J. Fluid Mech. (2003), *vol.* 474, *pp.* 245–273.
DOI: 10.1017/S0022112002002586 Printed in the United Kingdom

The nonlinear evolution of zonally symmetric equatorial inertial instability

By STEPHEN D. GRIFFITHS†

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK

(Received 25 October 2001 and in revised form 8 August 2002)

The inertial instability of equatorial shear flows is studied, with a view to understanding observed phenomena in the Earth's stratosphere and mesosphere. The basic state is a zonal flow of stratified fluid on an equatorial β-plane, with latitudinal shear. The simplest self-consistent model of the instability is used, so that the basic state and the disturbances are zonally symmetric, and a vertical diffusivity provides the scale selection. We study the interaction between the inertial instability, which takes the form of periodically varying disturbances in the vertical, and the mean flow, where 'mean' is a vertical mean.

The weakly nonlinear regime is investigated analytically, for flows with an arbitrary dependence on latitude. An amplitude equation of the form $\mathrm{d}A/\mathrm{d}t = A - k^2 A \int |A|^2 \mathrm{d}t$ is derived for the disturbances, and the evolving stability properties of the mean flow are discussed. In the final steady state, the disturbances vanish, but there is a persistent mean flow change that stabilizes the flow. However, the magnitude of the mean flow change depends strongly on the initial conditions, so that the system has a long memory. The analysis is extended to include the effects of Rayleigh friction and Newtonian cooling, destroying the long-memory property.

A more strongly nonlinear regime is investigated with the help of numerical simulations, extending the results up to the point where the instability leads to density contour overturning. The instability is shown to lead to a homogenization of $f\overline{Q}$ around the initially unstable region, where f is the Coriolis parameter, and $\overline{Q}$ is the vertical mean of the potential vorticity. As the instability evolves, the line of zero $\overline{Q}$ moves polewards, rather than equatorwards as might be expected from a simple self-neutralization argument.

1. Introduction

Inertial instability arises from an imbalance of forces that occurs in rotating fluid systems. As diagnosed by Rayleigh (1917), the instability may occur if the magnitude of the absolute angular momentum decreases away from the rotation axis. The instability has been extensively studied since G. I. Taylor's celebrated laboratory experiments of the flow of a viscous fluid between concentric rotating cylinders (Taylor 1923; Drazin & Reid 1981, and references therein). In such a configuration, without background rotation or stratification, the instability is often referred to as centrifugal instability.

It has long been recognized that inertial instability might occur within the Earth's

† Present address: Department of Mathematical Sciences, Loughborough University, Loughborough LE11 3TU, UK.

atmosphere and oceans, both on planetary scales (e.g. Dunkerton 1981; Hitchman *et al.* 1987; Tomas & Webster 1997), and on smaller scales (e.g. Potylitsin & Peltier 1998). The inertial instability of planetary scale flows can be investigated by considering a thin layer of inviscid stably stratified fluid moving relative to the planet in a zonal flow, with speed $u(\phi)$ at latitude ϕ. It can be shown (e.g. Bowman & Shepherd 1995) that such a flow is stable to zonally symmetric perturbations if everywhere $fQ \geqslant 0$. Here, $f = 2\Omega \sin\phi$ is the Coriolis parameter, where Ω is the planetary angular velocity, and Q is the potential vorticity, which here takes the same sign as the component of the absolute vorticity normal to the planetary surface.

If $fQ < 0$, we say there is a region of *anomalous vorticity*, and the flow may be unstable. If the latitudinal shear is non-zero at the equator, then Q will be non-zero at the equator, and hence Q will be anomalous on one side or other of the equator. Such unstable flows may occur in the equatorial stratosphere and mesosphere, where at solstice the zonal flow has little vertical shear, but does have a strong and persistent cross-equatorial shear associated with the large-scale meridional circulation. The shear weakens through the equinoxes and appears again, with the opposite sign, at the following solstice, so that the equatorial region of anomalous vorticity is located in the winter hemisphere.

The first theoretical study of inertial instability in this equatorial context was made by Dunkerton (1981). He studied the simplest possible problem likely to be of interest: the linear instability of a uniformly stratified fluid with constant latitudinal shear and Prandtl number unity, on an equatorial β-plane. Both the basic flow and the disturbances were taken to be zonally symmetric. The linear instability theory was subsequently extended to systems with Prandtl number not equal to unity (Dunkerton 1982; see also Edwards & Richards 1999), to zonally asymmetric modes on a zonally symmetric basic state (Boyd & Christidis 1982; Dunkerton 1983), and to systems with a zonally asymmetric basic state (Dunkerton 1993; Clark & Haynes 1996).

The linear theory predicts that the instability will take the form of overturning cells in the meridional plane. Indeed, there are observations of structures in the equatorial upper stratosphere and lower mesosphere bearing a strong resemblance to unstable modes of the linear stability theory (e.g. Hitchman *et al.* 1987; Hayashi, Shiotani & Gille 1998; Smith & Riese 1999). Despite some unresolved questions about the preferred vertical scale of the instability (Griffiths 2002), the structures are attributable to inertial instability of the cross-equatorial shear, and it appears that there are several large-amplitude inertial instability events per year. Similar-looking cellular structures in the equatorial oceans may also be due to inertial instability (e.g. Hua, Moore & Le Gentil 1997).

Even though the theory of linear equatorial inertial instability is well developed for simple flow configurations, there are still unresolved questions as to the nonlinear evolution of the instability. Will the system reach a final steady state? If so, will it be a state of rest? What happens to the mean flow? Can anything be said about the inertial stability of the evolved system? Do the nonlinear structures resemble the linear structures? Previous studies have gone some way towards answering such questions. Hua *et al.* (1997) performed some numerical simulations of the nonlinear instability of a uniform shear flow, with oceanic applications in mind. In a system with Rayleigh friction, Newtonian cooling and enhanced horizontal diffusion, they showed how the instability equilibrates to a state with baroclinic zonal jets. Zhao & Ghil (1991) performed a weakly nonlinear analysis for a two-layer system,

with atmospheric applications in mind, although their formulation artifically restricts attention to modes less unstable than the fasting growing modes. Other relevant contributions have been made for the related case of symmetric baroclinic instability in the atmosphere (Thorpe & Rotunno 1989), and for an unstratified f-plane inertial instability, with oceanic applications in mind (Shen & Evans 1998). However, there remains a need to make some more general conclusions about the nonlinear instability.

Here, the nonlinear evolution of equatorial inertial instability is considered for a basic state with arbitrary latitudinal shear. The flow is zonally symmetric and hydrostatic, with Prandtl number unity, and there is no horizontal diffusion, Rayleigh friction or Newtonian cooling. In § 2, the stability problem is formulated, and the linear theory is reviewed. A set of non-dimensionalized equations is introduced, with just one control parameter, so that the linear stability properties can be simply classified. In § 3, the weakly nonlinear evolution is described, based on the results of a perturbation analysis. A detailed study of the mean-flow change and the evolving stability properties is made, providing a clear insight into the nature of nonlinear inertial instability. In § 4, the theoretical predictions are compared with numerical simulations of the nonlinear problem for a particular basic flow, and the simulations are extended into the moderately nonlinear regime. The nature of the mean-flow change is described. In § 5, the effects of Rayleigh friction and Newtonian cooling upon the nonlinear evolution are considered. The resulting analytic theory qualitatively reproduces the numerical results of Hua *et al.* (1997). In § 6, there is a summary and discussion of the results, and of their possible relevance to inertial instability in the equatorial stratosphere and mesosphere.

2. Formulation and linear stability

2.1. *The governing equations*

We study a stratified fluid in three dimensions, with background rotation, on an unbounded equatorial β-plane. We use a set of Cartesian coordinates, with the x-axis pointing eastwards, the y-axis pointing northwards, and the z-axis as a local vertical coordinate. The x- and z-directions are either periodic or unbounded. The fluid flow speeds are u, v, and w in the x-, y- and z-directions, respectively. We define a fluctuating buoyancy acceleration $\sigma(\boldsymbol{x}, t)$ such that the density ρ takes the form $\rho = \rho_{00} + \rho_0(z) - (\rho_{00}/g)\sigma(\boldsymbol{x}, t)$, where ρ_{00} is a constant reference value, and $\rho_0(z)$ describes the background stable stratification.

For the remainder of this study, we suppose that (i) the flow is zonally symmetric, i.e. $\partial/\partial x = 0$, (ii) the flow is hydrostatic, (iii) the flow can be modelled using the Boussinesq approximation (expected to be valid to the extent that the inertial instability occurs in thin vertical layers), and (iv) the diffusion coefficients for heat and momentum have the same value ν. Furthermore, for the structures to be studied, the vertical length scale is much smaller than the horizontal length scale, so we approximate the diffusion operator $\nu\nabla^2$ as $\nu\partial^2/\partial z^2$.

The restriction to zonal symmetry allows us to introduce a streamfunction $\psi(y, z, t)$ for the flow in the meridional plane. We write $v = -\partial\psi/\partial z$ and $w = \partial\psi/\partial y$, so that the continuity equation $\nabla \cdot \boldsymbol{u} = 0$ is automatically satisfied. It is then possible to eliminate the pressure from the formulation by introducing an equation for the x-component ξ of the relative vorticity. Using the traditional approximation for

large-scale geophysical flows (Phillips 1966), the complete set of equations is

$$\frac{\partial u}{\partial t} + \boldsymbol{u}\cdot\nabla u + f\frac{\partial \psi}{\partial z} = \nu\frac{\partial^2 u}{\partial z^2}, \tag{2.1a}$$

$$\frac{\partial \xi}{\partial t} + \boldsymbol{u}\cdot\nabla \xi - f\frac{\partial u}{\partial z} - \frac{\partial \sigma}{\partial y} = \nu\frac{\partial^2 \xi}{\partial z^2}, \tag{2.1b}$$

$$\frac{\partial \sigma}{\partial t} + \boldsymbol{u}\cdot\nabla \sigma + N^2\frac{\partial \psi}{\partial y} = \nu\frac{\partial^2 \sigma}{\partial z^2}, \tag{2.1c}$$

where $f = \beta y$, $\xi = \partial^2\psi/\partial z^2$, and $N^2(z) = -(g/\rho_{00})\partial\rho_0/\partial z > 0$.

We will study the stability of a flow $u = u_0(y)$, $\psi = \sigma = 0$, with constant $\partial\rho_0/\partial z$ and therefore constant buoyancy frequency N, to zonally symmetric disturbances u', ψ' and σ'. An important quantity will be the potential vorticity Q, which we define as

$$Q = \frac{(f\boldsymbol{e}_z + \nabla\times\boldsymbol{u})\cdot\nabla\rho}{\partial\rho_0/\partial z}, \tag{2.2}$$

where the denominator, $\partial\rho_0/\partial z$, is a constant included simply for convenience. The basic state potential vorticity $Q_0(y) = f - \partial u_0/\partial y$ is simply equal to the vertical component of the absolute vorticity.

2.2. *The case of uniform shear*

Separable solutions of the linear system can be found analytically when the basic flow has uniform latitudinal shear. It is useful to review this problem, originally solved by Dunkerton (1981), since these results motivate the theoretical framework to be used in the following sections.

We write the uniform shear flow as $u_0 = \Lambda y$, and take $\Lambda > 0$ for definiteness. This leads to a band of anomalous vorticity occupying $0 < y < \Lambda/\beta$. Ignoring terms quadratic in disturbance amplitude in (2.1*a*)–(2.1*c*), we look for latitudinally bounded disturbances of the form $\psi'(y,z,t) = \mathrm{Re}\{\Psi(y)\,\mathrm{e}^{imz+st}\}$. Thus, Re($s$) is the growth rate, and m is the vertical wavenumber, which we take to be positive without loss of generality. The system then reduces to a Hermite equation for the latitudinal structure $\Psi(y)$ with a discrete set of eigenmodes, the most unstable of which has dispersion relation

$$(s + \nu m^2)^2 = \frac{\Lambda^2}{4} - \frac{N\beta}{m}. \tag{2.3}$$

Introducing a coordinate Y shifted to the centre of the unstable region, and the equatorial Rossby length L, defined by

$$Y = y - \frac{\Lambda}{2\beta}, \quad L(m) = \left(\frac{N}{\beta m}\right)^{1/2}, \tag{2.4a,b}$$

the most unstable mode is of the form

$$\psi' \propto \mathrm{Re}\{\mathrm{e}^{-(1/2)(Y/L)^2}\mathrm{e}^{(imz+st)}\}, \quad \sigma' \propto \mathrm{Re}\{Y\,\mathrm{e}^{-(1/2)(Y/L)^2}\mathrm{e}^{(imz+st)}\}.$$

There is an overturning motion in vertically stacked cells of height π/m, and of latitudinal scale L. In the regions of rising and sinking air there are density perturbations, which are largest at $Y = \pm L$, i.e. at $y = \frac{1}{2}\Lambda/\beta \pm L$. The density perturbations give rise to the temperature perturbations observed in events of inertial instability in the equatorial atmosphere. Using (2.3), we see that all growing modes have $m > 4N\beta/\Lambda^2$,

or equivalently using (2.4*b*) that $L < \frac{1}{2}\Lambda/\beta$, so the maximum density perturbations both lie within the initially unstable region.

If $\nu = 0$, then the maximum possible growth rate $s = \frac{1}{2}\Lambda$ is attained as $m \to \infty$. Thus, for a real fluid we must consider the effect of diffusion, however small. As pointed out by Dunkerton (1981), the diffusivity may be interpreted as an eddy diffusivity either resulting from an earlier occurrence of the instability itself on much smaller scales, or from some background turbulence. As is well known, the concept of eddy diffusivity and eddy Prandtl number are of questionable validity, but the present state of knowledge does not seem to justify anything more elaborate. For this study, we can choose to interpret the diffusivity as a molecular diffusivity, or to cautiously interpret the diffusivity as an eddy diffusivity.

