigma^2 + \tfrac{1}{2}h^2(\bar{\vartheta} - \tfrac{1}{3}\vartheta_\sigma) + hh_0\bar{\vartheta}. \tag{4.3}$$

Energy conservation is also guaranteed for the new model, since

$$\partial_t E + \nabla \cdot (h\bar{\boldsymbol{u}}\bar{b} + \tfrac{1}{3}h\boldsymbol{u}_\sigma b_\sigma) = 0. \tag{4.4}$$

The zonal momentum density is given by

$$M := h(\bar{u} - f_0 y - \tfrac{1}{2}\beta y^2) \tag{4.5}$$

in the β-plane case ($f = f_0 + \beta y$). From the new equations (3.9) it follows that the rate of change is

$$\partial_t M + \nabla \cdot (\bar{\boldsymbol{u}}M + \tfrac{1}{3}\boldsymbol{u}_\sigma h u_\sigma) + \tfrac{1}{2}\partial_x(h^2\bar{\vartheta}) + h\bar{\vartheta}(\partial_x h_0) = 0. \tag{4.6}$$

Consequently if D is a zonal channel *and* $\partial_x h_0 = 0$ (i.e. all boundaries are x-symmetric) then total zonal momentum is conserved. For a sphere, the equivalent of

the zonal momentum is the angular momentum around the Earth's axis, with density $M := h\cos\theta(\bar{u} + 2\Omega R\cos\theta)$, where Ω, R and θ are the Earth's rotation velocity, radius and latitude, respectively. A similar conservation law is easily found from the equations of motion (3.9) in spherical coordinates.

The original system (2.1) also has a family of integrals of motion, namely the Casimirs $\mathscr{C}[A] := \iint h\overline{A(\vartheta)}\,\mathrm{d}a$, where $A(\vartheta)$ is an *arbitrary* function of the buoyancy, in particular, $\mathscr{I}_n$ defined above correspond to $\mathscr{C}[\vartheta^n]$ for $n = 0, 1$. The new approximate model (3.9) does not conserve all these functionals. However, in addition to volume $\mathscr{I}_0$ and mass $\mathscr{I}_1$, the new model does conserve buoyancy variance

$$\mathscr{I}_2 := \int\!\!\int_D h\overline{\vartheta^2}\,\mathrm{d}a = \int\!\!\int_D h(\bar{\vartheta}^2 + \tfrac{1}{3}\vartheta_\sigma^2)\,\mathrm{d}a \tag{4.7}$$

since, in fact,

$$\partial_t\left[h(\bar{\vartheta}^2 + \tfrac{1}{3}\vartheta_\sigma^2)\right] + \nabla\cdot\left[h\bar{\boldsymbol{u}}(\bar{\vartheta}^2 + \tfrac{1}{3}\vartheta_\sigma^2) + \tfrac{2}{3}h\boldsymbol{u}_\sigma\vartheta_\sigma\bar{\vartheta}\right] = 0. \tag{4.8}$$

In summary, the approximate model IL^{1}PEM satisfies the laws of volume, mass, energy, momentum and buoyancy-variance conservation, and, moreover, the dependence of these integrals of motion on the dynamical fields is exactly the same as those in the original model IL$^\infty$PEM, evaluated with the fields of the ansatz (3.3). Ertel's theorem, however, needs a more elaborate discussion, and is the subject of §4.2.

4.1. *Forcing*

Forcing is introduced in the equations (3.9) of the new model in a way which is compatible with its conservation laws for energy, momentum and mass (or heat and salt content). For instance, let the system represent the upper layer of the ocean (case B in table 2), and assume a wind stress $\boldsymbol{\tau}$ acts at the surface ($\sigma = 1$) and a friction force acts at the base of the active layer ($\sigma = -1$): equations (3.9) are modified as

$$\left.\begin{aligned}\partial_t\bar{\boldsymbol{u}} + \ldots &= -r(\bar{\boldsymbol{u}} - \boldsymbol{u}_\sigma) + \boldsymbol{\tau}/h,\\ \partial_t\boldsymbol{u}_\sigma + \ldots &= -3r(\boldsymbol{u}_\sigma - \bar{\boldsymbol{u}}) + 3\boldsymbol{\tau}/h,\end{aligned}\right\} \tag{4.9}$$

where r could be a constant or some function of h and $|\bar{\boldsymbol{u}} - \boldsymbol{u}_\sigma|$, which is the jump in speed between the base of the active layer and the passive layer below. This is illustrated in figure 2(*a*). The coefficients of $\boldsymbol{\tau}/h$ in (4.9) are chosen so that the energy and momentum conservation equations, (4.4) and (4.6), are modified into

$$\left.\begin{aligned}\partial_t E + \ldots &= \boldsymbol{\tau}\cdot(\bar{\boldsymbol{u}} + \boldsymbol{u}_\sigma) - rh(\bar{\boldsymbol{u}} - \boldsymbol{u}_\sigma)^2,\\ \partial_t M + \ldots &= \boldsymbol{\tau}\cdot\hat{\boldsymbol{x}} - rh(\bar{\boldsymbol{u}} - \boldsymbol{u}_\sigma)\cdot\hat{\boldsymbol{x}}.\end{aligned}\right\} \tag{4.10}$$

Notice that the work done by the wind and the interfacial drag is proportional to the top and bottom velocities, $\bar{\boldsymbol{u}} + \boldsymbol{u}_\sigma$ and $\bar{\boldsymbol{u}} - \boldsymbol{u}_\sigma$, respectively.

On the other hand, a buoyancy input $\Gamma(\boldsymbol{x}, t)$ through $\sigma = 1$ can be added to the present model by rewriting the first two equations in (3.9) as

$$\left.\begin{aligned}\partial_t\bar{\vartheta} + \cdots &= \Gamma(\boldsymbol{x}, t)/h,\\ \partial_t\vartheta_\sigma + \cdots &= \nu\Gamma(\boldsymbol{x}, t)/h.\end{aligned}\right\} \tag{4.11}$$

Any value of ν is compatible with the equation

$$\partial_t\left(h\bar{\vartheta}\right) + \nabla\cdot\left(h(\bar{\boldsymbol{u}}\bar{\vartheta} + \tfrac{1}{3}\boldsymbol{u}_\sigma\vartheta_\sigma)\right) = \Gamma. \tag{4.12}$$

The choice $\nu = 1$ means that buoyancy is introduced linearly with depth with a

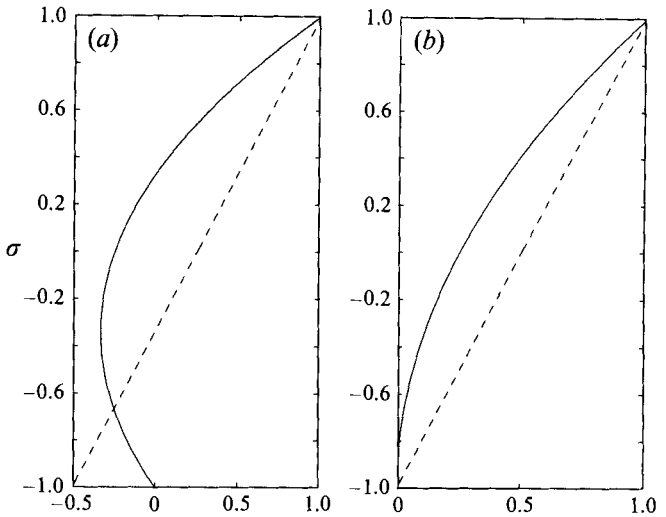

FIGURE 2. (*a*) The graph shows how stress at $\sigma = 1$ is introduced in the model (e.g. wind stress in the upper ocean layer). The solid line represents the vertical flux and the dashed one its vertical divergence, which is the forcing (both are normalized to unit surface value). (*b*) A similar situation, for the buoyancy (or heat) flux.

vanishing value at $\sigma = -1$ ($\partial_t(\bar{\vartheta} - \vartheta_\sigma) + \cdots = 0$). See figure 2(*b*). One typical example of this arises in the study of the thermodynamics of the upper layer of the ocean: Assuming that the density is controlled by the temperature, i.e. the buoyancy can be written as $\vartheta = g\alpha_T\,(T(\boldsymbol{x}, \sigma, t) - T_{down})$ where T is the temperature of the active layer, the buoyancy forcing is $\Gamma = g\alpha_T(\rho C_p)^{-1}Q$, where Q represents the heat input into the upper layer, α_T is the thermal expansion coefficient and C_p is the specific heat at constant pressure. In this case (4.12) gives the rate of change of the local heat content.

4.2. *Potential vorticity*

Vorticity is usually introduced via Stokes' theorem. However, the way it appears naturally in this work is through Ertel's theorem. What one needs is a covariant (coordinate-independent) representation of Ertel's operator $\mathscr{L} = q^\mu \partial(.)/\partial x^\mu$ where x^μ is *any* set of coordinates and $q^\mu = \mathscr{L} x^\mu$ *defines* the components of the absolute vorticity. In Cartesian coordinates ($x^\mu = \boldsymbol{x}, z$), which are orthogonal,†

$$2\mathscr{L} = (\hat{\boldsymbol{z}} \times \partial_z \boldsymbol{u}) \cdot \nabla + (f + \hat{\boldsymbol{z}} \cdot \nabla \times \boldsymbol{u})\partial_z, \tag{4.13}$$

whereas in σ-coordinates ($x^\mu = \boldsymbol{x}, \sigma$), which are *not* orthogonal,

$$\mathscr{L} = \boldsymbol{q} \cdot \tilde{\nabla} + q\,\partial_\sigma \tag{4.14}$$

with

$$\boldsymbol{q} := \frac{\hat{\boldsymbol{z}} \times \partial_\sigma \boldsymbol{u}}{h}, \quad q := \frac{f + \hat{\boldsymbol{z}} \cdot \tilde{\nabla} \times \boldsymbol{u}}{h}. \tag{4.15}$$

† Notice that $-\hat{\boldsymbol{z}} \times \nabla w$ is not included in the horizontal vorticity (because of the hydrostatic approximation). The factor 2 in (4.13) is introduced in order to have a simple expression for $\mathscr{L}$ in σ-coordinates; in case A of table 1, this factor equals −2. However, for the purposes of this paper what matters is the simpler expressions (4.14) and (4.15).

A very general vorticity theorem states that for any scalar s

$$\frac{\mathrm{D}(\mathscr{L}s)}{\mathrm{D}t} = \mathscr{L}\frac{\mathrm{D}s}{\mathrm{D}t} + \tfrac{1}{2}[\vartheta, s]^z. \tag{4.16}$$

If s is a Lagrangian constant $\mathrm{D}s/\mathrm{D}t = 0$ and $[\vartheta, s]^z = 0$, then $\mathscr{L}s$ is also conserved following fluid particles; this is known as Ertel's theorem.

Thus, in the IL$^\infty$PEM the 'density–potential vorticity' $\mathscr{L}\vartheta$, *is* conserved because $\mathrm{D}\vartheta/\mathrm{D}t = 0$ and obviously $[\vartheta, \vartheta]^z \equiv 0$. This Lagrangian constant must be distinguished from q $(= \mathscr{L}\sigma)$ defined in (4.15), which may be called the 'σ-potential vorticity'. In the HLPEM (ϑ = constant, $\boldsymbol{u} \equiv \bar{\boldsymbol{u}}$), $\mathrm{D}\sigma/\mathrm{D}t = 0$ and obviously $[\vartheta, \sigma]^z \equiv 0$: Ertel's theorem guarantees conservation of q, i.e. $\mathrm{D}q/\mathrm{D}t = 0$. With density gradients within the layer, though, this potential vorticity is *not* conserved because, in general, neither $\mathrm{D}\sigma/\mathrm{D}t = 0$ (there is a velocity shear) nor $[\vartheta, \sigma]^z = 0$ (isopycnals are not parallel to the surface). Instead, using $s = \sigma$ in (4.16) it follows that

$$\frac{\mathrm{D}q}{\mathrm{D}t} = \mathscr{L}\mu + h^{-1}[v, \vartheta]^\sigma, \tag{4.17}$$

where the decomposition (2.4) of the Jacobian has been used. Use of $s = \boldsymbol{x}$ in (4.16) gives the rate of change of vector $\boldsymbol{q}$ $(= \mathscr{L}\boldsymbol{x})$ as

$$\frac{\mathrm{D}\boldsymbol{q}}{\mathrm{D}t} = \mathscr{L}\boldsymbol{u} + h^{-1}\hat{\boldsymbol{z}} \times (\partial_\sigma v \tilde{\nabla}\vartheta - \partial_\sigma \vartheta \tilde{\nabla} v), \tag{4.18}$$

with a similar use of (2.4). These two equations give the evolution of the vorticity field, in σ-coordinates: the first term on the right-hand side represents vortex-stretching and the rest is the baroclinic torque.