If $\nu \neq 0$, and since $\kappa = \nu$, the stability properties can be characterized by a single non-dimensional parameter ϵ, given by

$$\epsilon = \left(\frac{2\nu N^2\beta^2}{\Lambda^5}\right)^{1/3}. \tag{2.5}$$

There is instability if and only if $\epsilon < \epsilon_c = 5^{-5/6}$. Then viewed as a function of m, the growth rate has a single maximum at $m = m_*$, where

$$m_*(\epsilon) = \frac{1}{\epsilon}\frac{N\beta}{\Lambda^2}\hat{m}_*(\epsilon).$$

Here, $\hat{m}_*(\epsilon)$ is a non-dimensional function, satisfying $\hat{m}_*(0) = 1$, and monotonically increasing to $5^{1/6}$ as $\epsilon \to \epsilon_c$ (Griffiths 2002). Thus, the most unstable wavenumber scales like

$$m_s = \frac{1}{\epsilon}\frac{N\beta}{\Lambda^2} = \left(\frac{N\beta}{2\nu\Lambda}\right)^{1/3}. \tag{2.6}$$

We see the competition between stratification and the β-effect encouraging smaller vertical scales, and diffusion and cross-equatorial shear encouraging larger vertical scales.

2.3. *A more general formulation for arbitrary shear flows*

The scales for inertial instability on a uniform shear flow can be used to construct a set of non-dimensionalized equations for studying equatorial inertial instability in more general cases. We retain constant $\partial\rho_0/\partial z$, but allow $u_0(y)$ to have variable shear. However, we assume that near to the equator the shear has a typical magnitude Λ that is important in determining the growth rate and structure of the inertially unstable modes. Thus, we expect that (i) the growth rate will scale like $\frac{1}{2}\Lambda$, (ii) the modes will be centred a distance of order Λ/β from the equator, (iii) the most unstable vertical wavenumber will scale like m_s, given by (2.6), and (iv) the latitudinal structure will scale like $L_s = L(m_s)$, given by (2.4*b*). Hence, we introduce non-dimensional coordinates $\hat{y}$, $\hat{z}$ and $\hat{t}$ defined by

$$\hat{y} = \frac{y - \frac{1}{2}(\Lambda/\beta)\hat{y}_c}{L_s}, \quad \hat{z} = m_s z, \quad \hat{t} = \tfrac{1}{2}\Lambda t.$$

We fix $\hat{y}_c$, assumed to be of $O(1)$, to be the centre of the unstable region, i.e. the latitude at which $(-fQ_0)$ is maximized.

It is also worth introducing non-dimensional physical variables $\hat{u}$, $\hat{\psi}$, $\hat{\sigma}$, $\hat{v}$, $\hat{w}$ and $\hat{\xi}$, which we do according to

$$u = L_s \frac{\Lambda}{2}\hat{u}, \quad \psi = \frac{L_s}{m_s}\frac{\Lambda}{2}\hat{\psi}, \quad \sigma = \frac{N^2}{m_s}\hat{\sigma}, \quad v = L_s\frac{\Lambda}{2}\hat{v}, \quad w = \frac{1}{m_s}\frac{\Lambda}{2}\hat{w}, \quad \xi = m_s L_s \frac{\Lambda}{2}\hat{\xi}.$$

Thus, when $\hat{v} = 1$, a fluid parcel will move one non-dimensional latitudinal distance in one non-dimensional time unit. The other scalings are chosen to be compatible with this. We also introduce a non-dimensionalized potential vorticity $\hat{Q} = 2Q/\Lambda$, so that

$$\hat{Q} = \left(\hat{f} - \frac{\partial \hat{u}}{\partial \hat{y}}\right)\left(1 + \frac{\partial \hat{\sigma}}{\partial \hat{z}}\right) + \frac{\partial \hat{u}}{\partial \hat{z}}\frac{\partial \hat{\sigma}}{\partial \hat{y}}, \tag{2.7}$$

from (2.2). With these scalings, the full set, (2.1a)–(2.1c), of zonally symmetric Boussinesq equations, with Prandtl number unity, is

$$\frac{\partial \hat{u}}{\partial \hat{t}} + \hat{\boldsymbol{u}} \cdot \hat{\nabla}\hat{u} + \hat{f}\frac{\partial \hat{\psi}}{\partial \hat{z}} = \epsilon \frac{\partial^2 \hat{u}}{\partial \hat{z}^2}, \tag{2.8a}$$

$$\frac{\partial \hat{\xi}}{\partial \hat{t}} + \hat{\boldsymbol{u}} \cdot \hat{\nabla}\hat{\xi} - \hat{f}\frac{\partial \hat{u}}{\partial \hat{z}} - 4\epsilon \frac{\partial \hat{\sigma}}{\partial \hat{y}} = \epsilon \frac{\partial^2 \hat{\xi}}{\partial \hat{z}^2}, \tag{2.8b}$$

$$\frac{\partial \hat{\sigma}}{\partial \hat{t}} + \hat{\boldsymbol{u}} \cdot \hat{\nabla}\hat{\sigma} + \frac{\partial \hat{\psi}}{\partial \hat{y}} = \epsilon \frac{\partial^2 \hat{\sigma}}{\partial \hat{z}^2}, \tag{2.8c}$$

where

$$\hat{f}(\hat{y}, \epsilon) = \hat{y}_c + 2\epsilon^{1/2}\hat{y}, \quad \hat{v} = -\frac{\partial \hat{\psi}}{\partial \hat{z}}, \quad \hat{w} = \frac{\partial \hat{\psi}}{\partial \hat{y}}, \quad \hat{\xi} = \frac{\partial^2 \hat{\psi}}{\partial \hat{z}^2}. \tag{2.9}$$

The system is characterized by just one non-dimensional parameter $\epsilon = (2\nu N^2 \beta^2/\Lambda^5)^{1/3}$, a ratio of the stabilizing influences in the system – diffusion, stratification, and the β-effect – to the destabilizing influence of cross-equatorial shear. We can view ϵ as a sort of non-dimensional diffusivity, and as ϵ decreases we expect the system to become more inertially unstable.

In this non-dimensionalization, the terms associated with the horizontal component of the Coriolis parameter, neglected under the traditional approximation used here, are of magnitude $\epsilon^{1/2}\Omega/N$. Since for instability we expect $\epsilon \lesssim 1$, it is reasonable to neglect these terms provided $\Omega/N \ll 1$. This condition will certainly be satisfied in the stratosphere and mesosphere. However, for studying oceanic inertial instability, it may not be justifiable to neglect these terms, as discussed by Hua *et al.* (1997).

2.4. *Linear stability revisited*

It is worth examining the linear stability problem within this formulation. Omitting the hats from the non-dimensional variables, as will be done from hereon, the non-dimensionalized linear equation for the disturbance streamfunction ψ' is $\mathscr{L}\psi' = 0$, where

$$\mathscr{L}\left(-\mathrm{i}\frac{\partial}{\partial z}, \frac{\partial}{\partial t}, \epsilon, \frac{\partial}{\partial y}, y\right) = 4\epsilon\frac{\partial^2}{\partial y^2} + \left(\frac{\partial}{\partial t} - \epsilon\frac{\partial^2}{\partial z^2}\right)^2 \frac{\partial^2}{\partial z^2} + fQ_0\frac{\partial^2}{\partial z^2}. \tag{2.10}$$

Here, $Q_0(y,\epsilon) = f - \partial u_0/\partial y$ is the basic state potential vorticity, which, like $f = y_c + 2\epsilon^{1/2}y$, will depend on ϵ as well as y. For latitudinally bounded disturbances of the form $\psi'(y,z,t) = \mathrm{Re}\{\Psi(y)\,\mathrm{e}^{imz+st}\}$, (2.10) reduces to

$$\mathscr{L}\left(m, s, \epsilon, \frac{\partial}{\partial y}, y\right)\Psi(y) = 0, \quad \Psi \to 0 \text{ as } |y| \to \infty, \tag{2.11}$$

where, from (2.10),

$$\mathscr{L}\left(m, s, \epsilon, \frac{\partial}{\partial y}, y\right) = 4\epsilon\frac{\partial^2}{\partial y^2} - m^2(s + \epsilon m^2)^2 - m^2 f Q_0. \tag{2.12}$$

Multiplying (2.11) by Ψ^*, the complex conjugate of Ψ, and integrating across the domain gives

$$(s + \epsilon m^2)^2 = \frac{\displaystyle\int_{-\infty}^{\infty}(-fQ_0)|\Psi|^2\mathrm{d}y - (4\epsilon/m^2)\int_{-\infty}^{\infty}|\mathrm{d}\Psi/\mathrm{d}y|^2\mathrm{d}y}{\displaystyle\int_{-\infty}^{\infty}|\Psi|^2\mathrm{d}y}. \tag{2.13}$$

We note that

(i) $(s + \epsilon m^2)^2$ is real. Thus, from (2.12), $\mathscr{L}(m, s, \epsilon, \partial/\partial y, y)$ is a real operator, and we can choose $\Psi(y)$ to be a real function.

(ii) If $fQ_0 > 0$ everywhere, then $(s + \epsilon m^2)^2 < 0$ and the flow is linearly stable to zonally symmetric disturbances.

(iii) If $fQ_0 < 0$ somewhere, then instability is possible, and the maximum possible growth rate is $((-fQ_0)_{max})^{1/2}$. However, if m is too small then the stratification inhibits the vertical motion necessary to make the parcel exchange, and $s < 0$. The stabilizing effects of stratification become negligible as $m \to \infty$, but if $\epsilon \neq 0$ then in this limit the growth rate decreases without bound owing to the vertical diffusion. Thus, for sufficiently large ϵ, it is possible that no unstable modes will exist. A region of anomalous vorticity need not be inertially unstable – diffusion is able to completely stabilize the system.

Furthermore, as shown in the Appendix, for the unstable modes, the maximum growth rate increases monotonically as ϵ decreases, and as $\epsilon \to 0$ there exists a mode such that $s^2 \to (-fQ_0)_{max}$. Since as $\epsilon \to \infty$ all vertical wavenumbers becomes stable, for some value of ϵ the system is neutrally stable.

2.5. *The dispersion relation*

Because of the particular form of $\mathscr{L}$, we can write

$$\mathscr{L}(m, s, \epsilon, \partial/\partial y, y)\Psi(y) = D(m, s, \epsilon)\Psi(y), \tag{2.14}$$

for some function $D(m, s, \epsilon)$. The dispersion relation is just $D(m, s, \epsilon) = 0$, although, in general, it is impossible to write down the function D. Equation (2.14) is particularly useful because it enables us to relate the unknown function D directly to the known linear operator $\mathscr{L}$, albeit via an unknown eigenfunction $\Psi(y)$. For instance, multiplying (2.14) by $\Psi(y)$, integrating across the domain, differentiating with respect to s,

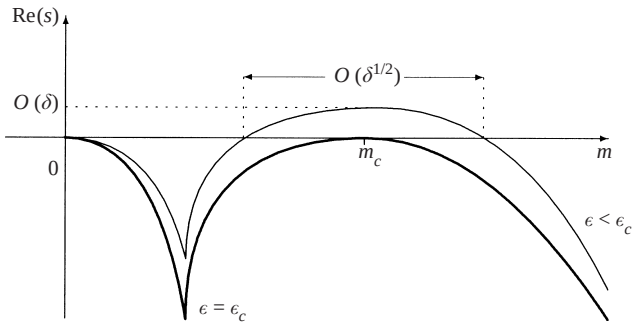

FIGURE 1. The dispersion relation. The thick line shows $s(m)$ at critical conditions $\epsilon = \epsilon_c$; there is a single global maximum at wavenumber m_c. The thin line shows $s(m)$ at weakly unstable conditions $\epsilon = \epsilon_c(1 - \delta/\delta_0)$; there is an envelope of unstable wavenumbers around $m = m_c$.

and using the self-adjoint property of $\mathscr{L}$ yields

$$D_s = \frac{\displaystyle\int_{-\infty}^{\infty} \Psi \mathscr{L}_s \Psi \, \mathrm{d}y}{\displaystyle\int_{-\infty}^{\infty} \Psi^2 \mathrm{d}y}, \tag{2.15}$$

where $D_s = \partial D/\partial s$, and $\mathscr{L}_s$ is the partial differential of $\mathscr{L}(m, s, \epsilon, \partial/\partial y, y)$ with respect to s. From (2.12), $\mathscr{L}_s = -2m^2(s + \epsilon m^2)$, independent of y, so that

$$D_s = -2m^2(s + \epsilon m^2). \tag{2.16}$$

Thus, without solving (2.11) for $\Psi(y)$, we have derived information about the function $D(m, s, \epsilon)$. Such information is useful in interpreting the dispersion relation (see § 3.1), and will prove essential to the weakly nonlinear analysis (see § 3.5).

3. The weakly nonlinear analysis

We now turn to understanding the nonlinear evolution of inertial instability. We first aim to understand the simplest possible problem, that is the nonlinear evolution when the system is only weakly unstable. In this regime, the parameter ϵ is just less than the critical value for stability, which might correspond to the existence of a weak horizontal shear, or a strong vertical diffusivity. Thus, we automatically exclude the study of strongly nonlinear instability. However, we are able to investigate this weakly nonlinear regime analytically, and the generality and completeness of our analysis will enable us to draw strong conclusions. At the outset it is worth noting that even though there might be a relatively strong vertical diffusion, there will be no horizontal diffusion within the system, so mean flow changes are brought about solely by nonlinear momentum advection.

3.1. *Weakly unstable conditions*

We continue to study the inertial instability of a zonal flow $u_0(y)$ with arbitrary dependence on latitude. As described in § 2.4, if $fQ_0 < 0$ somewhere, there is a critical value $\epsilon = \epsilon_c$ such that the linear growth rate $s \leqslant 0$ for all m, and such that $s = 0$ for some $m = m_c$. We assume that this critical wavenumber m_c is unique. The situation is illustrated in figure 1.

If we differentiate the dispersion relation $D(m, s, \epsilon) = 0$ with respect to m at fixed ϵ, viewing s as a function of m, we see that $\partial s/\partial m = -D_m/D_s$. Since $s(m)$ has a maximum at $m = m_c$ when $\epsilon = \epsilon_c$, and since from (2.16) $D_s(m_c, 0, \epsilon_c) = -2\epsilon_c m_c^4$, we must have $D_m(m_c, 0, \epsilon_c) = 0$. Similarly, $\partial^2 s/\partial m^2 = -D_{mm}/D_s$ at $m = m_c$, and, by the same reasoning, we must have $D_{mm}(m_c, 0, \epsilon_c) \leqslant 0$. We will assume that $D_{mm}(m_c, 0, \epsilon_c) < 0$. We summarize by writing

$$D(m_c, 0, \epsilon_c) = 0, \quad D_m(m_c, 0, \epsilon_c) = 0, \quad D_{mm}(m_c, 0, \epsilon_c) < 0. \tag{3.1}$$

For the remainder of the analysis, we will be interested in the flow regime when ϵ is just less than ϵ_c, that is when ϵ is just marginally smaller than the critical value for neutral stability. We set $\epsilon = \epsilon_c(1 - \delta/\delta_0)$, with $\delta \ll 1$, where

$$\delta_0 = \epsilon_c \left.\frac{D_\epsilon}{D_s}\right|_{(m_c,0,\epsilon_c)} > 0 \tag{3.2}$$

is a useful normalizing factor (using (3.1) and the definition of $\mathscr{L}(m, s, \epsilon)$, it is possible to check that $\delta_0 > 0$). Expanding the dispersion relation in a Taylor series about the point $m = m_c$, $s = 0$, $\epsilon = \epsilon_c$, and using (3.1), we see that for wavenumbers m close to m_c,

$$s \approx \delta - \tfrac{1}{2}(m - m_c)^2 \left.\frac{D_{mm}}{D_s}\right|_{(m_c,0,\epsilon_c).} \tag{3.3}$$

The parameter δ measures the strength of the instability, and is the leading-order growth rate of the most unstable modes. Setting $s = 0$ in (3.3), we see that the range of unstable wavenumbers $(m - m_c)$ scales like $\delta^{1/2}$, assuming that $D_s/D_{mm} = O(1)$ at $(m_c, 0, \epsilon_c)$.

3.2. *Formulation*

In this weakly nonlinear regime, an $O(\delta^{1/2})$ range of wavenumbers around $m = m_c$ is unstable. Whilst the most complete description of the instability would involve following the evolution of such an envelope of unstable modes, here the evolution of only the most unstable mode, in isolation, is described. Thus, the evolution will be determined by the self-interaction of the most unstable mode with its harmonics. For a first study of the weakly nonlinear regime, this approach will prove sufficient. The more general problem treating an envelope of unstable modes will be discussed briefly in § 5.2.

We will use the nonlinear equations of motion, (2.8*a*)–(2.8*c*). To benefit from the assumption of weak nonlinearity, the variables must be properly scaled. Since the growth rate $s \sim \delta$, we must introduce a slow-time variable

$$\tau = \delta t.$$

Since this is a nonlinear problem, it is also important to correctly scale the perturbation quantities, which we expect to grow only to a small (non-dimensional) amplitude. It is possible to show that nonlinear saturation terms become as important as the linear growth terms when the perturbations are of order δ, and hence that the optimal scaling for the analysis will be $u' \sim \psi' \sim \sigma' \sim \delta$. Further, we expand these rescaled quantities in terms of δ, and therefore we write:

$$u' = \delta u_1 + \delta^2 u_2 + \cdots, \quad \psi' = \delta\psi_1 + \delta^2\psi_2 + \cdots, \quad \sigma' = \delta\sigma_1 + \delta^2\sigma_2 + \cdots. \tag{3.4}$$

With these changes, equations (2.8*a*)–(2.8*c*) become

$$\delta\frac{\partial u_1}{\partial\tau}+\delta\boldsymbol{u}_1\cdot\nabla u_1+\left(Q_c-\frac{\delta\epsilon_c}{\delta_0}\frac{\partial Q_0}{\partial\epsilon}\bigg|_{\epsilon=\epsilon_c}\right)\frac{\partial}{\partial z}(\psi_1+\delta\psi_2)$$
$$=\epsilon_c\left(1-\frac{\delta}{\delta_0}\right)\frac{\partial^2}{\partial z^2}(u_1+\delta u_2)+O(\delta^2),\quad(3.5a)$$

$$\delta\frac{\partial\xi_1}{\partial\tau}+\delta\boldsymbol{u}_1\cdot\nabla\xi_1-\left(f_c-\frac{\delta\epsilon_c}{\delta_0}\frac{\partial f}{\partial\epsilon}\bigg|_{\epsilon=\epsilon_c}\right)\frac{\partial}{\partial z}(u_1+\delta u_2)-4\epsilon_c\left(1-\frac{\delta}{\delta_0}\right)$$
$$\times\frac{\partial}{\partial y}(\sigma_1+\delta\sigma_2)=\epsilon_c\left(1-\frac{\delta}{\delta_0}\right)\frac{\partial^2}{\partial z^2}(\xi_1+\delta\xi_2)+O(\delta^2),\quad(3.5b)$$

$$\delta\frac{\partial\sigma_1}{\partial\tau}+\delta\boldsymbol{u}_1\cdot\nabla\sigma_1+\frac{\partial}{\partial y}(\psi_1+\delta\psi_2)=\epsilon_c\left(1-\frac{\delta}{\delta_0}\right)\frac{\partial^2}{\partial z^2}(\sigma_1+\delta\sigma_2)+O(\delta^2),\quad(3.5c)$$

where $f_c(y)=f(y,\epsilon=\epsilon_c)=y_c+2\epsilon_c^{1/2}y$, and $Q_c(y)=Q_0(y,\epsilon=\epsilon_c)$. The linear operator $\mathscr{L}(-\mathrm{i}\partial/\partial z,\partial/\partial t,\epsilon)$ becomes

$$\mathscr{L}\left(-\mathrm{i}\frac{\partial}{\partial z},\delta\frac{\partial}{\partial\tau},\epsilon_c\left(1-\frac{\delta}{\delta_0}\right)\right)=\mathscr{L}\left(-\mathrm{i}\frac{\partial}{\partial z},0,\epsilon_c\right)+\delta\mathscr{L}_s\left(-\mathrm{i}\frac{\partial}{\partial z},0,\epsilon_c\right)\frac{\partial}{\partial\tau}$$
$$-\frac{\delta\epsilon_c}{\delta_0}\mathscr{L}_\epsilon\left(-\mathrm{i}\frac{\partial}{\partial z},0,\epsilon_c\right)+O(\delta^2).$$

Note that the dependence of the operators on $\partial/\partial y$ and y has been suppressed, since in the analysis (m,s,ϵ) gives the only active dependency. Now retaining terms quadratic in the disturbance amplitude, the master equation for the disturbance streamfunction is

$$\mathscr{L}\psi_1=\delta\left(\mathscr{N}_1-\mathscr{L}_s\frac{\partial\psi_1}{\partial\tau}+\frac{\epsilon_c}{\delta_0}\mathscr{L}_\epsilon\psi_1-\mathscr{L}\psi_2\right)+O(\delta^2),\quad(3.6)$$

where all the operators are evaluated at $(-\mathrm{i}\partial/\partial z,0,\epsilon_c)$, and where

$$\mathscr{N}_1=\epsilon_c\frac{\partial^2}{\partial z^2}(\boldsymbol{u}_1\cdot\nabla\xi_1)-4\epsilon_c\frac{\partial}{\partial y}(\boldsymbol{u}_1\cdot\nabla\sigma_1)-f_c\frac{\partial}{\partial z}(\boldsymbol{u}_1\cdot\nabla u_1).\quad(3.7)$$

It is possible to examine the limit $\delta\to0$ by gathering together terms of like order in δ in each of these equations. The general procedure will be to do this first for (3.6), enabling ψ_i to be evaluated, and then to use (3.5*a*)–(3.5*c*) to evaluate u_i and σ_i. The initial conditions will be that at $\tau=0$ the spatial structure of the disturbance is given by that of the most unstable linear mode.

3.3. *Leading-order equations*

The leading-order terms of (3.6) are

$$\mathscr{L}(-\mathrm{i}\partial/\partial z,0,\epsilon_c,\partial/\partial y,y)\psi_1=0.$$

We look for a mode of the form $\psi_1(y,z,\tau)=\mathrm{Re}\{\Psi_1(y,\tau)\,\mathrm{e}^{\mathrm{i}mz}\}$, where $m=m_0+\delta m_1+O(\delta^2)$. For this form of solution $-\mathrm{i}\partial/\partial z=m$, so that

$$\mathscr{L}(-\mathrm{i}\partial/\partial z,0,\epsilon_c)=\mathscr{L}(m_0,0,\epsilon_c)+\delta m_1\mathscr{L}_m(m_0,0,\epsilon_c)+O(\delta^2).\quad(3.8)$$

Thus, the leading-order master equation becomes

$$\mathscr{L}(m_0, 0, \epsilon_c, \partial/\partial y, y)\Psi_1 = 0, \quad \Psi_1 \to 0 \text{ as } |y| \to \infty.$$

For there to exist bounded solutions to this equation some eigencondition must be satisfied. Here, the relevant one will, of course, be that $m_0 = m_c$ (the critical wavenumber). The function $\Psi_1(y, \tau)$ will have the latitudinal structure of the corresponding neutral eigenfunction, which we write simply as $\Psi_c(y)$. As described in § 2.4, we can choose $\Psi_c(y)$ to be real, and we normalize it so that

$$\int_{-\infty}^{\infty} \Psi_c^2 \mathrm{d}y = 1. \tag{3.9}$$

Thus, the leading-order streamfunction is

$$\psi_1(y, z, \tau) = \mathrm{Re}\{A(\tau)\Psi_c(y)\,\mathrm{e}^{\mathrm{i}mz}\},$$

where $m = m_c + O(\delta)$, and where $A(\tau)$ is a complex amplitude function. We cannot determine the behaviour of $A(\tau)$ at this level of approximation.

From (3.5*a*) and (3.5*c*) the leading-order zonal momentum and buoyancy equations are

$$Q_c \frac{\partial \psi_1}{\partial z} = \epsilon_c \frac{\partial^2 u_1}{\partial z^2}, \quad \frac{\partial \psi_1}{\partial y} = \epsilon_c \frac{\partial^2 \sigma_1}{\partial z^2}.$$

On integration with respect to z they yield the expressions

$$u_1 = \mathrm{Re}\left\{-\frac{\mathrm{i}}{m_c \epsilon_c} Q_c A \Psi_c \mathrm{e}^{\mathrm{i}mz}\right\} + \bar{u}_1(y, \tau), \quad \sigma_1 = \mathrm{Re}\left\{-\frac{1}{\epsilon_c m_c^2} A \frac{\mathrm{d}\Psi_c}{\mathrm{d}y} \mathrm{e}^{\mathrm{i}mz}\right\} + \bar{\sigma}_1(y, \tau),$$

where $\bar{u}_1(y, \tau)$ and $\bar{\sigma}_1(y, \tau)$ are functions of integration, to be determined. Substituting these expressions into the leading-order vorticity equation, obtainable from (3.5*b*), reveals that $\bar{\sigma}_1 = \bar{\sigma}_1(\tau)$, and hence that $\bar{\sigma}_1 = 0$ (by conservation of mass). However, we cannot determine $\bar{u}_1(y, \tau)$, that is the vertical mean part of the leading-order perturbation zonal velocity, at this level of approximation. Because of the diffusive nature of the leading-order balance, this function need not vanish, and indeed it will be crucial to the process of nonlinear equilibration. However, if we wish to recover the linear mode at $\tau = 0$, then we must have $\bar{u}_1(y, 0) = 0$.

3.4. *First-order equations*

Before evaluating the first-order streamfunction and associated fields, it is advisable to complete the leading-order solution. We need to evaluate $\bar{u}_1(y, \tau)$, and we can do this by considering the first-order zonal momentum balance, obtainable from (3.5*a*):

$$\frac{\partial u_1}{\partial \tau} + \boldsymbol{u}_1 \cdot \nabla u_1 + Q_c \frac{\partial \psi_2}{\partial z} - \frac{\epsilon_c}{\delta_0} \left.\frac{\partial Q_0}{\partial \epsilon}\right|_{\epsilon = \epsilon_c} \frac{\partial \psi_1}{\partial z} = \epsilon_c \frac{\partial^2 u_2}{\partial z^2} - \frac{\epsilon_c}{\delta_0} \frac{\partial^2 u_1}{\partial z^2}. \tag{3.10}$$

A little calculation shows that

$$\boldsymbol{u}_1 \cdot \nabla u_1 = \mathrm{Re}\left\{\frac{1}{2\epsilon_c} \frac{\mathrm{d}}{\mathrm{d}y}\left(Q_c \Psi_c^2\right) |A|^2 + \mathrm{i}m_c \overline{Q}_1 \Psi_c A \mathrm{e}^{\mathrm{i}mz} - \frac{1}{2\epsilon_c} \frac{\mathrm{d}Q_c}{\mathrm{d}y} \Psi_c^2 A^2 \mathrm{e}^{2\mathrm{i}mz}\right\}, \tag{3.11}$$

where $\overline{Q}_1 = -\partial \bar{u}_1 / \partial y$. Hence, taking the vertical average of (3.10), and integrating with respect to τ implies that

$$\bar{u}_1(y, \tau) = -\frac{1}{2\epsilon_c} \frac{\mathrm{d}}{\mathrm{d}y}(Q_c \Psi_c^2) \int_0^{\tau} |A(\tau')|^2 \mathrm{d}\tau', \tag{3.12}$$

using the boundary condition $\bar{u}_1(y, 0) = 0$. The leading-order solution is now complete.