In the approximate model IL^{1}PEM, derived from the ansatz (3.3),

$$\boldsymbol{q} = \bar{\boldsymbol{q}} = \frac{\hat{\boldsymbol{z}} \times \boldsymbol{u}_\sigma}{h} \tag{4.19}$$

(notice that there is no representation for $\boldsymbol{q}_\sigma$) and $q = \bar{q} + \sigma\, q_\sigma$ with

$$\bar{q} = \frac{f + \hat{\boldsymbol{z}} \cdot \nabla \times \bar{\boldsymbol{u}}}{h} \quad \text{and} \quad q_\sigma = \frac{\hat{\boldsymbol{z}} \cdot \nabla \times \boldsymbol{u}_\sigma}{h}. \tag{4.20}$$

Given an arbitrary scalar field $s = \bar{s} + \sigma s_\sigma$, $\mathscr{L}s$ is represented by $\overline{\mathscr{L}s} = \bar{\boldsymbol{q}} \cdot \nabla\bar{s} + \bar{q}s_\sigma$ and $(\mathscr{L}s)_\sigma = \bar{\boldsymbol{q}} \cdot \nabla s_\sigma + q_\sigma s_\sigma$. In order to check the validity of theorem (4.16) for the new (approximate) model, it is useful to analyse its derivation for the exact model in σ-coordinates (2.7), which is done in Appendix A. Following the same steps as in Appendix A, but with equations (3.9) of the new model, it is found that

$$\left.\begin{aligned} \overline{\mathrm{D}q/\mathrm{D}t} &= \bar{\boldsymbol{q}} \cdot \nabla\bar{\mu} - q_\sigma\bar{\mu} + h^{-1}\overline{[v, \vartheta]^\sigma}, \\ (\mathrm{D}q/\mathrm{D}t)_\sigma &= -3\bar{q}\bar{\mu} + h^{-1}\left([v_\sigma, \bar{\vartheta}] + [\bar{v}, \vartheta_\sigma]\right), \\ \overline{\mathrm{D}\boldsymbol{q}/\mathrm{D}t} &= \bar{\boldsymbol{q}} \cdot \nabla\bar{\boldsymbol{u}} + \bar{q}\boldsymbol{u}_\sigma + h^{-1}\hat{\boldsymbol{z}} \times (v_\sigma \nabla\bar{\vartheta} - \vartheta_\sigma \nabla\bar{v}). \end{aligned}\right\} \tag{4.21}$$

The third equation is an exact representation of (4.18). The first two may not appear to be a good representation of (4.17), but in fact they are. In order to show this, it is necessary to go back to the continuity equation (2.9). In §3 that equation was used to find $\mu = 2h^{-1}(1 - \sigma^2)\nabla \cdot (h\boldsymbol{u}_\sigma)$ and thus $\bar{\mu} = \frac{1}{3}h^{-1}\nabla \cdot (h\boldsymbol{u}_\sigma)$, which is the only quantity needed in the evolution equations (3.9) of the new model, i.e. at the level of approximation represented by the ansatz (3.3). However, in the derivation of the vorticity equation, outlined in Appendix A, the σ-derivative of μ is also needed. In

the approximate model this is given by

$$\partial_\sigma \mu = -3\sigma\bar{\mu}. \tag{4.22}$$

Even though $\mu_\sigma := \overline{\partial_\sigma \mu} = 0$, $\overline{\sigma\partial_\sigma \mu} = -\bar{\mu}$; notice than this satisfies the trivial equality $\overline{\partial_\sigma(\sigma\mu)} = 0$. Consequently, the first terms on the right-hand side of the $\bar{q}$ and q_σ equations in (4.21) are the correct representation of those of (4.17), namely $\overline{\mathscr{L}\mu} = \bar{\boldsymbol{q}}\cdot\nabla\bar{\mu} + \overline{q\partial_\sigma\mu} \equiv \bar{\boldsymbol{q}}\cdot\nabla\bar{\mu} - q_\sigma\bar{\mu}$ and $\sigma(\mathscr{L}\mu)_\sigma = \sigma\boldsymbol{q}_\sigma\cdot\nabla\bar{\mu} + \bar{\boldsymbol{q}}\cdot\tilde{\nabla}(\sigma\mu_\sigma) + \bar{q}\partial_\sigma\mu \equiv -3\sigma\bar{q}\bar{\mu}$.

In summary, the new model satisfies the correct vorticity equation for $\bar{q} = \overline{\mathscr{L}x}$, $\bar{q} = \overline{\mathscr{L}\sigma}$, and $q_\sigma = (\mathscr{L}\sigma)_\sigma$. However, it does not have an exact representation of the general theorem (4.16), i.e. for a general $\mathscr{L}s$, for the following reason. Let $\mathrm{D}a/\mathrm{D}t = A$ and $\mathrm{D}b/\mathrm{D}t = B$, with $a = \bar{a} + \sigma a_\sigma$, and similarly for A, b, and B. It is easy to show that $\mathrm{D}(ab)/\mathrm{D}t = Ab + aB$ is correctly represented in the mean, i.e. $\overline{\mathrm{D}(ab)/\mathrm{D}t} = \overline{Ab + aB}$, but all four evolution equations (i.e. those for $\bar{a}$, a_σ, $\bar{b}$ and b_σ) are needed in order to get this result. Consequently, there is one element missing in (4.21) to allow a derivation of $\overline{\mathrm{D}(\mathscr{L}s)/\mathrm{D}t}$: the evolution equation of $\boldsymbol{q}_\sigma$. In fact, at this level of approximation there is not even a representation for $\boldsymbol{q}_\sigma$ (one would need to know $\partial_{\sigma\sigma}\boldsymbol{u}$ for that). Therefore, the only potential vorticity whose evolution is correctly represented by the new model is $q = \mathscr{L}\sigma$ (which, however, is not conserved). The evolution of the 'density potential vorticity' $\mathscr{L}\vartheta$ (which in the exact model IL$^\infty$PEM is conserved) is not correctly represented by the new model for lack of knowledge of the vertical *curvature* of the velocity field. Neither does $\overline{\mathrm{D}(\mathscr{L}\vartheta)/\mathrm{D}t}$ vanish identically nor is $\iint h\overline{\mathscr{L}\vartheta}$ guaranteed to be constant.

5. Hamiltonian structure

Let me review the main points of Hamiltonian theory (see for instance Shepherd 1990; Ripa 1993*b*). If a dynamical system, described by some fields $\varphi^a(\boldsymbol{x}, t)$, is Hamiltonian, then there exist a functional $\mathscr{H}[\varphi^a, \boldsymbol{x}, t]$ and a Poisson tensor (operator) $\boldsymbol{J}^{ab}$ such that the evolution equations can be written in the form

$$\frac{\partial\varphi^a}{\partial t} = \boldsymbol{J}^{ab}\frac{\delta\mathscr{H}}{\delta\varphi^b} \tag{5.1}$$

(with implicit summation over repeated indices). The x-momentum $\mathscr{M}[\varphi^a, \boldsymbol{x}, t]$, when it exists, is such that

$$\frac{\partial\varphi^a}{\partial x} = -\boldsymbol{J}^{ab}\frac{\delta\mathscr{M}}{\delta\varphi^b}. \tag{5.2}$$

Thus, $\mathscr{H}$ and $\mathscr{M}$ are related to t- and x-translations. Finally, a Casimir $\mathscr{C}[\varphi^a, \boldsymbol{x}]$ satisfies

$$\boldsymbol{J}^{ab}\frac{\delta\mathscr{C}}{\delta\varphi^b} = 0, \tag{5.3}$$

i.e. it does not produce any transformation of the dynamical field. If there exist one or more Casimirs, then the Hamiltonian system is *singular*. In such a case, $\mathscr{H}$ and $\mathscr{M}$ are defined modulo the addition of Casimirs (with the appropriate units); $\mathscr{H} - \alpha\mathscr{M} + \mathscr{C}$ (with α some constant) is a Hamiltonian in a frame moving with velocity α along x, i.e.

$$\frac{\partial\varphi^a}{\partial t} + \alpha\frac{\partial\varphi^a}{\partial x} = \boldsymbol{J}^{ab}\frac{\delta(\mathscr{H} - \alpha\mathscr{M} + \mathscr{C})}{\delta\varphi^b}. \tag{5.4}$$

All Casimirs are integrals of motion. If the system is invariant under t-translations

Model	Casimirs
HLPEM	$\iint h A(\bar{q})$
IL^{0}PEM	$\iint h A_1(\bar{\vartheta}) + h\bar{q} A_2(\bar{\vartheta})$
IL^{1}PEM	$\iint h \overline{(a_0 + a_1\vartheta + a_2\vartheta^2)}$
IL$^\infty$PEM	$\iint h \overline{A(\vartheta, \mathscr{L}\vartheta)}$

NB $A(\cdots)$ and a_k are arbitrary.

TABLE 3.

(x-translations), then $\mathscr{H}$ ($\mathscr{M}$) is conserved. Let $\mathscr{H} - \alpha\mathscr{M} + \mathscr{C} =$ constant for any initial condition: most stability/instability theorems can be derived by searching for a constant α and a Casimir $\mathscr{C}$ such that the first variation of this integral of motion from a certain steady basic state vanishes: $\delta(\mathscr{H} - \alpha\mathscr{M} + \mathscr{C}) = 0$. Thus, if the second variation $\delta^2(\mathscr{H} - \alpha\mathscr{M} + \mathscr{C})$ is sign definite, then the basic state is stable (e.g. Shepherd 1990; Ripa 1990); conversely, if the basic state is unstable, then the second variation must be sign indefinite and can be used to make statements on the structure of growing perturbations (Ripa 1992*b*). The possibility of deriving these stability/instability theorems for a certain singular Hamiltonian system is somehow linked to the number of Casimirs available to construct those integrals of motion.

The model in this paper is Hamiltonian, with $\mathscr{H}$ equal to the total energy

$$\mathscr{H}[\bar{\vartheta}, \vartheta_\sigma, h, \bar{\boldsymbol{u}}, \boldsymbol{u}_\sigma] = \iint_D E \, \mathrm{d}a \tag{5.5}$$

and a Poisson tensor $\boldsymbol{J}^{ab}$ given in Appendix B. If the coasts are zonal, then $\mathscr{M}$ from (4.5) is a momentum in the sense of (5.2), independent of whether $\mathscr{M}$ is conserved or not, i.e. of whether $\partial h_0/\partial x$ vanishes or not. Finally, $\mathscr{I}_n = \iint h\overline{\vartheta^n}$ ($n = 0, 1, 2$) are Casimirs. However, no vorticity-related Casimirs have been found for the new system (see Appendix B) because of its failure to represent $\mathrm{D}(\mathscr{L}\vartheta)/\mathrm{D}t = 0$, as explained in §4.2. The implication is that it is not possible to solve for $\delta(\mathscr{H} - \alpha\mathscr{M} + \mathscr{C}) = 0$ at a basic state with currents (except for a trivial one, discussed in §6).

Given the importance of the existence of Casimirs for the derivation of stability/instability theorems, it is interesting to compare these integrals of motion for the different models discussed here; see table 3. In the fully three-dimensional case, the most general Casimir is the integral of an arbitrary function A of the buoyancy ϑ and the density potential vorticity $\mathscr{L}\vartheta$; in the present model A is reduced to just a parabola in ϑ. In the particular case of a uniform ϑ-field, that integral reduces to the volume $\iint h$. However, in this case (HLPEM) there are additional Casimirs, namely, the integral of an arbitrary function A of the σ-potential vorticity $\bar{q}$. Finally, when $\vartheta = \bar{\vartheta}(\boldsymbol{x}, t)$ (IL^{0}PEM), the function A must be linear in $\bar{q}$, but with coefficients A_1 and A_2 which may depend on $\bar{\vartheta}$.

Notice that q does not appear in the Casimirs of either IL$^\infty$PEM or IL^{1}PEM, and therefore its absence is certainly not a limitation of the latter; one may wonder why is it present in the Casimirs of the HLPEM and the IL^{0}PEM. From the exact equation (4.17) it follows that

$$\tilde{\partial}_t(hq) + \tilde{\nabla} \cdot (h\boldsymbol{u}q) + \partial_\sigma(h\mu q) = \tilde{\nabla} \cdot (\mu\boldsymbol{q} + \hat{\boldsymbol{z}} \times \nu\tilde{\nabla}\vartheta), \tag{5.6}$$

whereas from the approximate system (4.21) it can be derived that

$$\partial_t(h\bar{q}) + \nabla\cdot(h\overline{uq}) = \nabla\cdot(\bar{\mu}\bar{\boldsymbol{q}} + \hat{z}\times\overline{v\nabla\vartheta}), \tag{5.7}$$

where $\overline{ab} = \bar{a}\bar{b} + \frac{1}{3}a_\sigma b_\sigma$. It is clear that (5.7) is an exact representation of the vertical average of (5.6). The rate of change of $\iint h\bar{q}$ is driven by the term in the right-hand side, more precisely by the boundary integral of the normal component of $\bar{\mu}\bar{\boldsymbol{q}}$ plus the tangential component of $\overline{v\nabla\vartheta}$. For the HLPEM μ, $\boldsymbol{q}$ and $\nabla\vartheta$ vanish identically. In the approximation represented by the IL^{0}PEM, μ and $\boldsymbol{q}$ vanish for lack of resolution of the vertical gradients, whereas ϑ can be chosen to be position independent along each rigid boundary (Ripa 1993*a*); that is why $\iint h\bar{q}$ is conserved in systems without vertical density gradients inside the active layer.

6. Free energy

The negative sign in the second-to-last term in the energy density (4.3) might lead to the impression that the new system could be spontaneously unstable, i.e. to be able to grow without limit. That is not necessarily correct. In order to address this problem one must have an integral of motion which is quadratic, to the lowest order, in the deviation from a (steady) reference state; the sign indefiniteness or definiteness of this integral determines whether or not the fields can grow without bound. Thus, consider the following decomposition of the dynamical fields:

$$\begin{pmatrix}\bar{\vartheta}\\ \vartheta_\sigma\\ h\\ \bar{\boldsymbol{u}}\\ \boldsymbol{u}_\sigma\end{pmatrix} = \begin{pmatrix}\bar{\Theta}(\boldsymbol{x})\\ \Theta_\sigma(\boldsymbol{x})\\ H(\boldsymbol{x})\\ \bar{\boldsymbol{U}}(x)\\ \boldsymbol{U}_\sigma(\boldsymbol{x})\end{pmatrix} + \begin{pmatrix}\delta\bar{\vartheta}(\boldsymbol{x},t)\\ \delta\vartheta_\sigma(\boldsymbol{x},t)\\ \delta h(\boldsymbol{x},t)\\ \delta\bar{\boldsymbol{u}}(\boldsymbol{x},t)\\ \delta\boldsymbol{u}_\sigma(\boldsymbol{x},t)\end{pmatrix}, \tag{6.1}$$

where the reference state $[\bar{\Theta}, \Theta_\sigma, H, \bar{\boldsymbol{U}}, \boldsymbol{U}_\sigma]$ has yet to be specified, and the following integral of motion:

$$\mathscr{E}_f := \mathscr{E} - \alpha\mathscr{M} + \sum_{n=0}^{2} a_n\mathscr{I}_n, \tag{6.2}$$

where α and the a_n are constants (cases with $\alpha \neq 0$ are only considered when $\mathscr{M}$ is conserved, i.e. $\partial_x h_0 = 0$ and for zonal coasts). The question is: can we choose these constants so that the first variation of $\mathscr{E}_f$ from the reference state vanishes? If such an $\mathscr{E}_f$ can be constructed, it will represent an integral of motion which is quadratic, to the lowest order, in the deviation from this reference state.

In order for this integral of motion to be extreme, $\delta\mathscr{E}_f = 0$, its functional derivatives with respect to all dynamical fields must vanish at the reference state. Thus (for the β-plane case), first,

$$\frac{\delta\mathscr{E}_f}{\delta\bar{\boldsymbol{u}}} = 0 \Rightarrow \bar{U} = \alpha, \bar{V} = 0; \quad \frac{\delta\mathscr{E}_f}{\delta\boldsymbol{u}_\sigma} = 0 \Rightarrow \boldsymbol{U}_\sigma = 0, \tag{6.3}$$

i.e. the velocity field in the reference state must be zonal and uniform. (In the case of spherical geometry, where $\mathscr{M}$ is the angular momentum, α is then a uniform angular velocity of the reference state.) Secondly,

$$\frac{\delta\mathscr{E}_f}{\delta\vartheta_\sigma} = 0 \Rightarrow \tfrac{1}{2}N_r^{-2} = a_2. \tag{6.4}$$

This equation implies that the spatial structure of the fields $\Theta_\sigma(\boldsymbol{x})$ and $H(\boldsymbol{x})$ cannot

be arbitrary, since their ratio must be constant, so that the squared Brunt–Väisälä field is uniform, $N_r^2 = 2\Theta_\sigma/H = \text{const.}$, in the reference state. Thirdly,

$$\frac{\delta \mathscr{E}_f}{\delta \bar{\vartheta}} = 0 \Rightarrow h_0 + \tfrac{1}{2}H + N_r^{-2}\bar{\Theta} = -a_1. \tag{6.5}$$

Now, the pressure force acting on the velocity shear $\boldsymbol{u}_\sigma$ is $(\nabla p)_\sigma = \vartheta_\sigma \nabla(h_0 + \frac{1}{2}h) + \frac{1}{2}h\nabla\vartheta$. Taking ϑ_σ as common factor and using equation (6.4) for a_2 it follows that, at the reference state, $(\nabla P)_\sigma = -\Theta_\sigma \nabla a_1$; since a_1 is constant this force vanishes. Therefore, the evolution equation for $\boldsymbol{u}_\sigma$ is satisfied in the reference state. Finally,

$$\frac{\delta \mathscr{E}_f}{\delta h} = 0 \Rightarrow \tfrac{1}{2}\bar{U}^2 - \bar{U}(f_0 y + \tfrac{1}{2}\beta y^2) + \tfrac{1}{2}N_r^{-2}\left(\bar{\Theta} - \Theta_\sigma\right)^2 = a_0. \tag{6.6}$$

The depth-averaged pressure gradient is $\overline{\nabla p} = (\bar{\vartheta} - \frac{1}{3}\vartheta_\sigma)\nabla h + \frac{1}{2}h\nabla(\bar{\vartheta} - \frac{1}{3}\vartheta_\sigma) + \bar{\vartheta}\nabla h_0$; at the reference state then $\overline{\nabla P} = -\bar{\Theta}\nabla a_1 + \nabla[\bar{U}(f_0 y + \frac{1}{2}\beta y^2) - \frac{1}{2}\bar{U}^2 - a_0] \equiv \bar{U}\,\nabla(f_0 y + \frac{1}{2}\beta y^2)$, which is easily shown to be equal to $-f\hat{\boldsymbol{z}} \times \bar{\boldsymbol{U}}$, i.e. the evolution equation of $\bar{\boldsymbol{u}}$ is also satisfied in the reference state.

Notice that $h_0(\boldsymbol{x})$ and the four constants (α, a_0, a_1, a_2) fully determine the structure of the reference state. If there is neither topography, $h_0 = 0$, nor current, $\alpha = 0$, then it can be shown, from (6.3)–(6.6), that this solution has uniform $\bar{\Theta}$-, Θ_σ-, and H-fields. On the other hand, with a mean flow, $\alpha \neq 0$, the fields $\bar{\Theta}$, Θ_σ, and H (as well as h_0) must be x-independent in order for their evolution equations to be satisfied. In either case the evolution equations for $\bar{\vartheta}, \vartheta_\sigma$, and h are trivially satisfied in the reference state. Of course, not all steady states qualify as a 'reference state' to build $\mathscr{E}_f$ because the left-hand side in the equations above for α, a_0, a_1, and a_2 may not be constant for some of them.

Using the values of α, a_0, a, and a_2 from (6.3)–(6.6) and subtracting a trivial constant (the value of $\mathscr{E} - \alpha\mathscr{M} + \sum a_n \mathscr{I}_n$ at the reference state) it follows that the integral of motion just constructed has the form

$$\text{IL}^1\text{PEM}: \begin{cases} \mathscr{E}_f = \tfrac{1}{2}\displaystyle\int\!\!\int_D \{h(\delta\bar{\boldsymbol{u}})^2 + \tfrac{1}{3}h(\delta \boldsymbol{u}_\sigma)^2 + [g_r - \tfrac{1}{6}N_r^2 H](\delta h)^2 \\ \qquad\qquad + h\,\delta h[\delta\bar{\vartheta} - \tfrac{1}{3}\delta\vartheta_\sigma] + hN_r^{-2}[(\delta\bar{\vartheta})^2 + \tfrac{1}{3}(\delta\vartheta_\sigma)^2]\}. \end{cases} \tag{6.7}$$

where $g_r(\boldsymbol{x}) := \bar{\Theta}$. The available potential energy density, i.e. the last three terms, can be rewritten in the form

$$[g_r - \tfrac{1}{2}N_r^2 H - \tfrac{1}{3}N_r^2 \delta h](\delta h)^2 + hN_r^{-2}(\delta\bar{\vartheta} + \tfrac{1}{2}N_r^2\delta h)^2 + \tfrac{1}{3}hN_r^{-2}(\delta\vartheta_\sigma - \tfrac{1}{2}N_r^2\delta h)^2, \tag{6.8}$$

which shows explicitly that the lowest-order $\mathscr{E}_f$ (quadratic) is positive definite if and only if

$$g_r > \tfrac{1}{2}N_r^2 H > 0 \tag{6.9}$$

$(\bar{\Theta} > \Theta_\sigma > 0)$, which are the conditions for the buoyancy to be everywhere positive and for the density to increase with depth. Given any initial condition, a 'nearby' reference state with the characteristics discussed here can be found. Then conservation of $\mathscr{E}_f$ implies that the system cannot deviate much from it, i.e. it cannot 'explode'.† This is illustrated in figure 3 where the centre point symbolizes the reference state. Other equilibrium points, two stable ones (elliptic) and two probably unstable (hyperbolic),

† Since $\mathscr{E}_f$ has cubic terms in $[\delta\bar{\vartheta}, \delta\vartheta_\sigma, \delta h, \delta\bar{\boldsymbol{u}}, \delta\boldsymbol{u}_\sigma]$, in addition to the quadratic ones, this is not a statement on a well-defined *distance* in state space. This represents an important difference between primitive equations and quasi-geostrophic models, for which the free energy is exactly quadratic.

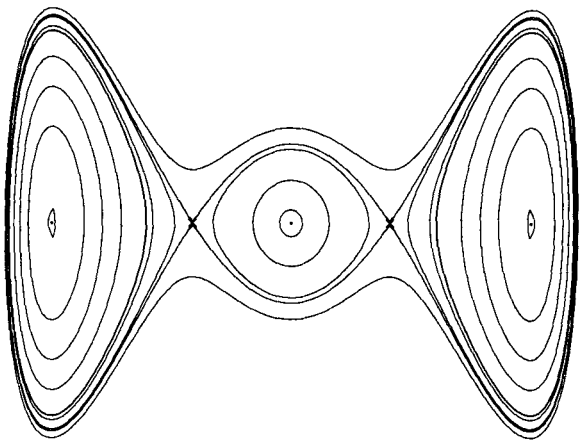

FIGURE 3. Contour lines of an integral of motion in a hypothetical (two-dimensional) state space. The system can never move arbitrarily far from the centre point, which is then a suitable 'reference state' to prove the boundness of the system evolution. Four other equilibrium states are shown (two stable and two probably unstable). In a system with more degrees of freedom (e.g. a fluid!) an equilibrium state may be an elliptic or hyperbolic point of some integrals of motion but not so of others.

are shown in the same figure; there may be equilibrium points that are neither elliptic nor hyperbolic points of a particular integral of motion, i.e. its first variation does not vanish at that point. Most stability and instability theorems can be derived by finding an integral of motion with an elliptic or hyperbolic point at an equilibrium state which corresponds to non-uniform currents. In order to find such an integral of motion, for systems like those discussed in this paper, it is necessary to have Casimir integrals of motion which depend on the vorticity field. As explained in §5 these integrals of motion do not exist for the system in this paper (see table 3) and therefore it is apparently not possible to derive stability theorems by 'Arnol'd's method' (Arnol'd 1965, 1966).

7. Comparison with other models

Here I compare the free energy of the present model with that of the other systems discussed in this paper; in all cases I will consider a reference state with neither current nor topography, $h_0 \equiv 0$, for simplicity. For the model in this paper, this corresponds to $\alpha = 0$ in equations (6.3)–(6.6) above, which require all three parameters g_r, N_r^2, and H to be constant.

Let me start with the continuously stratified case, which it is taken here as the 'exact' one. From (2.2), (4.2) and the Casimir integral of motion in table 3, it follows that the free energy must be of the form

$$\mathscr{E}_f = \frac{1}{2}\int\int_D h\,\overline{[\boldsymbol{u}^2 + h(1-\sigma)\vartheta + A(\vartheta)]}, \tag{7.1}$$

where the arbitrary function $A(\vartheta)$ is chosen so that the first variation of $\mathscr{E}_f$ vanishes at $[\vartheta, h, \boldsymbol{u}] = [\Theta(\sigma), H, \mathbf{0}]$. For a linear stratification in the reference state,

$$\Theta(\sigma) = g_r + \tfrac{1}{2}N_r^2 H\,\sigma, \tag{7.2}$$

it can be shown that $A(\vartheta)$ is a second-order polynomial, and that

$$\text{IL}^{\infty}\text{PEM}: \mathscr{E}_f = \frac{1}{2}\int\!\!\int_D \Big\{h\overline{(\delta \boldsymbol{u})^2} + [g_r - \tfrac{1}{6}N_r^2 H](\delta h)^2 + h\delta h\overline{(1-\sigma)\delta\vartheta} + hN_r^{-2}\,\overline{(\delta\vartheta)^2}\Big\}. \tag{7.3}$$

This integral evaluated with the perturbation fields of the form (3.3) gives precisely the free energy of the present model (6.7). (Notice that if $N_r^2 \equiv 0$ then it is not possible to construct this integral of motion, quadratic to the lowest order in the deviation from the reference state.)

The free energy for the classical shallow water equations (3.1), or the HLPEM, is

$$\text{HLPEM}: \mathscr{E}_f = \frac{1}{2}\int\!\!\int_D \big\{h\bar{\boldsymbol{u}}^2 + g_r\delta h^2\big\}. \tag{7.4}$$

This is the limit of (7.3) for a uniform buoyancy field, in the sense that both $N_r^2 \to 0$ and $\delta\vartheta \to 0$, but in such a way that $\delta\vartheta/N_r^2 \to 0$. The free energy for the IL^{0}PEM, (3.2), is

$$\text{IL}^0\text{PEM}: \mathscr{E}_f = \frac{1}{2}\int\!\!\int_D \Big\{h\bar{\boldsymbol{u}}^2 + \big(\bar{\vartheta}^{1/2}h - g_r^{1/2}H\big)^2\Big\} \tag{7.5}$$

(see Ripa 1995*a*). Notice that the quadratic part of $\mathscr{E}_f$ is positive definite for HLPEM but not for IL^{0}PEM. Variations of h and $\bar{\vartheta}$ which leave $h^2\bar{\vartheta}$ unaltered do not change the free energy; for infinitesimal perturbations this corresponds to the *force-compensating mode*, $2g_r h' + H\bar{\vartheta}' = 0$. Spontaneous growth of such perturbations is not prevented by conservation of $\mathscr{E}_f$. Figure 4 depicts how the present model and the HLPEM have a positive definite $\mathscr{E}_f$, whereas the IL^{0}PEM has a non-negative definite $\mathscr{E}_f$.