We turn to calculating the first-order streamfunction. The $O(\delta)$ terms of the master equation (3.6) are

$$\mathscr{L}\left(-\mathrm{i}\frac{\partial}{\partial z},0,\epsilon_c,\frac{\partial}{\partial y},y\right)\psi_2=\mathscr{N}_1-\mathscr{L}_s\frac{\partial\psi_1}{\partial\tau}+\frac{\epsilon_c}{\delta_0}\mathscr{L}_\epsilon\psi_1-m_1\mathscr{L}_m\psi_1. \tag{3.13}$$

The final term comes from the $O(\delta)$ term of (3.8), and the operators on the right-hand side are now all evaluated at $(m_c,0,\epsilon_c)$, accurate to $O(\delta)$. To progress from (3.13) we clearly need to evaluate $\mathscr{N}_1$, defined in (3.7). Noting that

$$\boldsymbol{u}_1\cdot\nabla\xi_1=0,\quad \boldsymbol{u}_1\cdot\nabla\sigma_1=\mathrm{Re}\left\{\frac{\mathrm{i}}{2m_c\epsilon_c}\left(\Psi_c\frac{\mathrm{d}^2\Psi_c}{\mathrm{d}y^2}-\left(\frac{\mathrm{d}\Psi_c}{\mathrm{d}y}\right)^2\right)A^2\mathrm{e}^{2\mathrm{i}mz}\right\},$$

and using (3.11), we can do this, and hence show that a more explicit version of (3.13) is

$$\begin{aligned}\mathscr{L}\left(-\mathrm{i}\frac{\partial}{\partial z},0,\epsilon_c\right)\psi_2=&\left(-\frac{\mathrm{d}A}{\mathrm{d}\tau}\mathscr{L}_s+\frac{\epsilon_c}{\delta_0}A\mathscr{L}_\epsilon-m_1A\mathscr{L}_m+m_c^2f_c\overline{Q}_1A\right)\Psi_c\mathrm{e}^{\mathrm{i}mz}\\&-\frac{2\mathrm{i}}{m_c}\left(\frac{\mathrm{d}}{\mathrm{d}y}\left(\Psi_c\frac{\mathrm{d}^2\Psi_c}{\mathrm{d}y^2}-\left(\frac{\mathrm{d}\Psi_c}{\mathrm{d}y}\right)^2\right)-\frac{m_c^2}{2\epsilon_c}f_c\frac{\mathrm{d}Q_c}{\mathrm{d}y}\Psi_c^2\right)A^2\mathrm{e}^{2\mathrm{i}mz},\end{aligned} \tag{3.14}$$

where all the operators on the right-hand side are now evaluated at $(m_c,0,\epsilon_c)$.

3.5. *The solvability condition*

We will not attempt to solve (3.14). Enough useful information can be gained from the solvability condition associated with the component of the particular integral proportional to $\mathrm{e}^{\mathrm{i}mz}$. If we write this component of ψ_2 as $\mathrm{Re}\{\Psi_2(y,\tau)\mathrm{e}^{\mathrm{i}mz}\}$, then the leading-order $\mathrm{e}^{\mathrm{i}mz}$ component of (3.14) is

$$\mathscr{L}\Psi_2=-\frac{\mathrm{d}A}{\mathrm{d}\tau}\mathscr{L}_s\Psi_c+\frac{\epsilon_c}{\delta_0}A\mathscr{L}_\epsilon\Psi_c-m_1A\mathscr{L}_m\Psi_c+m_c^2f_c\overline{Q}_1A\Psi_c,$$

where all the operators are evaluated at $(m_c,0,\epsilon_c)$. Since $\mathscr{L}$ is self-adjoint, we have $\int_{-\infty}^{\infty}\Psi_c\mathscr{L}\Psi_2\mathrm{d}y=\int_{-\infty}^{\infty}\Psi_2\mathscr{L}\Psi_c\mathrm{d}y=0$, so necessarily

$$\begin{aligned}0=&-\left(\int_{-\infty}^{\infty}\Psi_c\mathscr{L}_s\Psi_c\mathrm{d}y\right)\frac{\mathrm{d}A}{\mathrm{d}\tau}+\frac{\epsilon_c}{\delta_0}\left(\int_{-\infty}^{\infty}\Psi_c\mathscr{L}_\epsilon\Psi_c\mathrm{d}y\right)A\\&-m_1\left(\int_{-\infty}^{\infty}\Psi_c\mathscr{L}_m\Psi_c\mathrm{d}y\right)A+m_c^2\left(\int_{-\infty}^{\infty}f_c\overline{Q}_1\Psi_c^2\mathrm{d}y\right)A.\end{aligned}$$

The integrals involving the operator $\mathscr{L}$ can be related to differentials of the dispersion relation $D(m,s,\epsilon)$, using expressions of the form (2.15) and the normalization condition (3.9). Then recalling, from (3.1), that $D_m(m_c,0,\epsilon_c)=0$, and substituting for δ_0 from (3.2), we find that equivalently we must have

$$\frac{\mathrm{d}A}{\mathrm{d}\tau}=A-k^2A\int_0^\tau|A(\tau')|^2\mathrm{d}\tau', \tag{3.15}$$

where, using (3.12) for $\bar{u}_1$ (and hence $\overline{Q}_1$), and (2.16) for $D_s(m_c,0,\epsilon_c)$,

$$k^2=\frac{1}{4\epsilon_c^2m_c^2}\int_{-\infty}^{\infty}f_c\Psi_c^2\frac{\mathrm{d}^2}{\mathrm{d}y^2}(Q_c\Psi_c^2)\,\mathrm{d}y. \tag{3.16}$$

Since $\Psi_c(y)$ is real, k^2 is real. It will be assumed that $k^2 > 0$, although it has not been proved that this must necessarily be the case for an arbitrary shear flow. If $k^2 > 0$ then, as will be shown below, the amplitude equation (3.15) implies that $A(\tau)$ will only grow to a magnitude of $O(1)$. This corresponds to a non-dimensional disturbance amplitude of $O(\delta)$, and hence no more terms need be considered in the amplitude equation – a consistent approach.

Since m_1 does not appear in (3.15) and (3.16), the amplitude equation describes the evolution of any mode with $|m - m_c| = O(\delta)$. It is natural to consider the evolution of the most unstable mode, but we are not restricted to this case alone.

3.6. *Equilibration*

The weakly nonlinear analysis is now complete, since we have a complete spatial and temporal description of the leading-order solution as $\delta \to 0$. The evolution equation (3.15) is to be solved subject to the initial condition $A = A_0$ at $\tau = 0$. Since k^2 is real, even though A may be complex, it will have constant argument. Thus, the problem reduces to solving for $|A|$, and we can easily show that

$$|A| = \frac{(1 + k^2|A_0|^2)^{1/2}}{k \cosh((1 + k^2|A_0|^2)^{1/2}(\tau - \tau_*))}, \quad \tau \geqslant 0, \tag{3.17}$$

where τ_* is a constant given by

$$\tau_* = \frac{1}{2(1 + k^2|A_0|^2)^{1/2}} \log\left(\frac{(1 + k^2|A_0|^2)^{1/2} + 1}{(1 + k^2|A_0|^2)^{1/2} - 1}\right). \tag{3.18}$$

The behaviour is quite simple. $A(\tau)$ grows to a maximum amplitude at $\tau = \tau_*$ and then decays towards zero as $\tau \to \infty$. The maximum amplitude $|A|_{max}$ depends on the initial amplitude $|A_0|$, and it is easy to see that

$$|A|_{max} = \frac{1}{k}(1 + k^2|A_0|^2)^{1/2}.$$

$1/k$ is an amplitude that naturally arises here – it is the value of $|A|_{max}$ that occurs in the problem with $|A_0| \to 0$. However, if $|A_0|$ is of the same order as $1/k$, then the instability grows to a larger amplitude. Physically, this scenario could occur if the instability were triggered by a large-amplitude wave with the correct sort of vertical and latitudinal structure, perhaps an equatorial Kelvin wave. In many weakly nonlinear systems, such as those governed by the Landau equation, the maximum disturbance amplitude does not depend so crucially on the initial conditions.

Quantities proportional to A, such as the streamfunction or the fluctuating buoyancy acceleration, tend to zero as $\tau \to \infty$. However, the magnitude of the mean flow change is

$$\int_0^\tau |A(\tau')|^2 \mathrm{d}\tau' = \frac{1}{k^2}\{1 + (1 + k^2|A_0|^2)^{1/2} \tanh((1 + k^2|A_0|^2)^{1/2}(\tau - \tau_*))\}.$$

Thus, the mean flow change increases monotonically for all τ, and tends to a constant non-zero amplitude, also dependent on $|A_0|$, as $\tau \to \infty$. In that sense the system has a long memory.

3.7. *Linear stability of the evolving mean flow*

Inertial instability is driven by a particular property of the mean flow, the particular property being related to regions where $f\overline{Q} < 0$, where the overbar denotes a vertical mean. During the course of the inertial instability, the mean flow is changed, and it

is useful to examine the significance of this change. In particular, does the evolved mean flow correspond to an inertially stable state?

To proceed, we note that during the instability the evolving potential vorticity is

$$Q(y,z,t) = Q_0(y) + \delta\left(Q_0(y)\frac{\partial\sigma_1}{\partial z} - \frac{\partial u_1}{\partial y}\right) + O(\delta^2),$$

using (2.7) and (3.4). Thus, $\overline{Q} = Q_0(y) + \delta\overline{Q}_1(y,\tau) + O(\delta^2)$, where $\overline{Q}_1 = -\partial\bar{u}_1/\partial y$, so that to leading order, changes in $\overline{Q}$ are solely due to changes in $\bar{u}$. Therefore, we analyse the action of the instability by considering the linear stability properties of the evolving mean flow $\bar{u} = u_0(y) + \delta\bar{u}_1(y,\tau)$. To do this, we treat the evolving mean flow as effectively frozen in time, and consider the growth of linear disturbances on this frozen state. From (2.10) and (2.11), the governing linear equation for the disturbance streamfunction $\phi(y,z,t)$ is

$$\left\{\mathscr{L}\left(-\mathrm{i}\frac{\partial}{\partial z}, \frac{\partial}{\partial t}, \epsilon_c\left(1-\frac{\delta}{\delta_0}\right), \frac{\partial}{\partial y}, y\right) + \delta f\overline{Q}_1\frac{\partial^2}{\partial z^2}\right\}\phi = 0. \tag{3.19}$$

Since the mean flow changes are only of $O(\delta)$, we can analyse the stability properties by a perturbation analysis. We write $\phi(y,z,t) = \mathrm{Re}\{\Phi(y)\mathrm{e}^{\mathrm{i}mz+st}\}$, and

$$\Phi(y) = \Phi_0(y) + \delta\Phi_1(y) + \cdots, \quad m = m_0 + \delta m_1 + \cdots, \quad s = s_0 + \delta s_1 + \cdots.$$

Then the leading-order terms of (3.19) are $\mathscr{L}(m_0, s_0, \epsilon_c, \partial/\partial y, y)\Phi_0(y) = 0$. We wish to look for the most unstable linear mode, and by our previous assumptions, at $\epsilon = \epsilon_c$ the maximum value of s_0 for latitudinally bounded solutions is $s_0 = 0$, when $m = m_c$. Therefore, we set $s_0 = 0$ and $m = m_c$, and the leading-order streamfunction $\Phi_0(y)$ is just the neutral streamfunction $\Psi_c(y)$.

The $O(\delta)$ terms of (3.19) are

$$\mathscr{L}\Phi_1 + m_1\mathscr{L}_m\Psi_c + s_1\mathscr{L}_s\Psi_c - (\epsilon_c/\delta_0)\mathscr{L}_\epsilon\Psi_c - m_c^2 f_c\overline{Q}_1\Psi_c = 0,$$

where the operators are all evaluated at $(m_c, 0, \epsilon_c)$. The solvability condition for $\Phi_1(y)$, formed by multiplying this equation by $\Psi_c(y)$ and integrating across the domain, is

$$s_1 = \left.\frac{\epsilon_c D_\epsilon}{\delta_0 D_s}\right|_{(m_c,0,\epsilon_c)} + \frac{m_c^2}{D_s(m_c,0,\epsilon_c)}\int_{-\infty}^{\infty} f_c\overline{Q}_1\Psi_c^2\,\mathrm{d}y \;\Leftrightarrow\; s_1 = 1 - k^2\int_0^\tau |A(\tau')|^2\mathrm{d}\tau',$$

using (2.16), (3.2), (3.12) and (3.16). Comparing with (3.15), the evolution equation for A, we see that

$$s_1 = \frac{\mathrm{d}A/\mathrm{d}\tau}{A} \;\Rightarrow\; s = \delta\frac{\mathrm{d}A/\mathrm{d}\tau}{A} + O(\delta^2),$$

since $s_0 = 0$. Thus, to first order in δ, the growth rate of linear disturbances on the evolving mean flow behaves in a simple fashion. Initially, $s = \delta$, the original growth rate, whatever the value of A_0. The instability has had no effect on the mean flow. At the maximum (streamfunction) amplitude, we see that $s = 0$: the mean flow is exactly neutralized to linear disturbances (to leading order in δ). As $\tau \to \infty$, $s \to -\delta(1 + k^2|A_0|^2)^{1/2}$. Hence, the final stability properties are dependent upon the initial conditions, and the larger $|A_0|$, the more stable is the final state.

Note that the above stability analysis is not exact for the complete evolving inertial instability, since we have ignored the presence of the sinusoidal perturbations of u', ψ' and σ', and of the time-dependence of the mean flow. However, as we approach the final steady state, both these effects disappear, and the evolving stability analysis becomes exact.

3.8. *Discussion*

The analysis has given us a clear and complete description of the flow. The disturbance streamfunction grows until the mean flow is neutralized. At this instant, there is a small disturbance on a neutral mean flow, and naturally, the disturbance neither grows nor decays. However, the continuing overturning motion induces a further mean flow change, so that the flow becomes slightly stabilized. The disturbance streamfunction will then start to decay exponentially. Yet the mean flow change satisfies $\partial\bar{u}/\partial\tau \propto \delta|A|^2$, and is insensitive to whether the mode is growing or decaying. Hence, throughout the decay there is a further stabilizing mean flow change. This mean flow change persists, and the final mean flow is stable.

Even though in these weakly unstable conditions the disturbances are almost in a diffusive balance, the mean flow is z-independent and the mean flow change is not directly related to the vertical diffusion. Recalling that there is no horizontal diffusion in the system, the mean flow change is due solely to the nonlinear momentum advection term $-\partial\overline{(u'v')}/\partial y$. This suggests that if we perform numerical simulations of the nonlinear instability, we should be careful to ensure that the mean flow changes are indeed independent of any artifical horizontal diffusion introduced for numerical stability alone.

4. Numerical simulations

There are good reasons for performing some numerical simulations of the nonlinear evolution of the instability. The first is to assess the accuracy of the weakly nonlinear analysis. Were the equations derived and solved without error? A second reason is to assess the applicability of the analysis. Over what range of δ do the predicted results hold true? A third reason is to extend the results into a more unstable regime. Is the same qualitative behaviour seen for larger δ? Finally, numerical simulations may help us to see some features of the neutralization process that we might otherwise have missed.

4.1. *The numerical model*

The inertial instability to be simulated is centred near the equator, and dies away exponentially in latitude. There are two important consequences. First, the horizontal domain may be truncated to a finite width. Providing the domain is wide enough, the boundary effects will be negligible. Secondly, it is possible to use a Fourier basis for the latitudinal structure, even though there will be a discontinuity in the fields between the northern and southern boundaries of the domain. As explained by Boyd (2000), since the fields are exponentially small at the boundaries, the discontinuity will be exponentially small, and hence the error in the Fourier approximation will be exponentially small too. The vertical domain is periodic, so it is natural to use a Fourier basis to represent the vertical structure. The nonlinear terms are evaluated pseudospectrally.

A fourth-order Runge–Kutta method is used to time-step the advective, Coriolis, and buoyancy terms. This totally explicit treatment of the buoyancy terms leads to a stability restriction on the dimensional time step Δt of the form $N\Delta t \lesssim 1$, although the high order of the scheme means that this limit is not too restrictive in practice. The vertical diffusion is applied in an arbitrarily stable infinite-order step. For a Fourier component with dimensional vertical wavenumber m, at each step this takes the form $\phi \to \exp(-\nu m^2 \Delta t)\phi$. There is no explicit horizontal diffusion.

4.2. *The experimental set-up*

The dimensional basic state $u_0 = \Lambda y$ is used for all the simulations. With the non-dimensionalization of § 2.3, taking $\hat{y}_c = 1$, this corresponds to

$$u_0 = 2y + \epsilon^{-1/2}, \qquad Q_0 = 2\epsilon^{1/2}y - 1, \Rightarrow (-fQ_0) = 1 - 4\epsilon y^2 \text{ since } f = 1 + 2\epsilon^{1/2}y.$$

The initially unstable region is $|y| < \frac{1}{2}\epsilon^{-1/2}$. From § 2.2, the critical linear conditions are $\epsilon_c = 5^{-5/6}$ and $m_c = 5^{1/6}$, and the corresponding normalized eigenfunction is

$$\Psi_c(y) = \left(\frac{m_c}{\pi}\right)^{1/4} \exp(-\tfrac{1}{2}m_c y^2).$$

The dispersion relation for this mode, defined via (2.14), is $D = m^2 - 4\epsilon m - m^2(s+\epsilon m^2)^2$. Hence, we may explicitly evaluate the coefficients in the amplitude equation. We find, from (3.16), that $k^2 = 5^{7/12} \times (8\pi)^{-1/2} \approx 0.510$, and, from (3.2), that the normalizing factor $\delta_0 = 3 \times 5^{-1/2}$. Thus, we are in a position to quantitatively test the weakly nonlinear theory. The approach will be first to demonstrate close agreement between the theory and the numerical simulations for $\delta \ll 1$, and then to test the applicability of the theory for $\delta \leqslant 0.3$. This will extend the results to the point at which the instability leads to density contour overturning.

The simulations were performed in the domain $|y| < 6$, so that the centre of the domain is coincident with the centre of the unstable region, and the domain width is approximately six times that of the unstable region. The domain height is taken to be $2\pi/m_*$, where m_* is the vertical wavenumber of the most unstable linear mode. Each simulation is initialized with an equatorially confined disturbance of vertical wavenumber m_*, and run for some time with the nonlinear terms switched off, to extract the fastest growing linear mode. The linear growth rate is typically in agreement with the exact value to at least five significant figures. Once this mode has been extracted, it is rescaled to an appropriate start-up amplitude, and the nonlinear terms are switched on.

4.3. *The weakly unstable regime*

4.3.1. *Verification of the time evolution*

Simulations were performed with $|A_0| = 10^{-4}$ (i.e. in the regime $k|A_0| \ll 1$) for $\delta = 0.005, 0.01, 0.02, 0.035, 0.05, 0.07$ and 0.1. The resolution was 128 latitudinal grid points $\times$ 32 vertical grid points, with a non-dimensional time step $\Delta t = 0.01$. To assess the accuracy of the amplitude equations in this range, it is useful to consider the function

$$A_p(t) = \frac{m_*}{2\pi} \int_0^{2\pi/m_*} 2\exp(-\mathrm{i}m_* z) \left(\int_{-\infty}^{\infty} \Psi_c(y)\psi(y,z,t)\,\mathrm{d}y \right) \mathrm{d}z,$$

where $\psi(y,z,t)$ is the numerical streamfunction. $A_p(t)$ is the projection of the numerical streamfunction onto the leading-order analytical streamfunction. According to the weakly nonlinear theory, $A_p = \delta A + O(\delta^2)$.

For $\delta = 0.005$, A_p is given by the theoretical prediction δA to within 2.5% throughout the whole time evolution. To test the applicability of the theory for larger δ, the maximum projected streamfunction amplitude $|A_p|_{max}$ is compared with the maximum predicted amplitude $\delta|A|_{max} = \delta/k$. The results, shown in figure 2, indicate that this relationship does hold good, although the $O(\delta^2)$ correction appears to be quite large. Also shown in figure 2 is the time evolution of $|A_p|$ at $\delta = 0.1$, along with the theoretical prediction. The quantitative deviation between experiment and theory is becoming quite significant.

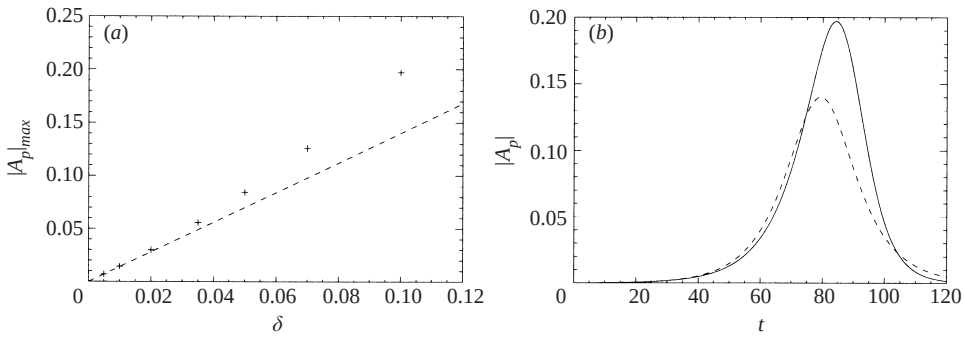

FIGURE 2. (*a*) The maximum projected numerical amplitude $|A_p|_{max}$ (+, data points) and the predicted maximum amplitude $\delta|A|_{max} = \delta/k$ (– – –) against δ, for the simulations with $k|A_0| \ll 1$. The difference is proportional to δ^2. (*b*) The time evolution of the projected amplitude $|A_p|$ for the simulation at $\delta = 0.1$: ——, numerical result; – – –, theoretical prediction.

4.3.2. *Verification of the latitudinal structure*

For $\delta = 0.005$, the latitudinal structure of the numerically calculated fields is in good agreement with the theoretical predictions. To assess the difference for larger δ, we consider the quantity

$$U_p = \lim_{t\to\infty} \left(\int_{-\infty}^{\infty} (\bar{u}(y,t) - u_0(y,\epsilon) - \delta\bar{u}_1(y,\tau))^2 \mathrm{d}y \right)^{1/4},$$

where $\bar{u}(y,t)$ is the numerically calculated mean zonal velocity. Here, $\delta\bar{u}_1(y,\tau)$ is the predicted mean flow change, obtainable from (3.12) as

$$\delta\bar{u}_1(y,\tau) = -\delta \left(\frac{5}{\pi}\right)^{1/2} (1 + 5^{7/12}y - 2\times 5^{1/6}y^2)\exp(-m_c y^2) \int_0^{\tau} |A(\tau')|^2 \mathrm{d}\tau'.$$

According to the weakly nonlinear theory, $U_p = O(\delta)$ as $\delta \to 0$. Figure 3(*a*) shows U_p versus δ, verifying that the theory does correctly predict both the magnitude and the latitudinal structure of the mean flow change as $\delta \to 0$.

By $\delta = 0.1$ the theory is at the limit of its quantitative usefulness. This is not only evident in the magnitude of the perturbation fields, but also in their latitudinal structure. Figure 3(*b*) shows the final mean flow change and the theoretical prediction for $\delta = 0.1$. Although the theoretical prediction has the correct qualitative behaviour, there are now significant deviations between it and the numerical result, especially for $y > 0$.

It is worth examining some of the consequences of these mean flow changes. The plot of the mean flow change shows that the instability induces a westward jet north of $y = 0$ (the centre of the unstable region), and an eastward jet south of $y = 0$. Thus, the disturbance has a mean negative shear around $y = 0$. Recall that the basic state has a positive shear across the whole domain, thus producing a region of anomalous vorticity around $y = 0$. For the evolved state the negative shear of the mean flow change around $y = 0$ has reduced the magnitude of the anomalous vorticity there. This is the inertial neutralization that has been described by the evolving linear stability analysis of § 3.7. Further aspects of these neutralized mean flows will be discussed in § 4.6.

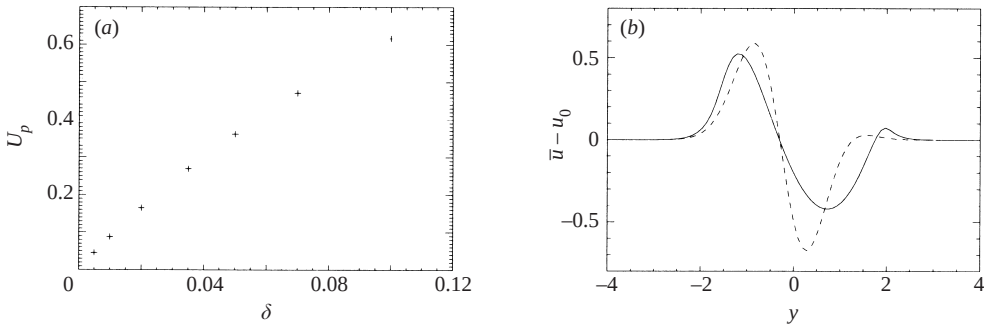

FIGURE 3. (*a*) The quantity U_p versus δ, for the simulations with $k|A_0| \ll 1$. (*b*) The evolved mean flow change for the simulation at $\delta = 0.1$: ——, numerical result; – – –, theoretical prediction.

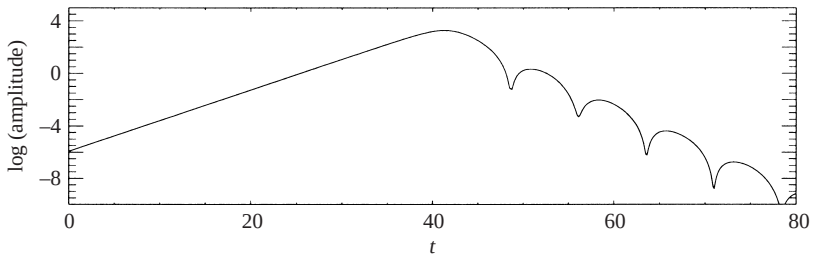

FIGURE 4. Disturbance amplitude (that is the globally integrated absolute streamfunction) versus time, for the simulation at $\delta = 0.25$.

4.4. *A moderately unstable regime*

For $\delta > 0.1$, the instability can no longer be described quantitatively by the weakly nonlinear amplitude equations. However, in the range $0.1 < \delta \leqslant 0.3$, the instability is not strong enough to lead to density contour overturning, and this regime is accessible to simple numerical simulations. Results will be given for simulations performed at $\delta = 0.15$ and 0.20 (at a resolution of 256 latitudinal grid points $\times$ 32 vertical grid points), and for $\delta = 0.25$ and 0.30. Owing to the formation of very large spatial gradients at $\delta = 0.25$ and 0.30, it was necessary to use 512 latitudinal grid points for these simulations. Furthermore, using a domain mapping, the local grid point density was increased between $y = 2$ and 3.

The simulations for $0.1 < \delta < 0.3$ display the same qualitative behaviour, and most aspects may be discussed together. The temporal behaviour is similar to the weakly unstable case. The disturbance streamfunction grows exponentially, reaches a maximum amplitude, and then decays exponentially to zero. The final state is one of uniform stratification, with no meridional flow, but with an adjusted mean zonal flow. In figure 4, the time evolution of the disturbance streamfunction is plotted for $\delta = 0.25$. The exponential decay is modified by an oscillatory behaviour. Whilst the streamfunction maintains approximately the same latitudinal structure during the decay, it periodically shrinks to a small amplitude and reappears with reversed sign. This oscillatory behaviour could presumably be described analytically if the weakly nonlinear analysis were carried out to the next order in δ.

The potential vorticity (PV) and density, at maximum disturbance amplitude, are shown in figure 5, for the $\delta = 0.25$ run. The fields are still very much dominated by structures with vertical wavenumbers zero and m_*, and even though the density

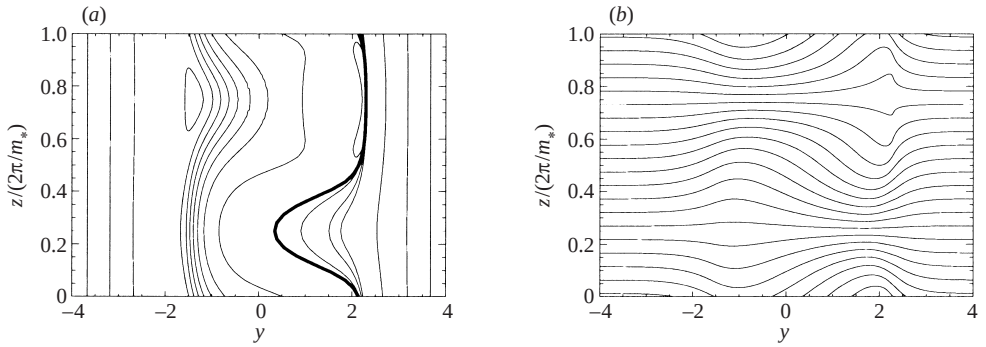

FIGURE 5. The instability at maximum disturbance amplitude, for $\delta = 0.25$. (*a*) Potential vorticity. (*b*) Density. Only the centre of the computational domain is shown. The equator is at $y = -1.08$, and the northern edge of the initially unstable region is at $y = 1.08$. The zero PV contour is indicated by a thicker line.

contours are considerably disturbed, they have not yet overturned. The vertical shear of the horizontal winds, mainly due to $\partial u/\partial z$, but also partially due to $\partial v/\partial z$, leads to regions where the Richardson number is small. The smallest values occur at, or near, maximum disturbance amplitude, in regions where the stratification is weak. For $\delta = 0.15, 0.20, 0.25$ and 0.30, the minimum values are $Ri = 1.5, 0.88, 0.57$ and 0.23, respectively.