The HLPEM (3.1) is a particular case of both (3.2) and (3.9), obtained for any initial condition with a uniform buoyancy field ($\bar{\vartheta} \equiv g_r$, $\vartheta_\sigma \equiv 0$). This condition is preserved by the dynamics and makes (7.5) coincide with (7.4). Notice, however, that the IL^{0}PEM (3.2) is *not* a particular case of the new model (3.9) and certainly (7.5) is not the limit of (6.7) when $\vartheta_\sigma \to 0$ because this implies $N_r^{-2} \to \infty$. In this sense, the step from the – now classical – system (3.2) to the new one (3.9) is different than that from (3.1) to (3.2).

7.1. *Reduced model with* $\boldsymbol{u}_\sigma \equiv 0$

In their original model, Schopf & Cane (1983) had an intermediate layer in which they allowed variable density stratification but neglected the velocity shear, i.e. $\vartheta_\sigma \neq 0$ but $\boldsymbol{u}_\sigma = 0$ in the notation in this paper. This type of model can be obtained from (3.9), neglecting all $\boldsymbol{u}_\sigma$ as well as the equation for this field, which gives

$$\left.\begin{aligned} &(\partial_t + \bar{\boldsymbol{u}}\cdot\nabla)\bar{\vartheta} = 0,\\ &(\partial_t + \bar{\boldsymbol{u}}\cdot\nabla)\vartheta_\sigma = 0,\\ &\partial_t h + \nabla\cdot(h\bar{\boldsymbol{u}}) = 0,\\ &(\partial_t + \bar{\boldsymbol{u}}\cdot\nabla)\bar{\boldsymbol{u}} + f\hat{\boldsymbol{z}}\times\bar{\boldsymbol{u}} + \tfrac{1}{2}h^{-1}\nabla[h^2(\bar{\vartheta} - \tfrac{1}{3}\vartheta_\sigma)] = 0. \end{aligned}\right\} \tag{7.6}$$

The energy density of this reduced system is $E = \frac{1}{2}h\bar{\boldsymbol{u}}^2 + \frac{1}{2}h^2(\bar{\vartheta} - \frac{1}{3}\vartheta_\sigma)$, whereas the Casimir integrals of motion have the form $\mathscr{C} = \int\!\!\int h A(\bar{\vartheta}, \vartheta_\sigma)$, where $A(\bar{\vartheta}, \vartheta_\sigma)$ is completely arbitrary. Comparing the Casimirs with those of the continuously stratified case, it is seen that this system has 'far too many' because $A(\bar{\vartheta}, \vartheta_\sigma)$ need not

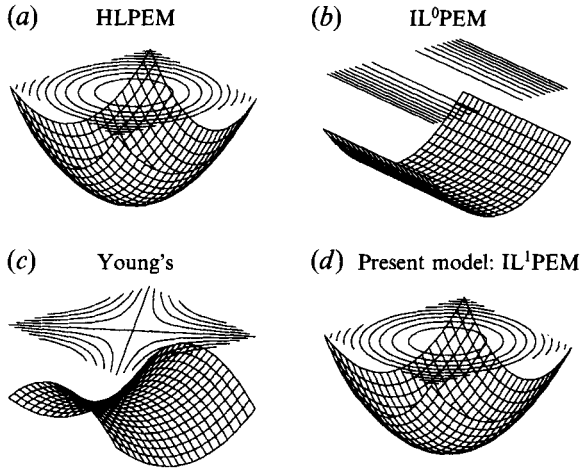

FIGURE 4. Topology of the free energy $\mathscr{E}_f$ for different types of models discussed in the text, sketched as contour lines and three-dimensional surfaces in a two-dimensional state space. The HLPEM, with constant density, have an $\mathscr{E}_f$ which is positive definite in the deviation from a reference state, as the example in figure 3. A similar structure is found for the $\mathscr{E}_f$ of the present model, IL^{1}PEM, which allows for both horizontal and vertical variations of density. The IL^{0}PEM, with only horizontal density variability, has a non-negative definite $\mathscr{E}_f$: there are 'zero energy' variations of the density and depth fields, which leave the flow unaltered. Finally, the $f\tau \to \infty$ limit of Young's (1994) model has sign-indefinite $\mathscr{E}_f$: the system may spontaneously 'explode'.

be a vertical average of a function of $\bar{\vartheta} + \sigma\,\vartheta_\sigma$. At this level, with neither topography nor forcing, this model is equivalent to the IL^{0}PEM (3.2), with $\bar{\vartheta}$ replaced by $\bar{\vartheta} - \frac{1}{3}\vartheta_\sigma$, including in (7.5), which gives a free energy of the form

$$\mathscr{E}_f = \frac{1}{2}\int\!\!\int_D \left\{ h\bar{\boldsymbol{u}}^2 + \left(h(\bar{\vartheta} - \vartheta_\sigma/3)^{1/2} - H(g_r - N_r^2 H/6)^{1/2}\right)^2 \right\}, \tag{7.7}$$

which is non-negative definite.

7.2. *Reduced model with* $\vartheta_\sigma \equiv 0$

The limit $f\tau \to \infty$ of Young's (1994) model yields the opposite case: $\vartheta_\sigma = 0$ but $\boldsymbol{u}_\sigma \neq 0$. That model is meant to be valid for low frequencies, where the velocity field is diagnosed from the geostrophic balances

$$\left.\begin{aligned} f\hat{\boldsymbol{z}} \times \bar{\boldsymbol{u}}^g + \nabla(g_r \delta h + \tfrac{1}{2} H \delta\bar{\vartheta}) = 0, \\ f\hat{\boldsymbol{z}} \times \boldsymbol{u}_\sigma^g + \tfrac{1}{2} H \nabla(\delta\bar{\vartheta}) = 0 \end{aligned}\right\} \tag{7.8}$$

(the latter is the *thermal wind* relation alluded to above). Young's model satisfies $\mathrm{d}\mathscr{E}_f/\mathrm{d}t = 0$ where

$$\mathscr{E}_f = \frac{1}{2}\int\!\!\int_D \left\{ H(\bar{\boldsymbol{u}}^g)^2 - \tfrac{1}{3}H(\boldsymbol{u}_\sigma^g)^2 + g_r\left(\delta h + \tfrac{1}{2}H\,\delta\bar{\vartheta}/g_r\right)^2 \right\} \tag{7.9}$$

is shown to be a Hamiltonian to his equations in Ripa (1995*b*); its structure in state space is compared in figure 4 with that of $\mathscr{E}_f$ from the present model, the HLPEM and the IL^{0}PEM. The negative sign in the second term in (7.9), with respect to (6.7), indicates that his system can 'explode' by itself, e.g. through the mechanism of 'explosive resonant triads' (see the Introduction).

A model with $\vartheta_\sigma = 0$ but $\boldsymbol{u}_\sigma \neq 0$, not restricted to low frequencies, can be obtained

from (3.9) neglecting all ϑ, as well as the evolution equation of this field, namely

$$\left.\begin{aligned}
&(\partial_t + \bar{\boldsymbol{u}}\cdot\nabla)\bar{\vartheta} = 0,\\
&\partial_t h + \nabla\cdot(h\bar{\boldsymbol{u}}) = 0,\\
&(\partial_t + \bar{\boldsymbol{u}}\cdot\nabla)\bar{\boldsymbol{u}} + \tfrac{1}{3}\boldsymbol{u}_\sigma\cdot\nabla\boldsymbol{u}_\sigma + f\hat{\boldsymbol{z}}\times\bar{\boldsymbol{u}} + \tfrac{1}{2}h^{-1}\nabla(h^2\bar{\vartheta}) = 0,\\
&(\partial_t + \bar{\boldsymbol{u}}\cdot\nabla)\boldsymbol{u}_\sigma + \boldsymbol{u}_\sigma\cdot\nabla\bar{\boldsymbol{u}} + f\hat{\boldsymbol{z}}\times\boldsymbol{u}_\sigma + \tfrac{1}{2}h\nabla\bar{\vartheta} = 0.
\end{aligned}\right\} \tag{7.10}$$

The energy density of this reduced system is $E = \frac{1}{2}h\bar{\boldsymbol{u}}^2 + \frac{1}{6}h\boldsymbol{u}_\sigma^2 + \frac{1}{2}h^2\bar{\vartheta}$, whereas the Casimir integrals of motion have the form $\mathscr{C} = \iint h\,A(\bar{\vartheta})$, where $A(\bar{\vartheta})$ is completely arbitrary. The free energy is then

$$\mathscr{E}_f = \frac{1}{2}\iint_D \left\{h\bar{\boldsymbol{u}}^2 + \tfrac{1}{3}h\boldsymbol{u}_\sigma{}^2 + \left(\bar{\vartheta}^{1/2}h - g_r^{1/2}H\right)^2\right\}, \tag{7.11}$$

which is non-negative definite, and differs with Young's integral of motion (7.9) in the sign of the second term.

8. Waves

In order to investigate the solutions of the new system, I will start with the simplest problem: linear waves superimposed on a state of no motion. This will be chosen as the reference state of §6 for a problem without topography, $h_0 = 0$. Equations (6.6), with $\bar{U} = 0$, and (6.4) imply that $\bar{\Theta} - \Theta_\sigma$ is a constant, whereas equations (6.6) and (6.5) imply that $\bar{\Theta} + \Theta_\sigma$ is a constant. Consequently, all fields in the reference state are uniform, and the linear wave is defined by

$$\begin{pmatrix}\bar{\vartheta}\\ \vartheta_\sigma\\ h\\ \bar{\boldsymbol{u}}\\ \boldsymbol{u}_\sigma\end{pmatrix} = \begin{pmatrix} g_r\\ \frac{1}{2}N_r^2H\\ H\\ \boldsymbol{0}\\ \boldsymbol{0}\end{pmatrix} + \varepsilon\begin{pmatrix}\bar{\vartheta}'(\boldsymbol{x},t)\\ \vartheta_\sigma'(\boldsymbol{x},t)\\ h'(\boldsymbol{x},t)\\ \bar{\boldsymbol{u}}'(\boldsymbol{x},t)\\ \boldsymbol{u}_\sigma'(\boldsymbol{x},t)\end{pmatrix} + O(\varepsilon^2), \tag{8.1}$$

where g_r, N_r^2, and H are constants, in the range $g_r > \frac{1}{2}N_r^2H > 0$. Upon substitution in (3.9) the following set is obtained to $O(\varepsilon)$:

$$\left.\begin{aligned}
&\partial_t\bar{\vartheta}' + \tfrac{1}{6}N_r^2H\,\nabla\cdot\boldsymbol{u}_\sigma' = 0,\\
&\partial_t\vartheta_\sigma' = 0,\\
&\partial_t h' + H\,\nabla\cdot\bar{\boldsymbol{u}}' = 0,\\
&\partial_t\bar{\boldsymbol{u}}' + f\hat{\boldsymbol{z}}\times\bar{\boldsymbol{u}}' + (g_r - \tfrac{1}{6}N_r^2H)\nabla h' + \tfrac{1}{2}H\nabla(\bar{\vartheta}' - \tfrac{1}{3}\vartheta_\sigma') = 0,\\
&\partial_t\boldsymbol{u}_\sigma' + f\hat{\boldsymbol{z}}\times\boldsymbol{u}_\sigma' + \tfrac{1}{4}N_r^2H\,\nabla h' + \tfrac{1}{2}H\nabla\bar{\vartheta}' = 0.
\end{aligned}\right\} \tag{8.2}$$

The second term in the equation for $\bar{\vartheta}'$ is the σ-advection of the reference buoyancy $\bar{\mu}'\Theta_\sigma$ since

$$\bar{\mu}' = \tfrac{1}{3}\nabla\cdot\boldsymbol{u}_\sigma'. \tag{8.3}$$

It is easy to show that the linearized system has an integral of motion which is the lowest-order term (quadratic) of $\mathscr{E}_f$ from (6.7).