Near maximum disturbance amplitude, the maximum density perturbation occurs north of the initially unstable region, at $y \approx 2$ (that is $y \approx 1.4\Lambda/\beta$ in dimensional variables). Figure 6(*b*) shows the corresponding streamfunction and density perturbation, for $\delta = 0.25$. For comparison, the streamfunction and density perturbation are also shown during the linear growth and decay stages of the instability. During the linear growth, the streamfunction is symmetric about $y = 0$, the most unstable latitude, and the maximum density perturbations occur within the unstable region – a consequence of the linear instability theory, as noted in § 2.2. At maximum disturbance amplitude, the centre of the overturning motion has drifted northwards. So too has the northern density extremum, although the southern one has stayed at roughly the same latitude. During the linear decay, the streamfunction drifts back towards $y = 0$, but the density perturbation remains asymmetrical. This behaviour can be interpreted in terms of the way in which the flow is neutralized.

4.5. *Inertial neutralization*

Throughout the instability, we expect motion to be strong in regions of anomalous vorticity. Initially, $(-f\overline{Q})$ has a single maximum at $y = 0$, and the instability responds by centring the overturning motion about this latitude. In the linear regime, a perturbation velocity $v'(y,z,t) = V(y)\cos(mz)\mathrm{e}^{st}$ is accompanied by a perturbation PV $Q' = -v'(s+\epsilon m^2)^{-1}\partial Q_0/\partial y$. Thus, for the uniform shear flow, the mean PV changes according to

$$\frac{\partial \overline{Q}}{\partial t} = -\frac{\partial}{\partial y}(\overline{v'Q'}) = \frac{1}{(s+\epsilon m^2)}\frac{\partial}{\partial y}\left(\overline{v'^2}\frac{\partial Q_0}{\partial y}\right) = \frac{\epsilon^{1/2}\mathrm{e}^{2st}}{s+\epsilon m^2}\frac{\mathrm{d}V^2}{\mathrm{d}y},$$

since $Q_0 = 2\epsilon^{1/2}y - 1$. For the most unstable mode $V^2 \propto \exp(-my^2)$, so

$$\partial\overline{Q}/\partial t > 0 \quad \text{if} \quad y < 0, \qquad \partial\overline{Q}/\partial t < 0 \quad \text{if} \quad y > 0.$$

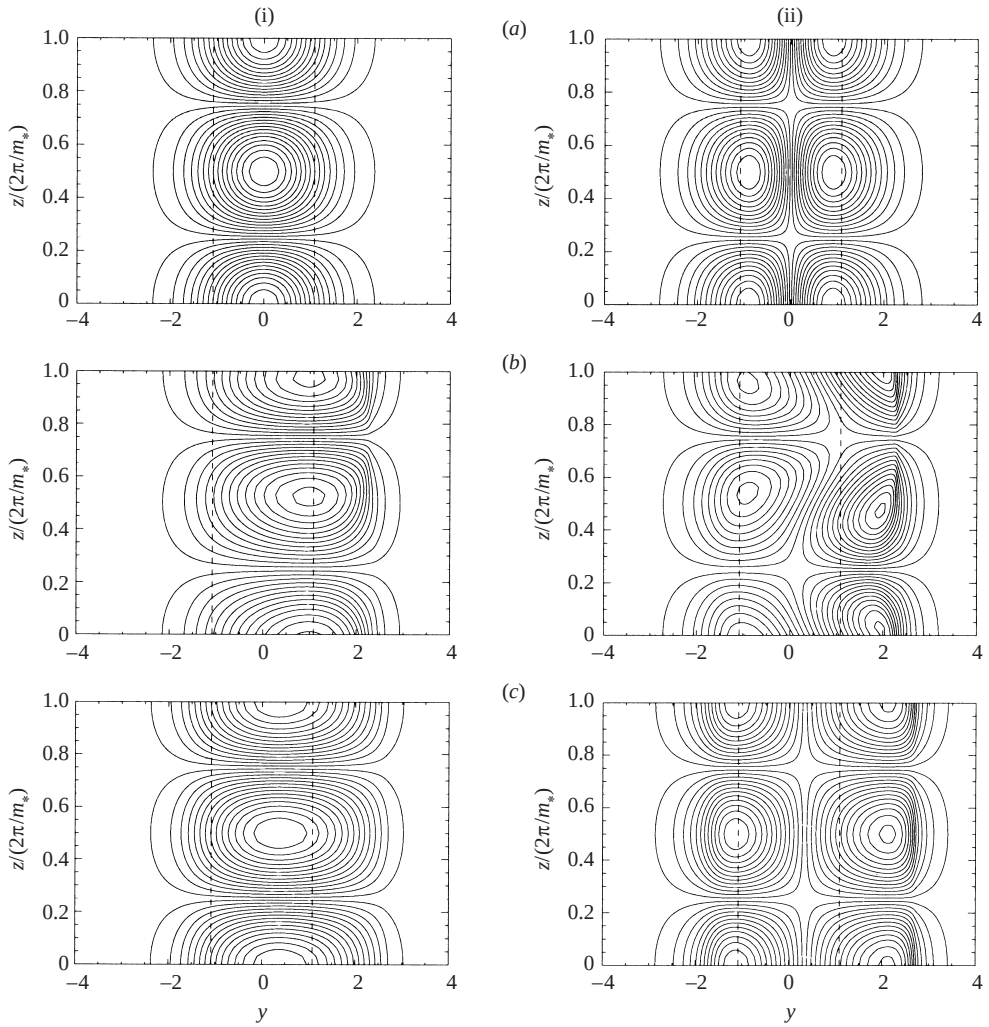

FIGURE 6. The instability at three different times during the $\delta = 0.25$ run. (i) Streamfunction. (ii) Density perturbation. (*a*) During the initial linear growth. (*b*) At maximum disturbance amplitude. (*c*) During the linear decay. The vertical dashed lines denote the boundaries of the initially unstable region.

The mean PV profile pivots about the most unstable latitude $y = 0$. The situation is illustrated in figure 7.

The new inertial stability is determined by $f\overline{Q}$. The pivoting means that for $y < 0$ the flow becomes locally more stable, whilst for $y > 0$ the flow becomes locally more unstable. The locally most unstable point shifts northwards. Hence, the natural response of the inertial instability is to shift northwards as well. We might anticipate that this sort of process will continue throughout the instability. As the cell moves further north, it will be mixing in increasingly less negative values of PV, so the neutralization will become more effective. Thus, the northward drift of the overturning motion, and the corresponding northward drift of the poleward density perturbation, are intimately connected with the neutralization process.

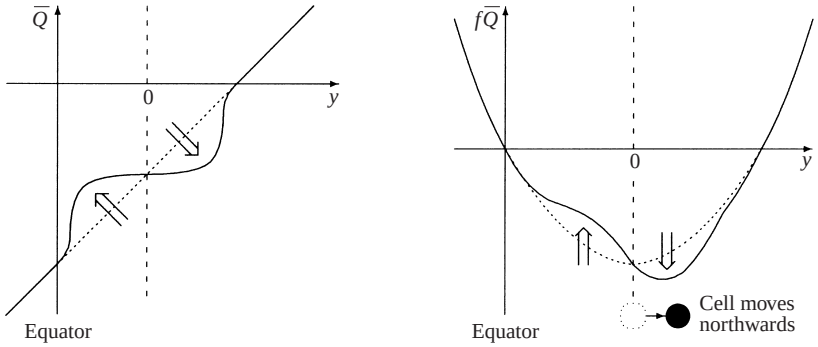

FIGURE 7. A schematic view of the neutralization process: ···, initial state; ——, initial mean flow change.

4.6. *The mean flow change*

What does an adjusted mean flow look like in the moderately nonlinear regime? The adjusted mean PV $\overline{Q}$ is shown in figure 8, for $\delta = 0.15$, 0.20 and 0.25. The profiles are taken at maximum disturbance amplitude (interpreted as neutralized profiles, since the disturbances are neither growing nor decaying), and at the end of the instability (interpreted as stabilized profiles).

The significance of these particular adjusted profiles is made apparent by considering the quantity $f\overline{Q}$. Recall that $f\overline{Q}$ determines the strength of the instability. As also shown in figure 8, in each of the stabilized profiles, $f\overline{Q}$ has been approximately homogenized to a small negative value over a broad region around the initially unstable region. This is the action of the inertial instability in the moderately unstable regime. It is clear that as δ increases, i.e. as the stabilizing forces decrease, the value of $(-f\overline{Q})$ in this homogenized region decreases, but does not reach zero. This is consistent with the idea that vertical diffusion can stabilize regions of anomalous vorticity. The inertial instability need not completely wipe out regions of anomalous vorticity, just reduce the value of $(-f\overline{Q})$. The higher the vertical diffusivity is, and hence the lower δ is, the greater the value of $(-f\overline{Q})$ that can be stabilized. Equivalently, the instability leads to a homogenization of the angular momentum gradient, rather than of the angular momentum itself.

Note that since the domain integrated PV is conserved, to balance the increase in the mean PV over the initially unstable region we must expect a decrease in the mean PV elsewhere. One such region is in the southern hemisphere, so it actually leads to an increase in $f\overline{Q}$ and enhanced inertial stability. The other such region, near $y = 1.5$, is in the northern hemisphere, and so it leads to a decrease in $f\overline{Q}$ and reduced inertial stability. However, in this region of reduced stability, $f\overline{Q}$ never falls to a lower value than elsewhere in the domain. Accompanying these two adjustment regions are a deepening minimum in the PV just south of the equator, and a corresponding maximum north of the initially unstable region. In three-dimensional reality these regions of reversed PV gradients might be wiped out by barotropic instability.

A noteworthy feature of the adjusted profiles, whether at the middle or end of the evolution, is that the line of zero mean PV actually moves polewards during the instability. This is consistent with the idea of a weakly diffusive stabilization, but contrary to the traditional view that inertial stability requires the line of zero PV to sit at the equator. It is also worth noting that as δ increases, and ϵ decreases, the most unstable mode becomes narrower. Despite this, the modes induce a mean

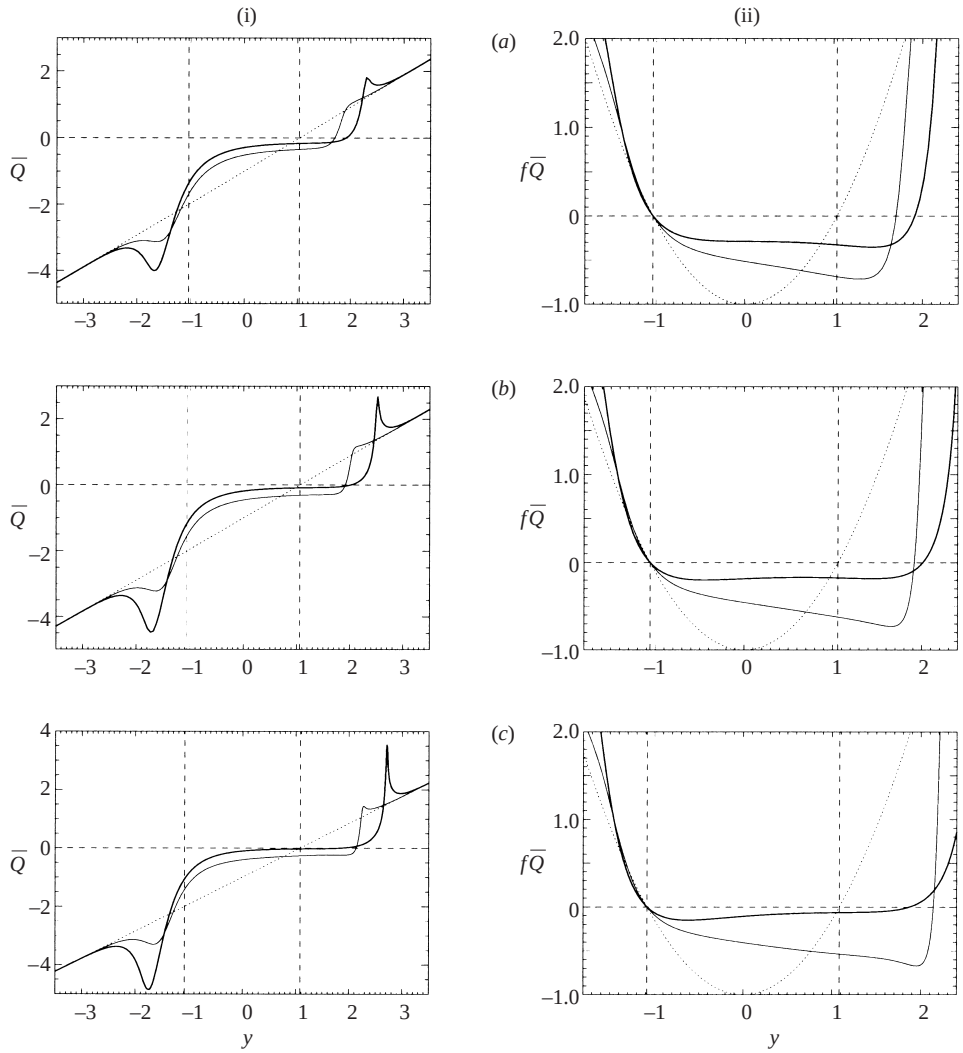

FIGURE 8. Mean flow profiles for (*a*) $\delta = 0.15$, (*b*) $\delta = 0.20$ and (*c*) $\delta = 0.25$. (i) Mean potential vorticity $\overline{Q}$, (ii) $f\overline{Q}$. In each plot the dotted curve is the initial profile, the thin solid curve is taken at maximum disturbance amplitude, and the thick solid curve is taken at the end of the instability. The vertical dashed lines denote the boundaries of the initially unstable region.

flow change over a wider and wider area. The idea that a latitudinally narrow mode can eventually come to completely neutralize a latitudinally broad region has been investigated in an unstratified f-plane inertial instability by Shen & Evans (1998).

5. Additional effects

5.1. *The inclusion of weak damping*

The problem studied in §§ 2–4 is the simplest self-consistent example of nonlinear inertial equilibration. However, it proves informative to allow for some sort of damping within the system, since such effects, no matter how weak, can radically alter the nonlinear evolution.

Following Hua *et al.* (1997), it is mathematically convenient to include such effects by adding terms that damp all the perturbation quantities at the same rate. Thus, we include the effects of weak damping in our analysis by choosing this rate to be $\delta\alpha$. Then, the linear growth rate changes according to $s \to s - \delta\alpha$, and hence when $\epsilon = \epsilon_c(1 - \delta/\delta_0)$ the leading-order growth rate of the most unstable mode will be $\delta(1 - \alpha)$. To study growing modes, we will require $0 < \alpha < 1$. We will formally set $\alpha = O(1)$, although this does not mean that we cannot subsequently choose it to be smaller than this.

The weakly nonlinear analysis of § 3 is modified by changing $\partial/\partial\tau$ to $\partial/\partial\tau + \alpha$. The latitudinal structure of all the perturbation quantities remains the same, but the time evolution is altered. In particular, writing

$$\bar{u}_1(y, \tau) = -\frac{1}{2\epsilon_c}\frac{\mathrm{d}}{\mathrm{d}y}(Q_c\Psi^2)B(\tau), \tag{5.1}$$

the mean flow equation, (3.12), becomes

$$\left(\frac{\mathrm{d}}{\mathrm{d}\tau} + \alpha\right) B = |A|^2, \quad B(0) = 0, \tag{5.2a}$$

using the boundary condition $\bar{u}_1(y, 0) = 0$. The evolution equation for $A(\tau)$ becomes

$$\left(\frac{\mathrm{d}}{\mathrm{d}\tau} + \alpha\right) A = A - k^2 AB, \quad A(0) = A_0, \tag{5.2b}$$

with k^2 given by (3.16). The evolution is now governed by the coupled equations (5.2*a*) and (5.2*b*). The undamped system may be recovered simply by setting $\alpha = 0$.

$B(\tau)$, which gives the magnitude of the mean flow change, and k^2 are real, so that even though $A(\tau)$ may be complex it will have constant argument. Once again, it is sufficient to determine the behaviour of $|A|$. For $0 < \alpha < 1$, the evolution is best illustrated by a phase portrait of $|A|$ versus B. There is an unstable fixed point at $(|A|, B) = (0, 0)$, and a stable fixed point at $(|A|, B) = (\alpha^{1/2}(1 - \alpha)^{1/2}/k, (1 - \alpha)/k^2)$. The behaviour of the system, illustrated in figure 9, is radically different to that when $\alpha = 0$. The system moves towards the stable fixed point as $\tau \to \infty$. Thus, $A(\tau)$ remains non-zero for large times, and hence there is a persistent response in the streamfunction and density perturbation fields. In contrast to the undamped case, the final state is independent of $|A_0|$. The long-memory property is destroyed.

We can assess the linear stability of the evolved mean flow using the approach of § 3.7. The evolving linear growth rate $s = \delta s_1 + O(\delta^2)$, where

$$s_1 = 1 - \alpha - k^2 B = \frac{\mathrm{d}A/\mathrm{d}\tau}{A}.$$

The stable fixed point corresponds exactly to $s_1 = 0$. Here, to leading order, the mean flow is exactly neutralized to linear disturbances. Thus, in phase space, the system moves from an unstable fixed point to a stable fixed point, corresponding to a mean flow change from a state of weak instability to one of neutral stability. As shown in figure 9, during the evolution, typically the magnitude B of the mean flow change will overshoot that required for neutral stability. However, the damping is continually trying to eradicate the mean flow change, so that B is subsequently reduced, leading to decaying oscillations about the stable fixed point.

This sort of damped regime was the focus of the numerical study by Hua *et al.* (1997). In their non-hydrostatic numerical simulations, typically $\delta \approx 0.3$ and $\alpha \approx 0.75$,

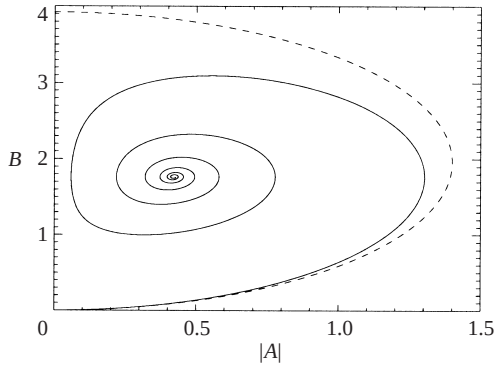

FIGURE 9. The streamfunction amplitude $|A(\tau)|$ against the mean flow change $B(\tau)$, as determined by the amplitude equations (5.2*a*) and (5.2*b*), for $|A_0| = 0.01$ and $k^2 = 0.510$: ——, evolution for $\alpha = 0.1$; – – –, evolution for $\alpha = 0$.

so that the system was close to equilibrium. They found an inertial instability with equilibration to a state with finite-amplitude baroclinic jets, exactly the sort of behaviour described by (5.2*a*) and (5.2*b*).

Hua *et al.* (1997) also used a large horizontal diffusivity ν_y in their simulations. As a consequence, even in their run at $\delta = 0.07$ with $\alpha = 0$ (see their figure 2), they found equilibration to a state with baroclinic jets (albeit of small amplitude). However, as we have seen in §§ 3 and 4, when $\alpha = 0$, the flow equilibrates to a barotropic state if $\nu_y = 0$. Presumably, their results were affected by the presence of horizontal diffusion, which we anticipate would act somewhat like a relaxational term on the mean flow.

Zhao & Ghil (1991) also considered the effects of Rayleigh friction and Newtonian cooling in a study of weakly nonlinear inertial instability. However, in their study, the weak nonlinearity came from artificially restricting the vertical wavenumber to be near the buoyancy cut-off wavelength, rather than by restricting the growth rate by vertical diffusion. Thus, the dynamics they found were quite different to those described here.

5.2. *An envelope of unstable modes*

As shown in § 3.1, and illustrated in figure 1, when $\epsilon = \epsilon_c(1 - \delta/\delta_0)$ there is a band of unstable wavenumbers of width $O(\delta^{1/2})$ around $m = m_c$. To take account of the possible simultaneous growth and interaction of all these modes, we can consider the evolution of a mode $\psi(y, z, t) = \mathrm{Re}\{A(Z, \tau)\Psi(y)\exp(im_c z)\}$, where $Z = \delta^{1/2}z$ is a slowly varying vertical coordinate. A modified weakly nonlinear analysis can be performed, similar to that of §§ 3.2–3.5, taking account of this new vertical structure. The details are given in Griffiths (2000). The horizontal structure of the mean flow change is the same as before, but the amplitude equations are

$$\frac{\partial B}{\partial \tau} = |A|^2 + \epsilon_c \frac{\partial^2 B}{\partial Z^2}, \quad \frac{\partial A}{\partial \tau} = A + \tfrac{1}{2} r^2 \frac{\partial^2 A}{\partial Z^2} - k^2 AB,$$

subject to $A(0) = A_0, B(0) = 0$, where $r^2 = D_{mm}(m_c, 0, \epsilon_c)/D_s(m_c, 0, \epsilon_c) > 0$, using (2.16) and (3.1). Here, $B(Z, \tau)$ is the amplitude of the mean flow change, as in (5.1). Owing to the diffusive form of the additional terms in this coupled system, it is anticipated that the dynamics will not differ considerably from those of the Z-independent system.

6. Conclusions

6.1. *Summary*

We have studied the zonally symmetric inertial instability of uniformly stratified equatorial shear flows of the form $u = u_0(y)$, with corresponding potential vorticity $Q_0(y)$. If $fQ_0 < 0$ somewhere, then the flow will be linearly unstable, provided the vertical diffusion is sufficiently small. However, as the vertical diffusion increases, the growth rate of the most unstable linear mode decreases monotonically, and eventually becomes zero. We chose to measure the strength of the instability by a non-dimensional parameter δ, related to the vertical diffusion, as well as to the strength of the stratification, the cross-equatorial shear and the β-effect, such that the most unstable linear mode has growth rate δ as $\delta \to 0$.

For the regime $\delta \ll 1$, we performed a weakly nonlinear analysis, deriving equations for the amplitude of the most unstable mode, and for the mean-flow change. By analysing the linear stability properties of the evolving mean flow, we gained a clear qualitative understanding of nonlinear inertial instability. There is disturbance growth then decay, accompanied by neutralization then stabilization of the mean flow. Further, the system has a long memory: the maximum disturbance amplitude and the magnitude of the mean-flow change depend on the initial disturbance amplitude. However, weak damping (in the form of Rayleigh friction and Newtonian cooling) radically alters the nonlinear evolution, leading to an evolved state with a neutrally stable mean flow, and a persistent streamfunction disturbance. The long-memory property is destroyed.

For the particular case of a uniform shear flow, numerical simulations showed that the quantitative applicability of the weakly nonlinear analysis is limited to just a small range of δ, perhaps $0 \leqslant \delta < 0.05$. However, the numerical simulations also suggest that the qualitative applicability is good, since for $0 \leqslant \delta < 0.3$ we see the same qualitative features. Further, in the moderately unstable regime, say $0.1 \leqslant \delta < 0.3$, the action of the instability on the mean flow is to create a region, around the initially most unstable latitude, in which $f\overline{Q}$ is homogenized to a small negative value. We might expect this sort of homogenization to be the generic action of inertial instability for any unstable mean flow in this regime, not just for that with uniform latitudinal shear. We also argued that, during the instability, the disturbances will necessarily shift polewards to effect an efficient neutralization, and noted that the line of zero mean potential vorticity moves polewards too, rather than equatorwards.

6.2. *Discussion*

The numerical simulations showed that for $\delta = 0.3$ there are regions of the flow in which the Richardson number falls to less than 0.25. Thus, in reality we would expect secondary Kelvin–Helmholtz instabilities to occur, and to significantly modify the evolution in this case. Since the Kelvin–Helmholtz instability occurs on small horizontal scales, and would have longitudinal dependence, large-scale zonally symmetric simulations are unable to capture this additional behaviour, and hence are unrealistic for $\delta \gtrsim 0.3$. Some possible consequences of this are discussed elsewhere (Griffiths 2002).

The imposition of zonal symmetry on both the analysis and the simulations leads to other restrictions. For instance, even though the uniform shear flow studied is barotropically stable, the simulations show that for $\delta \gtrsim 0.15$ the evolving mean flow tends to display regions of reversed potential vorticity gradients on both sides of the equator. Thus, if zonally asymmetric perturbations were permitted, the evolving

mean flow could be susceptible to barotropic shear instability. This idea has been somewhat explored by Limpasuvan *et al.* (2000), as a mechanism for the generation of the upper-stratospheric two-day wave. In a zonally symmetric model, it would be possible to investigate the effect of such instabilities on the evolution of the mean flow by incorporating a barotropic shear adjustment scheme (perhaps by adapting the approach of Haynes 1989).

The fully three-dimensional stability properties of the zonally symmetric flow have also been ignored. In the weakly nonlinear regime, it seems that zonally asymmetric modes of inertial instability are faster growing than zonally symmetric modes (Dunkerton 1983). Thus, even though the fastest growing zonally symmetric mode has been analysed, it may not be the fastest growing mode in three-dimensional reality. Whether the weakly nonlinear evolution of the zonally asymmetric modes is radically different from the zonally symmetric evolution remains to be seen.

6.3. *Relation to inertial instability in the stratosphere and mesosphere*

Structures attributable to equatorial inertial instability have been observed in the Earth's atmosphere at about 40–60 km altitude, that is in the upper stratosphere and lower mesosphere (e.g. Hitchman *et al.* 1987; Hayashi *et al.* 1998; Smith & Riese 1999). The observed temperature perturbations, with a vertical wavelength of about 10 km, typically appear close to the equator with antiphased structures about 30° into the winter hemisphere. Does this inertial instability fall into the weakly nonlinear regime considered here? Putting aside for the moment the physical basis for being in such a regime, and, in particular, the need for an unrealistically large vertical diffusion, let us compare the magnitude, temporal behaviour and spatial structure of the observations with those of the weakly nonlinear perturbations described here.

Given that the observed temperature perturbations are typically 5–8 K, we might be tempted to think that they are simply too large to be of a weakly nonlinear origin. However, for a structure with a vertical wavelength of 10 km, even vertical displacements of a fraction of a wavelength can lead to large temperature perturbations. We can quantify this. In the Boussinesq approximation, temperature perturbations T' are associated solely with density perturbations ρ', and not with pressure perturbations. Therefore, in dimensional variables, $T'/T_{00} \approx -\rho'/\rho_{00} = \sigma'/g$, where T_{00} is a background reference temperature. Or, in terms of the non-dimensional buoyancy acceleration fluctuation $\hat{\sigma}'$, the temperature perturbation will be $T' = (N^2 T_{00}/g m_s)\hat{\sigma}'$. For $\delta = 0.1$, and $u_0 = \Lambda y$, the maximum non-dimensional buoyancy acceleration fluctuation was found in the numerical simulations to be $\hat{\sigma}' = 0.27$. Hence, for a disturbance with a vertical wavelength of 10 km (i.e. $m_s \approx 6 \times 10^{-4}\,\mathrm{m}^{-1}$), and taking a background temperature $T_{00} = 260\,\mathrm{K}$ (typical of conditions around 50 km altitude), the maximum temperature perturbation $T' \approx 5\,\mathrm{K}$. Given that the maximum amplitude of the perturbations could be increased by some initial excitation, this scale is consistent with the observed temperature perturbations of 5–8 K.