A time-varying eigensolution of (8.2) will have $\vartheta_\sigma' = 0$; assuming a vertical normal mode structure

$$\begin{pmatrix} g_r h'\\ H\bar{\vartheta}'\end{pmatrix} = \begin{pmatrix} y_1\\ y_2\end{pmatrix} p^c(\boldsymbol{x},t), \qquad \begin{pmatrix}\bar{\boldsymbol{u}}'\\ \boldsymbol{u}_\sigma'\end{pmatrix} = \begin{pmatrix} x_1\\ x_2\end{pmatrix}\boldsymbol{u}^c(\boldsymbol{x},t), \tag{8.4}$$

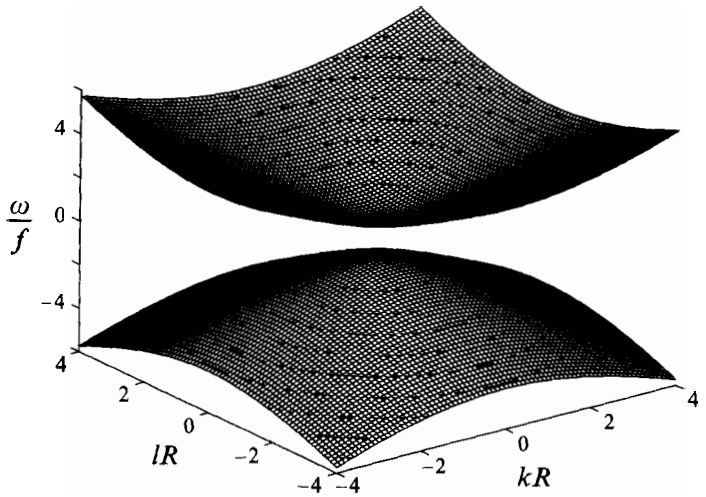

FIGURE 5. Non-dimensionalized dispersion relation for Poincaré waves: ω/f *vs.* (kR, lR), where f is the Coriolis parameter and $R = c/|f|$; the parameter c may have one of two values (see figure 7), depending on the vertical structure of the wave (see figure 8).

upon substitution in (8.2) it follows that $p^c(\boldsymbol{x}, t)$ and $\boldsymbol{u}^c(\boldsymbol{x}, t)$ satisfy the linearized shallow water equations

$$\left.\begin{aligned} \partial_t p^c + c^2 \nabla \cdot \boldsymbol{u}^c &= 0, \\ \partial_t \boldsymbol{u}^c + f\hat{\boldsymbol{z}} \times \boldsymbol{u}^c + \nabla p^c &= 0, \end{aligned}\right\} \tag{8.5}$$

if c^2 and $(x_1\ x_2)^T$ are an eigenvalue and eigenvector of

$$g_r H \begin{pmatrix} 1 - S/3 & S/6 \\ S/2 & S/6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = c^2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{8.6}$$

where $(y_1,\ y_2) = c^{-2} g_r H(x_1,\ \frac{1}{3} S\, x_2)$ and

$$S = \tfrac{1}{2} N_r^2 H / g_r \quad (0 < S < 1). \tag{8.7}$$

The solutions of (8.5) are well known (Pedlosky 1979; Gill 1982). For instance, at mid-latitudes assuming a common $\exp(\mathrm{i} \int \boldsymbol{k} \cdot \mathrm{d}\boldsymbol{x} - \mathrm{i}\omega t)$ phase dependence for all perturbation fields, the eigensolutions are found to be Poincaré

$$\omega^2 = f^2 + \boldsymbol{k}^2 c^2 \tag{8.8}$$

(see figure 5) and Rossby

$$\omega = \frac{\hat{\boldsymbol{z}} \cdot \nabla f \times \boldsymbol{k}}{\boldsymbol{k}^2 + f^2/c^2} \tag{8.9}$$

waves (see figure 6); with coasts, there are also Kelvin waves with celerity c. Similar results are obtained in the equatorial β-plane or the sphere.

The same solutions are found in the exact three-dimensional system (2.7), the only difference being that in the continuous case there are infinitely many vertical eigenmodes. More precisely,

$$\boldsymbol{u}' \propto \boldsymbol{u}^c(\boldsymbol{x}, t) \cos[m(\sigma - 1)], \quad m = \frac{N_r H}{2c}, \tag{8.10}$$

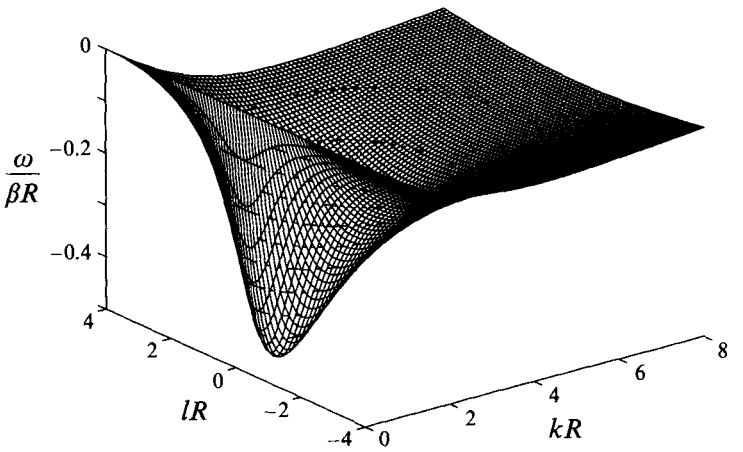

FIGURE 6. Non-dimensionalized dispersion relation for Rossby waves: $\omega/(\beta R)$ *vs.* (kR, lR), where β is the northward gradient of the Coriolis parameter.

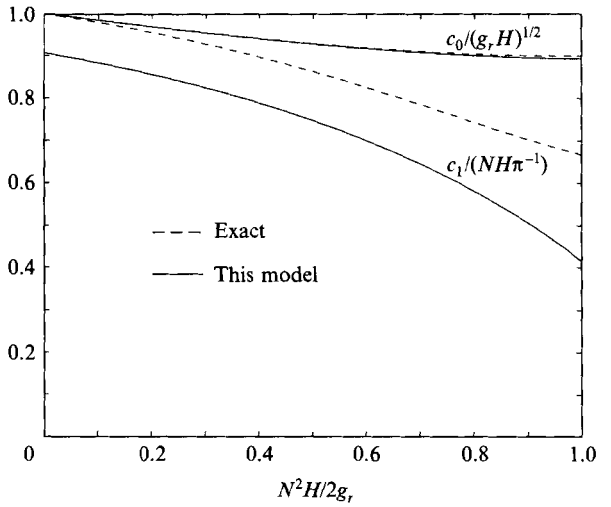

FIGURE 7. Separation constant c for both eigenmodes of the model of this paper (solid lines) compared with the first two eigenvalues of the continuously stratified model (dashed lines), as a function of the stratification N_r^2 of the reference state. In the reduced gravity case, a value of $N_r^2 H/2g_r = 1$ is huge because it implies a vanishing buoyancy jump at the base of the active layer relative to the density of the lower layer.

where m is determined by

$$m\tan(2m) = \frac{S}{1-S}; \tag{8.11}$$

see Appendix C.

Both values of the parameter c (which in the limit without rotation is the celerity of gravity waves) for the present model are compared in figure 7 with the first two eigenvalues for the exact continuously stratified model (8.11), calculated for the same stratification N_r^2 and buoyancy jump at the interface between both layers $g_r - \frac{1}{2}N_r^2 H$ (see, for instance, Ripa 1986). The comparison is excellent for the first mode and good for the second one, except for very strong stratifications, e.g. for $S = 1$ which corresponds to a vanishing buoyancy jump at $\sigma = -1$. Notice that typical values of

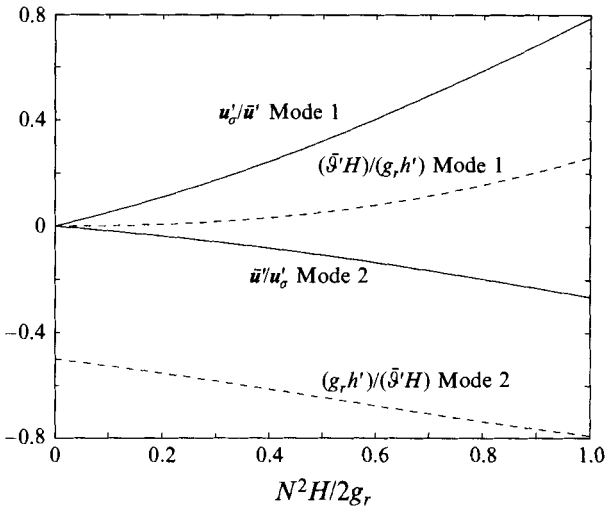

FIGURE 8. Vertical structure of the two eigenmodes of the model of this paper, as a function of the stratification N_r^2 of the reference state. Solid lines: ratio of the mean $\bar{u}'$ and shear u'_σ wave velocities. For the first mode (upper curve in figure 7) $|u'_\sigma| < |\bar{u}'|$ and therefore the velocity $\bar{u}' + \sigma u'_\sigma$ has no reversal with depth ($-1 \leqslant \sigma \leqslant 1$). For the second mode (lower curve in figure 7), there is always a reversal, closer to the free boundary (i.e. the surface in the rigid bottom case or the interface in the reduced gravity case). Dashed lines: ratio of the contributions to the horizontal force due to the gradient of density and the gradient of layer thickness. In the first mode, the second one dominates $g_r \nabla h' \gg H \nabla \bar{\vartheta}'$. The neutral modes of the IL^{0}PEM (symbolized by the 'horizontal' directions in figure 4*b*) correspond to $g_r h' = -\frac{1}{2}\bar{\vartheta}' H$, which is the $N_r^2 = 0$ point in the curve of the second mode.

S are quite small. For instance, the restratification scale of Tandon & Garrett (1994) gives $S = \frac{1}{2}(\nabla\bar{\vartheta}/\bar{\vartheta})^2 R_d^2 \ll 1$, where $R_d^2 = g_r H/f^2$ is very close to the deformation radius for the first mode.

In addition to the (Poincaré and Rossby) waves, there is a neutral solution, $\omega = 0$, of the linearized equations (8.2)

$$\left.\begin{aligned} \bar{u}' = 0, \quad u'_\sigma = 0, \\ (g_r - \tfrac{5}{12}N_r^2 H)h' - \tfrac{1}{6}H\vartheta'_\sigma = 0, \\ \tfrac{1}{2}N_r^2 h' + \bar{\vartheta}' = 0. \end{aligned}\right\} \tag{8.12}$$

Notice that this neutral mode has a finite free energy, unlike the case of the *force-compensating mode of the* IL^{0}PEM; see discussion after (7.5).

The two eigenvectors of (8.6) correspond to waves without and with flow reversal in the vertical, as shown in figure 8. For the first mode $0 \leqslant u_\sigma/\bar{u} < 0.8$, as a function of $N_r^2 H/2g_r$, and consequently $\bar{u} + \sigma u_\sigma$ cannot change sign with σ. On the other hand, for the second mode $-0.27 < \bar{u}/u_\sigma \leqslant 0$, resulting in a sign reversal of $\bar{u} + \sigma u_\sigma$ at a value of σ between 0 and 0.27 (depending on the value of $N_r^2 H/2g_r$). Recall that the waves of the present model have $\vartheta'_\sigma \equiv 0$: figures 7 and 8 show that in the limit of weak stratification, the first and second modes tend to the waves and the *force-compensating mode* of the IL^{0}PEM, namely

$$S = N_r^2 H/2g_r \to 0 : \begin{cases} c^2 = g_r H, & u'_\sigma = \bar{\vartheta}' = 0 \\ c^2 = N_r^2 H^2/12, & \bar{u}' = 0,\ 2g_r h' = -\bar{\vartheta}' H. \end{cases} \tag{8.13}$$

9. Baroclinic instability

In the previous section it was shown that the new model represents reasonably well the free waves superimposed on a motionless reference state. It is important to see how well the model reproduces known instability problems, particularly those of a parallel flow $\bar{\boldsymbol{u}} = \bar{U}\hat{\mathbf{x}}$ and $\boldsymbol{u}_\sigma = U_\sigma \hat{\mathbf{x}}$ in the basic state, instead of the resting state used to define the wave in (8.1). I will take both $\bar{U}$ and U_σ as constant, since the model is expected to deal correctly with instabilities related to the horizontal shear; the real test is in problems where the vertical structure of the perturbation is a crucial component.

I will compare the results of this model not only with those of a more exact calculation, but also with those of a two-layer HLPEM, which has about the same number of vertical degrees of freedom as the present one. The parameters of the two-layer model are chosen so that in the equilibrium state both layers have the same thickness (equal to $H/2$) and their buoyancies are equal to $\vartheta_1 = g_r + N_r^2 H/4$ and $\vartheta_2 = g_r - N_r^2 H/4$. The basic state whose stability is studied has uniform velocities given by $U_1 = \bar{U} + U_\sigma/2$ and $U_2 = \bar{U} - U_\sigma/2$. These values correspond to averaging of the buoyancy and velocity fields in each layer. The two-layer HLPEM has 6 fields, whereas the one-layer IL^{0}PEM and IL^{1}PEM have 4 and 7 fields respectively.

One particular area in which the new model can be tested is that of Kelvin–Helmholtz instability, which is posed using $f = 0$ in the evolution equations. However, the present model does not produce unstable normal modes for this problem. This is not surprising, since Kelvin–Helmholtz instability is related to the existence of a critical level, which is impossible to reproduce with the linear vertical profiles used here. I will then concentrate on baroclinic instability, choosing a non-vanishing value for the Coriolis parameter. With $f \neq 0$, the fields $(\bar{\vartheta}, \vartheta_\sigma, h)$ must have uniform y-slopes, in the basic state, to geostrophically balance the velocities $\bar{U}$ and U_σ.

Consider then the basic state and perturbation defined by

$$\begin{pmatrix} \bar{\vartheta} \\ \vartheta_\sigma \\ h \\ \bar{\boldsymbol{u}} \\ \boldsymbol{u}_\sigma \end{pmatrix} = \begin{pmatrix} g_r + Ay \\ \tfrac{1}{2} N_r^2 H_r + By \\ H_r + Cy \\ \bar{U}\hat{\mathbf{x}} \\ U_\sigma \hat{\mathbf{x}} \end{pmatrix} + \varepsilon \begin{pmatrix} \bar{\vartheta}'(\boldsymbol{x}, t) \\ \vartheta_\sigma'(\boldsymbol{x}, t) \\ h'(\boldsymbol{x}, t) \\ \bar{\boldsymbol{u}}'(\boldsymbol{x}, t) \\ \boldsymbol{u}_\sigma'(\boldsymbol{x}, t) \end{pmatrix} + O(\varepsilon^2). \tag{9.1}$$

Notice that there are three coefficients (A, B, C) to balance two velocities $(\bar{U}, U_\sigma)$: there is an extra degree of freedom, indicating that the new model can represent a non-uniform density field at both the top and bottom boundaries, something that, for instance, the two-layer model cannot do. In order to make a comparison with the other models, I will choose $B = 0$ in (9.1); however, this 'extra degree of freedom' might yield new instability phenomena, to be explored elsewhere. The other two coefficients are then calculated from the geostrophic balance of the basic flow, namely

$$\begin{pmatrix} f\bar{U} \\ fU_\sigma \end{pmatrix} + \begin{pmatrix} H_r/2 & g_r - N_r^2 H_r/6 \\ H_r/2 & N_r^2 H_r/4 \end{pmatrix} \begin{pmatrix} A \\ C \end{pmatrix} = 0,$$

where $\beta = 0$ is assumed, for simplicity.

If the perturbation is assumed proportional to $\exp \mathrm{i}\,[k(x - ct) \pm ly]$, then there are two non-dimensional measures of the horizontal wavenumber $|\boldsymbol{k}| = (k^2 + l^2)^{1/2}$, namely

$$\kappa_0^2 := (k^2 + l^2)\, g_r H_r / f_0^2, \quad \kappa_1^2 := (k^2 + l^2)\, N_r^2 H_r^2 / f_0^2 \equiv 2S\kappa_0^2. \tag{9.2}$$

For simplicity, I will assume a weak stratification ($S \ll 1$), which implies $\kappa_0 \gg \kappa_1$.

In this limit, κ_0 equals $|\boldsymbol{k}|$ scaled by the first Rossby radius, whereas that non-dimensionalization of $|\boldsymbol{k}|$ with the second deformation radius gives $\kappa_1/\sqrt{12}$ in the present model or κ_1/π for the exact one; see (8.13) and figure 7. I will now present the main results for the eigenvalue c; a detailed calculation is presented in Ripa (1995), where β-effects are also discussed.† Notice that if $U_\sigma = 0$ then c must be real (independent of $\bar{U}$) because in this case the free energy $\mathscr{E}_f$ is a positive definite integral of motion in the deviation from the basic state, which is therefore stable (see for instance Ripa 1990; Shepherd 1990). Long and short perturbations correspond formally to the limits $N_r^2 \to 0$ and $g_r \to \infty$.

9.1. *Long perturbations:* $\kappa_0 = O(1)$

With $\kappa_0 = O(1)$, i.e. $\kappa_1 \sim 0$, the eigenvalue c is given by

$$\mathrm{IL}^1 : \begin{cases} c = \left[\bar{U}(1+2\kappa_0^2) - U_\sigma \pm \Delta^{1/2}\right] / \left(2+2\kappa_0^2\right), \\ \Delta = \left(U_\sigma - \bar{U}\right)^2 - 4\kappa_0^2 U_\sigma \left[\bar{U} + \left(1+\kappa_0^2\right) U_\sigma/3\right]. \end{cases} \tag{9.3}$$

Growing and decaying normal modes ($\mathrm{Im}\,(kc) \neq 0$) are found for wavenumbers such that

$$\kappa_0^2 > \left[1 + 3\left(\frac{\bar{U}}{U_\sigma}\right)^2\right]^{1/2} - \frac{1}{2} - \frac{3}{2}\frac{\bar{U}}{U_\sigma}. \tag{9.4}$$

(Notice that the critical wavenumber vanishes for $\bar{U} = U_\sigma$.) Young & Chen (1995) obtained exactly the same result using the subinertial model of Young (1994), in the limit in which buoyancy is mixed but momentum is not ($f\tau \to \infty$; see the Introduction and table 1).‡ This result is supported by the more exact calculation of Fukamachi *et al.* (1995), which also used a linear profile for the basic flow but allowed for an arbitrary depth dependence for the perturbation. Their calculation corresponds to $\bar{U} = U_\sigma$, and their eigenvalue coincides with the value in (9.3) evaluated for $\bar{U} = U_\sigma$. The two-layer model, on the other hand, gives

$$\text{2-layer} : \begin{cases} c = \left[\bar{U}(1+2\kappa_0^2) - U_\sigma/2 \pm \Delta^{1/2}\right] / \left(2+2\kappa_0^2\right), \\ \Delta = \left(U_\sigma/2 - \bar{U}\right)^2 - 2\kappa_0^2 U_\sigma \left[\bar{U} + \left(1+\kappa_0^2\right) U_\sigma/2\right], \end{cases} \tag{9.5}$$

and growing and decaying normal modes require

$$\kappa_0^2 > \left[\frac{1}{2} + 2\left(\frac{\bar{U}}{U_\sigma}\right)^2\right]^{1/2} - \frac{1}{2} - \frac{\bar{U}}{U_\sigma}.$$

Finally, the IL^0 model (i.e. inhomogeneous layer but depth-independent fields) gives

$$\mathrm{IL}^0 : \begin{cases} c = \left[\bar{U}(1+2\kappa_0^2) - U_\sigma \pm \Delta^{1/2}\right] / \left(2+2\kappa_0^2\right), \\ \Delta = \left(U_\sigma - \bar{U}\right)^2 - 4\kappa_0^2 U_\sigma \bar{U}, \end{cases} \tag{9.6}$$

where U_σ is implicit, through the thermal wind balance (see Fukamachi *et al.* 1995; Ripa 1995*a*; and the 'slab model' limit ($f\tau \to 0$) in Young & Chen 1995). Instability

† Since these are low-frequency instabilities, the results can be obtained by using the geostrophic balance and by working with the equations for $\left(\bar{\vartheta}, \vartheta_\sigma, \bar{q}, q_\sigma\right)$.

‡ That model does not have $\boldsymbol{u}_\sigma$ as an explicit variable but, rather, is implicit through the thermal wind relation. The parameters (U, C, μ) of Young & Chen (1995) corresponding to this result are $U = \bar{U}$, $C = -U_\sigma$, and $\mu = 0$.

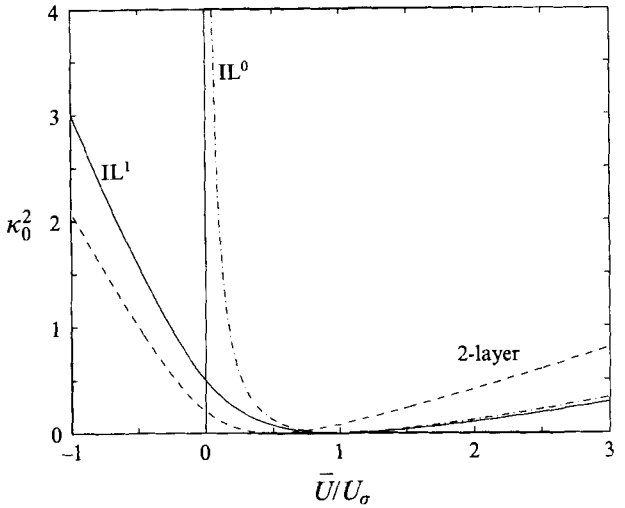

FIGURE 9. Minimum wavenumber, $\kappa_0^2 = (k^2 + l^2)g_r H_r/f^2$, for baroclinic instability, as a function of the ratio of the mean velocity $\bar{U}$ to the velocity shear U_σ in the basic state. (The magnitude of U_σ equals one half the velocity difference between the top and bottom of the active layer.) IL^1 : present model and $f\tau \to \infty$ limit of Young & Chen (1995); at $\bar{U} = U_\sigma$ the dispersion relation coincides with the exact result of Fukamachi *et al.* (1995). IL^0: vertically averaged model (Ripa 1995*a*, and $f\tau \to 0$ limit of Young & Chen 1995); this model predicts stability for $\bar{U}/U_\sigma \leqslant 0$.

is now restricted to

$$U_\sigma \bar{U} > 0,\ \kappa_0^2 > \frac{1}{4}\left(\left(\frac{U_\sigma}{\bar{U}}\right)^{1/2} - \left(\frac{\bar{U}}{U_\sigma}\right)^{1/2}\right)^2. \tag{9.7}$$

The main mechanism for this instability is the interaction of a topographic Rossby wave and a neutral *force-compensating mode* which are the modes obtained by making $\bar{U} = 0$ in (9.6) (Ripa 1995*a*).

The instability regions in parameter space of the different models considered here are compared in figure 9, whereas an example of a dispersion relation is shown in figure 10. The IL^0 and IL^1 models seem to have similar results for (large) positive values of $\bar{U}/U_\sigma$. Moreover, for $\bar{U} = U_\sigma$ they compare well with the exact IL$^\infty$ model. The reason for the discrepancy between the IL^0 and IL^1 models for negative $\bar{U}/U_\sigma$ constitues an interesting research subject.

9.2. *Short perturbations:* $\kappa_1 = O(1)$

For short wavelengths (compared with the deformation radius of the first vertical mode) top and bottom boundaries are effectively rigid, and one can compare with the exact result of the Eady problem (Gill 1982), which gives

$$\text{Eady}: c = \bar{U} \pm U_\sigma \left[\left\{(2/\kappa_1) - \tanh(\kappa_1/2)\right\}\left\{(2/\kappa_1) - \coth(\kappa_1/2)\right\}\right]^{1/2}. \tag{9.8}$$

A very important property of this dispersion relation is the short-wavenumber cutoff of the instability, namely c is real for $\kappa_1 > \kappa_1^{crit} \approx 2.3994$. (In particular, the flow is stable in a narrow enough channel.)

However, the calculation of Young & Chen (1995) gives only the $\kappa_0 \to \infty$ limit of (9.3), namely $c = \bar{U} \pm \mathrm{i}U_\sigma\sqrt{1/3}$ for all κ_1, which is the correct value of (9.8) for $\kappa_1 = 0$, but fails to provide a short-wavenumber cutoff. This represents a serious

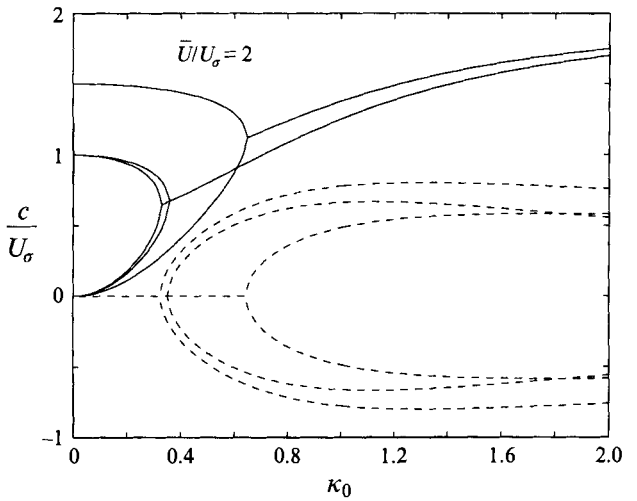

FIGURE 10. Dispersion relation of long perturbations in the baroclinic instability problem. The solid (dashed) line shows the real (imaginary) part of the complex 'phase speed' c. The three cases correspond to the IL^1, IL^0 and 2-layer models, ordered by the bifurcation point (see figure 9).

limitation of Young's (1994) model, because in a fully nonlinear calculation there may be a spurious growth of short perturbations. Fukamachi *et al.* (1995) also obtain only the $\kappa_1 = 0$ value of (9.8), and for $\bar{U} = U_\sigma$, namely $c = U_\sigma \left(1 \pm i\sqrt{1/3}\right)$.

The new model fares much better, since, in fact, it gives

$$\mathrm{IL}^1 : c = \bar{U} \pm U_\sigma \frac{\left(\left(\kappa_1^4 - 144\right)/3\right)^{1/2}}{\kappa_1^2 + 12}. \tag{9.9}$$

This formula is exact for $\kappa_1 = 0$ and has a high-wavenumber cutoff corresponding to $\kappa_1^{crit} = \sqrt{12} \approx 3.4641$, which is about 44% too large. The asymptotic value of $c - \bar{U}$ as $\kappa_1 \to \infty$ is about 43% too small. These discrepancies are not unexpected, because the present model restricts the vertical structure to a linear function σ whereas the solution of Eady's problem has a growing perturbation with an exponential depth dependence, namely a combination of $\exp\left(\pm\kappa_1\sigma/2\right)$.

Finally, the two-layer model gives

$$\text{2-layer} : c = \bar{U} \pm U_\sigma \frac{\left(\kappa_1^4 - 64\right)^{1/2}}{2\left(\kappa_1^2 + 8\right)}.$$

The short-wavenumber cutoff $\kappa_1^{crit} = \sqrt{8} \approx 2.8284$ is an overestimate by about 18% of Eady's one, the value of $c - \bar{U}$ for $\kappa_1 = 0$ is also overestimated (by about 15%) and the asymptotic value of $c - \bar{U}$ as $\kappa_1 \to \infty$ is one-half that of Eady's. The dispersion relations are compared in figure 11.