What about the timescale for the growth and decay of the disturbances? The weakly nonlinear (e-folding) timescale for growth (or decay) is $(\frac{1}{2}\delta\Lambda)^{-1}$. Near solstice, Λ can be as large as 5 day^{-1} (see, for instance, figure 3 of Hitchman & Leovy 1986), so that for $\delta = 0.1$, we obtain a timescale of about four days. The observed structures grow and decay over a period of about two weeks. Hence, the timescale is quite consistent with a weakly nonlinear instability.

Also note that the observed latitudinal and vertical scales are consistent with a weakly nonlinear inertial instability. For $\Lambda = 5\ \mathrm{day}^{-1}$, and taking $\beta = 2.3 \times 10^{-11}\,\mathrm{m}^{-1}\,\mathrm{s}^{-1}$, the unstable region extends from the equator to about 20° into the winter

hemisphere. According to the linear theory, the poleward temperature extremum must occur equatorward of 20°. However, the simulations showed how, during the nonlinear phase of the instability, this temperature extremum moves further polewards, to $y \approx 1.4\Lambda/\beta$, for $\delta \approx 0.2$. Thus, at maximum disturbance amplitude, for $\delta \approx 0.2$, we expect to see the maximum temperature perturbation about 30° into the winter hemisphere. This is consistent with the observations, whereas to explain the position of the poleward temperature extrema using the linear theory, we have to invoke unappealingly large horizontal shears.

However, the background conditions are unlikely to be consistent with a simple weakly nonlinear scenario. Taking $N = 0.02\,\mathrm{s}^{-1}$, for the most unstable vertical wavelength to be 10 km, the minimum shear required is 5.2 day^{-1}. The corresponding vertical diffusivity is $33\,\mathrm{m}^2\,\mathrm{s}^{-1}$, although this can be somewhat reduced if we consider even larger shears. Such diffusivities are usually regarded as unrealistically large. If somehow such a large diffusivity does exist in the upper stratosphere and lower mesosphere, then the magnitude, time scales and spatial structure of the observed instability are all consistent with a weakly nonlinear origin, of the type described here. It seems likely that this is not the case. The observed structures may well correspond to modes towards the buoyancy cut-off wavelength, with a relatively small linear growth rate compared with the inviscid maximum, but it may be that something other than diffusion is suppressing the shorter wavelength modes. A possible mechanism for this, involving the suppression of the shorter wavelength modes by a secondary Kelvin–Helmholtz instability, is discussed in Griffiths (2002). Even so, the basic nonlinear mechanisms highlighted in this simple study are likely to remain relevant in more complicated settings, such as is likely in the real atmosphere.

Part of this work was supported by NERC under award number GT4/96/40/M. Special thanks must be given to Michael McIntyre and Peter Haynes. Other helpful comments or assistance have been given by Mike Greenslade and Jacques Vanneste.

Appendix

We wish to investigate the variation of the growth rate s for inertially unstable modes as we vary ϵ, for a fixed basic flow in physical space. Unfortunately, in our non-dimensionalization, both f and Q_0 become functions of y and ϵ. In this Appendix, it proves convenient to use a non-dimensionalization in which fQ_0 remains constant as we vary ϵ, for a fixed basic flow in physical space. We ensure this by working with a rescaled (non-dimensional) coordinate $\tilde{y} = \epsilon^{1/2} y$, so that $f = y_c + 2\tilde{y}$, and Q_0 is a function of $\tilde{y}$ alone. With this small change, we write the eigenvalue equation (2.11) as

$$\frac{4\epsilon^2}{m^2}\frac{\mathrm{d}^2\Psi_i}{\mathrm{d}\tilde{y}^2} - fQ_0(\tilde{y})\Psi_i = \lambda_i\Psi_i, \quad |\Psi_i| \to 0 \text{ as } |\tilde{y}| \to \infty, \tag{A 1}$$

where $\lambda_i(m,\epsilon) = (s_i + \epsilon m^2)^2$, and $i \geqslant 0$. We assume that $Q_0 \sim f$ as $|\tilde{y}| \to \infty$. Then, according to Sturm–Liouville theory (e.g. Titchmarsh 1962), there exists a discrete, infinite spectrum of distinct eigenvalues. At given (m,ϵ), we order the eigenvalues so that $\lambda_i > \lambda_{i+1}$. The corresponding normalized eigenfunctions $\Psi_i(\tilde{y})$ satisfy $\int_{-\infty}^{\infty} \Psi_i\Psi_j \mathrm{d}\tilde{y} = \delta_{ij}$, and are complete for sufficiently well-behaved functions decaying to zero as $|\tilde{y}| \to \infty$.

We introduce the functional

$$I(\Phi, m, \epsilon) = \frac{\displaystyle\int_{-\infty}^{\infty} (-fQ_0)\Phi^2 \mathrm{d}\tilde{y} - (4\epsilon^2/m^2)\int_{-\infty}^{\infty} (\mathrm{d}\Phi/\mathrm{d}\tilde{y})^2 \mathrm{d}\tilde{y}}{\displaystyle\int_{-\infty}^{\infty} \Phi^2 \mathrm{d}\tilde{y}}, \tag{A 2}$$

where $\Phi(\tilde{y})$ is a sufficiently well-behaved function decaying to zero as $|\tilde{y}| \to \infty$. It is easy to see that $I(\Psi_i, m, \epsilon) = \lambda_i(m, \epsilon)$, and by writing $\Phi(\tilde{y}) = \sum_{i=0}^{\infty} c_i \Psi_i(\tilde{y})$ that

$$I(\Phi, m, \epsilon) = \frac{\sum \lambda_i c_i^2}{\sum c_i^2} \leqslant \lambda_0(m, \epsilon) \leqslant (-fQ_0)_{max}. \tag{A 3}$$

In particular, at $\epsilon = \epsilon_a$ suppose there exist growing modes, and let the maximum value of $s_0(m, \epsilon_a)$ with respect to m occur at $m = m_a$. We denote this maximum value by s_a, and the corresponding eigenfunction by Ψ_a. Then, when $\epsilon = \epsilon_b < \epsilon_a$,

$$\lambda_0(m_a, \epsilon_b) \geqslant I(\Psi_a, m_a, \epsilon_b) > I(\Psi_a, m_a, \epsilon_a) = \lambda_0(m_a, \epsilon_a),$$

the first inequality following using (A 3), and the second using (A 2) since $\epsilon_b < \epsilon_a$. Equivalently,

$$(s_0(m_a, \epsilon_b) + \epsilon_b m_a^2)^2 > (s_a + \epsilon_a m_a^2)^2.$$

Taking the positive square root for the growing modes, we have

$$s_0(m_a, \epsilon_b) > s_a + (\epsilon_a - \epsilon_b)m_a^2 > s_a \quad \text{since} \quad \epsilon_b < \epsilon_a.$$

The maximum growth rate s_b at $\epsilon = \epsilon_b$ satisfies $s_b \geqslant s_0(m_a, \epsilon_b)$. Thus, $s_b > s_a$, and the maximum growth rate monotonically increases as ϵ decreases.

We conclude by deriving the limiting value of s as $\epsilon \to 0$. Since we have taken the maximum value of $(-fQ_0)$ to occur at $y = 0$, and hence at $\tilde{y} = 0$,

$$0 \leqslant (-fQ_0)_{max} - (-fQ_0) = \int_0^{\tilde{y}} \frac{\mathrm{d}}{\mathrm{d}\tilde{y}}(fQ_0)\,\mathrm{d}\tilde{y} \leqslant \mathrm{sgn}(\tilde{y}) \int_0^{\tilde{y}} \left|\frac{\mathrm{d}}{\mathrm{d}\tilde{y}}(fQ_0)\right| \mathrm{d}\tilde{y}. \tag{A 4}$$

Since $f \sim Q_0 \sim 2\tilde{y}$ as $|\tilde{y}| \to \infty$, $|\mathrm{d}(fQ_0)/\mathrm{d}\tilde{y}| < 8C|\tilde{y}|$ for some constant $C > 0$. Then (A 4) $\Rightarrow 0 \leqslant (-fQ_0)_{max} - (-fQ_0) < 4C\tilde{y}^2$, so that, using (A 2),

$$(-fQ_0)_{max} - I(\Phi, m, \epsilon) \leqslant \frac{\displaystyle\int_{-\infty}^{\infty} 4C\tilde{y}^2\Phi^2 \mathrm{d}\tilde{y} + (4\epsilon^2/m^2)\int_{-\infty}^{\infty} (\mathrm{d}\Phi/\mathrm{d}\tilde{y})^2 \mathrm{d}\tilde{y}}{\displaystyle\int_{-\infty}^{\infty} \Phi^2 \mathrm{d}\tilde{y}}.$$

Taking $\Phi = \exp(-m\tilde{y}^2/2\epsilon)$ gives $(-fQ_0)_{max} - I \leqslant 2\epsilon(C+1)/m$. Then, using (A 3), we have

$$(-fQ_0)_{max} - \frac{2\epsilon(C+1)}{m} \leqslant \lambda_0(m, \epsilon) \leqslant (-fQ_0)_{max}$$

for the eigenvalue λ_0. Setting $m = 1$ and letting $\epsilon \to 0$, we see that $\lambda_0 \to (-fQ_0)_{max}$, or equivalently $s_0^2 \to (-fQ_0)_{max}$. Thus, if $(-fQ_0)_{max} < 0$, $s_0 \to ((-fQ_0)_{max})^{1/2}$ for some mode as $\epsilon \to 0$, i.e. the maximum possible growth rate is achieved.

REFERENCES

Bowman, J. C. & Shepherd, T. G. 1995 Nonlinear symmetric stability of planetary atmospheres. *J. Fluid Mech.* **296**, 391–407.

Boyd, J. P. 2000 *Chebyshev and Fourier Spectral Methods.* Dover.

Boyd, J. P. & Christidis, Z. D. 1982 Low wavenumber instability on the equatorial beta-plane. *Geophys. Res. Lett.* **9**, 769–772.

Clark, P. D. & Haynes, P. H. 1996 Inertial instability on an asymmetric low-latitude flow. *Q. J. R. Met. Soc.* **122**, 151–182.

Drazin, P. G. & Reid, W. H. 1981 *Hydrodynamic Stability*. Cambridge University Press.

Dunkerton, T. J. 1981 On the inertial stability of the equatorial middle atmosphere. *J. Atmos. Sci.* **38**, 2354–2364.

Dunkerton, T. J. 1982 The double-diffusive modes of symmetric stability on an equatorial beta-plane. *J. Atmos. Sci.* **39**, 1653–1657.

Dunkerton, T. J. 1983 A nonsymmetric equatorial inertial instability. *J. Atmos. Sci.* **40**, 807–813.

Dunkerton, T. J. 1993 Inertial instability of a nonparallel flow on an equatorial beta-plane. *J. Atmos. Sci.* **50**, 2744–2758.

Edwards, N. R. & Richards, K. J. 1999 Linear double-diffusive-inertial instability at the equator. *J. Fluid Mech.* **395**, 295–319.

Griffiths, S. D. 2000 Inertial instability in the equatorial stratosphere. PhD thesis, University of Cambridge. Available from Superintendent of Manuscripts, University Library, West Road, Cambridge, CB3 9DR, United Kingdom.

Griffiths, S. D. 2002 Nonlinear vertical scale selection in equatorial inertial instability. *J. Atmos. Sci.* (in press).

Hayashi, H., Shiotani, M. & Gille, J. C. 1998 Vertically stacked temperature disturbances near the equatorial stratopause as seen in cryogenic limb array etalon spectrometer data. *J. Geophys. Res.* **103**, 19 469–19 483.

Haynes, P. H. 1989 The effect of barotropic instability on the nonlinear evolution of a Rossby-wave critical layer. *J. Fluid Mech.* **207**, 231–266.

Hitchman, M. H. & Leovy, C. B. 1986 Evolution of the zonal mean state in the equatorial middle atmosphere during October 1978–May 1979. *J. Atmos. Sci.* **43**, 3159–3176.

Hitchman, M. H., Leovy, C. B., Gille, J. C. & Bailey, P. B. 1987 Quasi-stationary zonally asymmetric circulations in the equatorial lower mesosphere. *J. Atmos. Sci.* **44**, 2219–2236.

Hua, B. L., Moore, D. W. & Le Gentil, S. 1997 Inertial nonlinear equilibration of equatorial flows. *J. Fluid Mech.* **331**, 345–371.

Limpasuvan, V., Leovy, C. B., Orsolini, Y. J. & Boville, B. A. 2000 A numerical simulation of the two-day wave near the stratopause. *J. Atmos. Sci.* **57**, 1702–1717.

Phillips, N. A. 1966 The equations of motion for a shallow rotating atmosphere and the 'traditional approximation'. *J. Atmos. Sci.* **23**, 626–628.

Potylitsin, P. G. & Peltier, W. R. 1998 Stratification effects on the stability of columnar vortices on the f-plane. *J. Fluid Mech.* **335**, 45–79.

Rayleigh, Lord 1917 On the dynamics of revolving fluids. *Proc. R. Soc. Lond.* A **93**, 148–154.

Shen, C. Y. & Evans, T. E. 1998 Inertial instability of large Rossby number horizontal shear flows in a thin homogeneous layer. *Dyn. Atmos. Oceans* **26**, 185–208.

Smith, A. K. & Riese, M. 1999 Cryogenic infrared spectrometers and telescopes for the atmosphere (CRISTA) observations of tracer transport by inertially unstable circulations. *J. Geophys. Res.* **104**, 19 171–19 182.

Taylor, G. I. 1923 Stability of a viscous liquid contained between two rotating cylinders. *Phil. Trans. R. Soc. Lond.* A **223**, 289–343.

Thorpe, A. J. & Rotunno, R. 1989 Nonlinear aspects of symmetric instability. *J. Atmos. Sci.* **46**, 1285–1299.

Titchmarsh, E. C. 1962 *Eigenfunction Expansions, Part I*, 2nd edn. Oxford University Press.

Tomas, R. A. & Webster, P. J. 1997 The role of inertial instability in determining the location and strength of near-equatorial convection. *Q. J. R. Meteorol. Soc.* **123**, 1445–1482.

Zhao, J.-X. & Ghil, M. 1991 Nonlinear symmetric instability and intraseasonal oscillations in the tropical atmosphere. *J. Atmos. Sci.* **48**, 2552–2568.