In summary, the new model does not produce Kelvin–Helmholtz unstable modes (which require critical levels) but gives a fair representation of baroclinic instability. The latter includes a short-wavenumber cutoff similar to that encountered in Eady's problem. The free waves, discussed in the last section, and these results add confidence to the validity of IL^1PEM. The importance of the new model does not lie, of course, in reproducing known results (with a limited vertical structure) but in the possibility

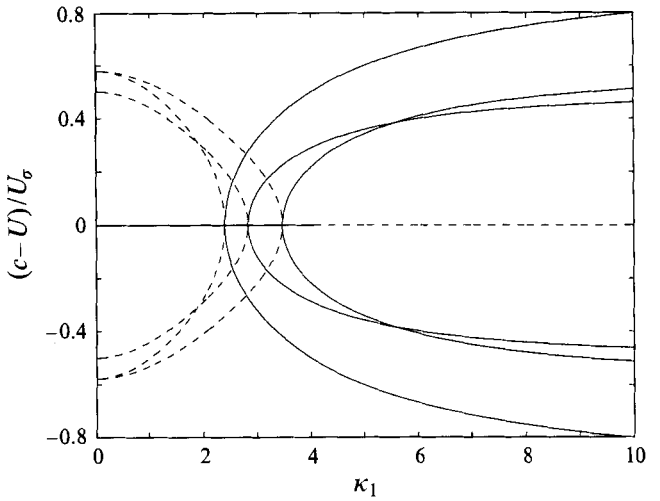

FIGURE 11. As in figure 10, for the short perturbations in the baroclinic instability problem. The three cases (ordered by the bifurcation point) correspond to the (exact) Eady solution and the approximations of the 2-layer (Phillips model) and IL^1 models to the problem.

of adding a new ingredient to these and other calculations, by means of the possibility of a horizontally varying stratification.

10. Conclusions

A new type of ocean circulation model proposed here has an active layer with variable thickness, within which horizontal velocity and density (or temperature and salinity) have a linear profile with depth, with coefficients which are function of horizontal position and time (the generalization to many layers is trivial). This system represents an advantage over the (fixed) level models, in the sense that even with one layer it is able to support Kelvin, Poincaré and short Rossby waves, i.e. phenomena for which horizontal divergence is important. The waves are a very good approximation to those of the first two vertical normal modes of the continuously stratified system.

The new model has an advantage over the simpler shallow water equations (with homogeneous layers) in that it is able to accommodate thermodynamic processes because it is not restricted to a constant density. Finally, the new system also represents an advantage over the now classical layer models with only lateral variations of density, because it can sustain the thermal wind balance (i.e. the vertical shear of the current associated with a horizontal density gradient) as well as distinguish between the densities at the top and bottom boundaries of the active layer (e.g. the temperatures at the ocean surface and at the base of the upper layer).

The model developed here satisfies the correct conservation equation for total energy, momentum, volume, mass and buoyancy variance, as well as the correct evolution equation for the three-dimensional vorticity field. However, this model does not include the law of (density) potential vorticity conservation. In order to have it, it would be necessary to deal with the vertical curvature of the velocity field. The integrals of motion are used to construct a 'free energy', which is conserved as well as being quadratic to the lowest order and positive definite in the deviation from a reference state with uniform Brunt-Väisälä frequency and (at most) a uniform flow.

The implication of this conservation law is that the free evolution of the system cannot lead to an 'explosion' in which the dynamical fields grow without limit. As a comparison, the model of Young (1994) can experience that type of explosion from a state of rest (through the mechanism of 'explosive resonant interactions') or by an unarrested baroclinic instability (i.e. which does not show the saturation seen in the quasi-geostrophic model; see Shepherd 1988).

The new model is not restricted to low frequencies, for instance Dr Amit Tandon (personal communication, 1994) has shown that the IL^{1}PEM gives the same results for the mixed layer restratification problem as the exact calculation of Tandon & Garrett (1994), which is a high-frequency problem, i.e. with time scales of the order of the inertial period $2\pi/|f|$. The real test of the validity of the present model will be its ability to reproduce observations in more complicated applications, particularly if it can achieve the same results than models with more complicated vertical structures.

This work was supported by CICESE's normal funding and by CONACyT (México) under grant 1282-T9204. The hospitality of Nora Velasco Fuentes during the early part of this work is sincerely appreciated. The final version of this work was written while on sabbatical leave at the Center of Earth and Ocean Sciences of the University of Victoria. I would like to thank CONACyT and NSERC (Canada) that made that visit possible. Rosalie Rutka and F. Javier Beroń V. have been very helpful with corrections to the paper.

Appendix A. Ertel's theorem in σ-coordinates

Operating with $\hat{\boldsymbol{z}} \times \partial_\sigma$ over the $\boldsymbol{u}$-equation in (2.7) written in the form

$$\mathrm{D}\boldsymbol{u}/\mathrm{D}t + f\hat{\boldsymbol{z}} \times \boldsymbol{u} + \tilde{\nabla} p + \vartheta \tilde{\nabla} v = 0 \tag{A 1}$$

and using $\partial_\sigma p = -\vartheta \partial_\sigma v$ one gets

$$\mathrm{D}(h\boldsymbol{q})/\mathrm{D}t - \hat{\boldsymbol{z}} \times (\partial_\sigma v \tilde{\nabla}\vartheta - \partial_\sigma \vartheta \tilde{\nabla} v) = -\partial_\sigma \boldsymbol{u} \cdot \tilde{\nabla}(\hat{\boldsymbol{z}} \times \boldsymbol{u}) + f\partial_\sigma \boldsymbol{u} + h\boldsymbol{q}\partial_\sigma \mu. \tag{A 2}$$

The terms on the right-hand side of this equation can be shown, using $\boldsymbol{a} \cdot \nabla(\hat{\boldsymbol{z}} \times \boldsymbol{b}) = \hat{\boldsymbol{z}} \times \boldsymbol{a}\, \nabla \cdot \boldsymbol{b} - \hat{\boldsymbol{z}} \times \boldsymbol{a} \cdot \nabla \boldsymbol{b} - \boldsymbol{a}\, \hat{\boldsymbol{z}} \cdot \nabla \times \boldsymbol{b}$, to be equal to $h(\boldsymbol{q}\tilde{\nabla} \cdot \boldsymbol{u} + \boldsymbol{q}\partial_\sigma \mu - \mathscr{L}\boldsymbol{u})$. Finally, using the continuity equation $\mathrm{D}h/\mathrm{D}t + h\tilde{\nabla} \cdot \boldsymbol{u} + h\partial_\sigma \mu = 0$ (4.18) follows. On the other hand, taking the curl of the $\boldsymbol{u}$-equation written in the form

$$\tilde{\partial}_t \boldsymbol{u} + \mu \partial_\sigma \boldsymbol{u} + hq\hat{\boldsymbol{z}} \times \boldsymbol{u} + \tilde{\nabla} b = v\tilde{\nabla}\vartheta \tag{A 3}$$

one gets

$$\tilde{\partial}_t(hq) + \tilde{\nabla} \cdot (hq\boldsymbol{u}) + \partial_\sigma(hq\mu) = h\boldsymbol{q} \cdot \tilde{\nabla}\mu + [v, \vartheta]^\sigma. \tag{A 4}$$

Using $\partial_t h + \tilde{\nabla} \cdot (h\boldsymbol{u}) + h\partial_\sigma \mu = 0$ (4.17) then follows. Finally, if $\mathrm{D}s/\mathrm{D}t = \dot{s}$, then

$$\mathrm{D}(\partial s)/\mathrm{D}t + (\partial \boldsymbol{u}) \cdot \nabla s + (\partial \mu)\partial_\sigma s = \partial \dot{s}, \tag{A 5}$$

where ∂ is ∇ or ∂_σ. Calculating $\nabla s \cdot (4.18) + \partial_\sigma s\,(4.17)$ and adding it to $\boldsymbol{q} \cdot \mathrm{D}(\nabla s)/\mathrm{D}t + q\mathrm{D}(\partial_\sigma s)/\mathrm{D}t$ it is easily found that $\nabla s \cdot (\mathscr{L}\boldsymbol{u}) + \partial_\sigma s(\mathscr{L}\mu)$ cancels with the terms originating in $(\partial \boldsymbol{u}) \cdot \nabla s + (\partial \mu)\partial_\sigma s$. Theorem (4.16) is then proved using the expansion (2.4) for the horizontal Jacobian. □

Appendix B. Hamiltonian structure

The Poisson tensor for the new model can be written, for simplicity, in the form $\boldsymbol{J} = \boldsymbol{J}_1 + \boldsymbol{J}_2$, where

$$\boldsymbol{J}_1 = -\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \nabla\cdot(& 0 \\ 0 & 0 & \nabla & \bar{q}\hat{\mathbf{z}}\times & q_\sigma\hat{\mathbf{z}}\times \\ 0 & 0 & 0 & q_\sigma\hat{\mathbf{z}}\times & 3\bar{q}\hat{\mathbf{z}}\times \end{pmatrix}, \tag{B 1}$$

$$\boldsymbol{J}_2 = -\begin{pmatrix} 0 & 0 & 0 & h^{-1}(\nabla\bar{\vartheta})\cdot & h^{-1}\nabla\cdot(\vartheta_\sigma \\ 0 & 0 & 0 & h^{-1}(\nabla\vartheta_\sigma)\cdot & 3h^{-1}(\nabla\bar{\vartheta})\cdot \\ 0 & 0 & 0 & 0 & 0 \\ -h^{-1}\nabla\bar{\vartheta} & -h^{-1}\nabla\vartheta_\sigma & 0 & 0 & h^{-1}\boldsymbol{u}_\sigma\nabla\cdot(\\ \vartheta_\sigma\nabla(h^{-1} & -3h^{-1}\nabla\bar{\vartheta} & 0 & \nabla(h^{-1}\boldsymbol{u}_\sigma\cdot & 0 \end{pmatrix}. \tag{B 2}$$

The Hamiltonian is the total energy, whose variational derivatives are given by

$$\frac{\delta\mathscr{H}}{\delta\bar{\vartheta}} = h(h_0 + \tfrac{1}{2}h),\ \frac{\delta\mathscr{H}}{\delta\vartheta_\sigma} = -\tfrac{1}{6}h^2,\ \frac{\delta\mathscr{H}}{\delta h} = \bar{b},\ \frac{\delta\mathscr{H}}{\delta\bar{\boldsymbol{u}}} = h\bar{\boldsymbol{u}},\ \frac{\delta\mathscr{H}}{\delta\boldsymbol{u}_\sigma} = \tfrac{1}{3}h\boldsymbol{u}_\sigma. \tag{B 3}$$

Using these and $\boldsymbol{J}$ in (5.1) the equation of the new model (3.9) is easily obtained. □

Given two 'admissible' functionals of state, $\mathscr{A}$ and $\mathscr{B}$, their Poisson bracket is defined by $\{\mathscr{A},\mathscr{B}\} := (\delta\mathscr{A}/\delta\varphi^a)\boldsymbol{J}^{ab}(\delta\mathscr{B}/\delta\varphi^b)$, and must be antisymmetric: $\{\mathscr{A},\mathscr{B}\} = -\{\mathscr{B},\mathscr{A}\}$; with (B 1–B 2) it follows that 'admissible' functionals $\mathscr{A}[\varphi^a,\boldsymbol{x},t]$ must satisfy

$$\frac{\delta\mathscr{A}}{\delta\bar{\boldsymbol{u}}}\cdot\hat{\boldsymbol{n}} = \frac{\delta\mathscr{A}}{\delta\boldsymbol{u}_\sigma}\cdot\hat{\boldsymbol{n}} = 0 \quad @\ \boldsymbol{x}\in\partial D. \tag{B 4}$$

The Hamiltonian in (5.5) is admissible as can be seen from (B 3) and the no-flow boundary condition, $\bar{\boldsymbol{u}}\cdot\hat{\boldsymbol{n}} = \boldsymbol{u}_\sigma\cdot\hat{\boldsymbol{n}} = 0 \quad @\ \boldsymbol{x}\in\partial D$. The momentum $\mathscr{M}$ is admissible if the coasts ∂D_j are zonal. In order to be conserved, it is further needed that $\{\mathscr{M},\mathscr{H}\} = 0 \Leftrightarrow \partial h_0/\partial x = 0$.

In order for $\mathscr{C}[\varphi^a]$ to be a Casimir, $\boldsymbol{J}^{ab}(\delta\mathscr{C}/\delta\varphi^b) = 0\ \forall a$, the following equalities must be satisfied:

$$\hat{\bar{\boldsymbol{u}}}\cdot\nabla\bar{\vartheta} + \hat{\boldsymbol{u}}_\sigma\cdot\nabla\vartheta_\sigma = 0, \tag{B 5}$$

$$\hat{\bar{\boldsymbol{u}}}\cdot\nabla\vartheta_\sigma + 3\hat{\boldsymbol{u}}_\sigma\cdot\nabla\bar{\vartheta} = 0, \tag{B 6}$$

$$\nabla\cdot\hat{\bar{\boldsymbol{u}}} = 0, \tag{B 7}$$

$$-\hat{\bar{\vartheta}}\nabla\bar{\vartheta} - \hat{\vartheta}_\sigma\nabla\vartheta_\sigma + h\nabla\hat{h} + h\bar{q}\hat{\mathbf{z}}\times\hat{\bar{\boldsymbol{u}}} + hq_\sigma\hat{\mathbf{z}}\times\hat{\boldsymbol{u}}_\sigma + \boldsymbol{u}_\sigma\nabla\cdot\hat{\boldsymbol{u}}_\sigma = 0, \tag{B 8}$$

$$\vartheta_\sigma\nabla(h^{-1}\hat{\bar{\vartheta}}) - 3h^{-1}\hat{\vartheta}_\sigma\nabla\bar{\vartheta} + \nabla(h^{-1}\boldsymbol{u}_\sigma\cdot\hat{\bar{\boldsymbol{u}}}) + q_\sigma\hat{\mathbf{z}}\times\hat{\bar{\boldsymbol{u}}} + 3\bar{q}\hat{\mathbf{z}}\times\hat{\boldsymbol{u}}_\sigma = 0, \tag{B 9}$$

where the short notation $\hat{\phi}_a := \delta\mathscr{C}/\delta\varphi^a$ has been used. It is easy to show that $\mathscr{I}_n = \iint h\overline{\vartheta^n}$ $(n = 0, 1, 2)$ are admissible and Casimirs. There do not seem to be more Casimir integrals of motion for the new system, though. For instance, consider the possibility $\mathscr{C} = \iint h\overline{\mathscr{L}\vartheta}$. It can be shown that conditions (B 5–B 8) are satisfied for this functional, but that (B 9) is not; (B 9) gives $\nabla\cdot(\bar{\boldsymbol{q}}\cdot\nabla\vartheta_\sigma + q_\sigma\vartheta_\sigma) = 0$, i.e. $(\mathscr{L}\vartheta)_\sigma =$ constant, a condition which is not preserved by the dynamics. A similar conclusion is reached with other putative Casimirs, e.g. $\iint h\overline{\mathscr{L}\vartheta^2}$: all conditions are satisfied, except for the last one (B 9).†

† From the point of view of geometrical mechanics, this means that these are not Casimirs, but generating functionals that transform along the 'direction' of $\boldsymbol{u}_\sigma$.

Appendix C. Vertical normal modes in σ-coordinates

Linearizing (2.7) around a state of no motion and with buoyancy $\Theta(\sigma)$ (i.e. $\vartheta = \Theta + \varepsilon\vartheta' + O(\varepsilon^2)$, etc.), gives

$$\left.\begin{aligned} &\tilde{\partial}_t\vartheta' + \tfrac{1}{2}N_r^2 H\,\mu' = 0,\\ &\tilde{\partial}_t h' + H\tilde{\nabla}\cdot\boldsymbol{u}' + H\partial_\sigma\mu' = 0,\\ &\tilde{\partial}_t\boldsymbol{u}' + f\hat{\boldsymbol{z}}\times\boldsymbol{u}' + \tilde{\nabla}p' + \tfrac{1}{2}(1-\sigma)\Theta\tilde{\nabla}h' = 0,\\ &\partial_\sigma p' = \tfrac{1}{2}H\vartheta' + \tfrac{1}{2}h'\Theta. \end{aligned}\right\} \tag{C 1}$$

where $\frac{1}{2}N_r^2(\sigma)H = \mathrm{d}\Theta/\mathrm{d}\sigma$. A (vertical) normal mode has a structure

$$\left.\begin{aligned} \boldsymbol{u}' &= \boldsymbol{u}^c(\boldsymbol{x},t)\,\left[\mathrm{d}F(\sigma)/\mathrm{d}\sigma\right],\\ p' &= p^c(\boldsymbol{x},t)\,\left[\mathrm{d}F(\sigma)/\mathrm{d}\sigma - \tfrac{1}{2}(1-\sigma)\,\Theta(\sigma)G(-1)\right], \end{aligned}\right\} \tag{C 2}$$

where $G(\sigma) = \Theta^{-1}(\mathrm{d}F/\mathrm{d}\sigma)$. The second term between square brackets guarantees that $p' = 0$ @ $\sigma = -1$, as required by (2.8). In order for the third equation in (C 1) to give $\partial_t\boldsymbol{u}^c + f\hat{\boldsymbol{z}}\times\boldsymbol{u}^c + \tilde{\nabla}p^c = 0$, it is necessary that

$$h' = p^c(\boldsymbol{x},t)G(-1). \tag{C 3}$$

The last and first equations in (C 1) then give

$$\left.\begin{aligned} \vartheta' &= \tfrac{1}{2}\,p^c(\boldsymbol{x},t)\,N_r^2(\sigma)\,H\,\left[(\sigma-1)H^{-1}G(-1) - c^{-2}F(\sigma)\right],\\ \mu' &= \partial_t p^c(\boldsymbol{x},t)\,\left[(\sigma-1)H^{-1}G(-1) - c^{-2}F(\sigma)\right], \end{aligned}\right\} \tag{C 4}$$

where the structure function $F(\sigma)$ must satisfy the following differential equation, whilst the boundary conditions are needed for μ' to vanish at $\sigma = \pm 1$, as required by (2.8):

$$\begin{cases} \mathrm{d}^2F(\sigma)/\mathrm{d}\sigma^2 + \tfrac{1}{4}N_r^2H^2c^{-2}F(\sigma) = 0, & -1 < \sigma < 1\\ F = 0, & \text{at } \sigma = 1\\ \Theta H\,F = -2c^2\left[\mathrm{d}F/\mathrm{d}\sigma\right], & \text{at } \sigma = -1 \end{cases} \tag{C 5}$$

Finally, the second equation (C 1) gives the second modal equation, $\partial_t p^c + c^2\nabla\cdot\boldsymbol{u}^c = 0$. The case discussed in the main text corresponds to

$$\Theta(\sigma) = g_r + \sigma\,\frac{N_r^2 H}{2} = g_r(1+\sigma\,S), \tag{C 6}$$

where g_r, N_r^2 and H are constants.

REFERENCES

ANDERSON, D. 1984 An advective mixed-layer model with applications to the diurnal cycle of the low-level East African jet. *Tellus* **36**, 278–291.

ANDERSON, D. & MCCREARY, J. 1985 Slowly propagating disturbances in a coupled ocean-atmosphere model. *J. Atmos. Sci.* **42**, 615–629.

ARNOL'D, V. 1965 Condition for nonlinear stationary plane curvilinear flows of an ideal fluid. *Dokl. Akad. Nauk. USSR* **162**, 975–978 (English transl: *Soviet Maths* **6**, 773–777, 1965).

ARNOL'D, V. 1966 On an apriori estimate in the theory of hydrodynamical stability. *Izv. Vyssh. Uchebn. Zaved. Matematika* **54**, 3–5 (English transl. *Am. Math. Soc. Transl.* Series 2, **79**, 267–269, 1969).

BALMASEDA, M. A., ANDERSON, D. & DAVEY, M. 1994 ENSO prediction using a dynamical ocean model coupled to statistical atmospheres. *Tellus* **46A**, 497–511.

CHERNIAWSKY, J. & HOLLOWAY, G. 1991 An upper-ocean general circulation model of the North Pacific: preliminary experiments. *Atmos.-Ocean* **29**, 737–784.

CHERNIAWSKY, J. Y., YUEN, C. W., LIN, C. A. & MYSAK, L. A. 1990 Numerical experiments with a wind- and buoyancy-driven two-and-a-half-layer upper ocean model. *J. Geophys. Res.* **95**, 16149–16167.

DARBY, M. & WILLMOTT, A. 1993 Vacillating ocean gyres: an instability mechanism in a thermodynamic reduced gravity ocean model. *Tellus*, submitted.

DEMPSEY, P. & ROTUNNO, R. 1988 Topographic generation of mesoscale vortices in mixed-layer models. *J. Atmos Sci.* **45**, 2961–2978.

FUKAMACHI, Y., MCCREARY, J. & PROEHL, J. 1995 Instability of density fronts in layer & continuously stratified models. *J. Geophys. Res.* **100**, 2559–2577.

GILL, A. 1982 *Atmosphere-Ocean Dynamics*. Academic Press.

KUENY, C. S. & MORRISON, P. J. 1994 Nonlinear instability and chaos in plasma wave-wave interactions. I. Introduction. *Tech. Rep. IFSR 682, The University of Texas.*

LAVOIE, R. L. 1972 A mesoscale numerical model of lake-effect storms. *J. Atmos. Sci.* **29**, 1025–1040.

MCCREARY, J. P., FUKAMACHI, Y. & KUNDU, P. 1991 A numerical investigation of jets and eddies near an eastern ocean boundary. *J. Geophys. Res.* **96**, 2515–2534.

MCCREARY, J. & KUNDU, P. 1988 A numerical investigation of the Somali current during the southwest monsoon. *J. Mar. Res.* **46**, 25–58.

MCCREARY, J. P., LEE, H. & ENFIELD, D. 1989 The response of the coastal ocean to strong offshore winds: with application to circulations in the gulfs of Tehuantepec and Papagayo. *J. Mar. Res.* **47**, 81–109.

MCCREARY, J. & LU, P. 1994 Interaction between the subtropical and equatorial ocean circulations: The subtropical cell. *J. Phys. Oceanogr.* **24**, 466–497.

MCCREARY, J. & YU, Z. 1992 Equatorial dynamics in a $2\frac{1}{2}$-layer model. *Prog. Oceanogr.* **29**, 61–132.

MCINTYRE, M. & SHEPHERD, T. 1987 An exact local conservation theorem for finite-amplitude disturbances to non-parallel shear flows, with remarks on Hamiltonian structure and on Arnol'd's stability theorems. *J. Fluid Mech.* **181**, 527–565.

MORRISON, P. J. & KOTSCHENREUTHER, M. 1990 The free energy principle, negative energy modes, and stability. In *Nonlinear World: IV Intl Workshop on Nonlinear and Turbulent Processes in Physics, Kiev, USSR, 1989* (ed. V. G. Bar'yakhtar, V. M. Chernousenko, N. S. Erokhin, A. G. Sitenko & V. E. Zakharov), pp. 910–932.

PEDLOSKY, J. 1979 *Geophysical Fluid Dynamics*. Springer.

RIPA, P. 1986 Evaluation of vertical structure functions for the analysis of oceanic data. *J. Phys. Oceanogr.* **16**, 223–232.

RIPA, P. 1990 Positive, negative and zero wave energy and the flow stability problem, in the eulerian and lagrangian-eulerian descriptions. *Pure Appl. Geophys.* **133**, 713–732.

RIPA, P. 1991 General stability conditions for a multi-layer model. *J. Fluid Mech.* **222**, 119–137.

RIPA, P. 1992*a* Wave energy-momentum and pseudoenergy-momentum conservation for the layered quasi-geostrophic instability problem. *J. Fluid Mech.* **235**, 379–398.

RIPA, P. 1992*b* Instability of a solid-body rotating vortex in a two layer model. *J. Fluid Mech.* **242**, 395–417.

RIPA, P. 1993*a* Conservation laws for primitive equations models with inhomogeneous layers. *Geophys. Astrophys. Fluid Dyn.* **70**, 85–111.

RIPA, P. 1993*b* Hamiltonian GFD. In *Geometrical Methods in Fluid Mechanics* (ed. R. Salmon), pp. 332–336. WHOI.

RIPA, P. 1993*c* Integrals of motion and stability-instability properties of ocean models. In *Trends in Oceanography* (ed. J. Menon), pp. 141–151, Research Trends (India).

RIPA, P. 1995*a* Linear waves in a one-layer ocean model with thermodynamics. *J. Geophys. Res. C*, in press.

RIPA, P. 1995*b* Low frequency approximation of a one-layer ocean model with thermodynamics *Rev. Mex. Fís.* submitted.

RIPA, P. 1995*c* A low-frequency one-layer model with variable velocity shear and stratification, *J. Phys. Oceanogr.* **41**, in press.

SCHOPF, P. & CANE, M. 1983 On ecuatorial dynamics, mixed layer physics and sea surface temperature. *J. Phys. Oceanogr.* **13**, 917–935.

SHEPHERD, T. 1988 Rigorous bounds on the nonlinear saturation of instabilities to parallel shear flows. *J. Fluid Mech.* **196**, 291–322.

SHEPHERD, T. 1990 Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics. *Adv. Geophys.* **32**, 287–338.

SZOEKE, R. DE & RICHMAN, J. 1984 On wind-driven mixed layers with strong horizontal gradients – A theory with applications to coastal upwelling. *J. Phys. Oceanogr.* **14**, 364–377.

TANDON, A. & GARRETT, C. 1994 Mixed layer restratification due to a horizontal density gradient. *J. Phys. Oceanogr.* **24**, 1419–1424.

VANNESTE, J. 1995 Explosive resonant interaction of Rossby waves and stability of multilayer quasi-geostrophic flow. *J. Fluid Mech.*, in press.

YOUNG, W. 1994 The subinertial mixed layer approximation. *J. Phys. Oceanogr.* **24**, 1812–1826.

YOUNG, W. & CHEN, L. 1995 Baroclinic instability and thermoaline gradient alignment in the mixed layer. *J. Phys. Oceanogr.*, submitted.