J. Fluid Mech. (1999), *vol.* 396, *pp.* 73–108. Printed in the United Kingdom

Nonlinear evolution of the elliptical instability: an example of inertial wave breakdown

By D. M. MASON AND R. R. KERSWELL
Department of Mathematics, University of Bristol, Bristol, BS8 1TW, UK

(Received 22 March 1999 and in revised form 30 April 1999)

A direct numerical simulation is presented of an elliptical instability observed in the laboratory within an elliptically distorted, rapidly rotating, fluid-filled cylinder (Malkus 1989). Generically, the instability manifests itself as the pairwise resonance of two different inertial modes with the underlying elliptical flow. We study in detail the simplest 'subharmonic' form of the instability where the waves are a complex conjugate pair and which at weakly supercritical elliptical distortion should ultimately saturate at some finite amplitude (Waleffe 1989; Kerswell 1992). Such states have yet to be experimentally identified since the flow invariably breaks down to small-scale disorder. Evidence is presented here to support the argument that such weakly nonlinear states are never seen because they are either unstable to secondary instabilities at observable amplitudes or neighbouring competitor elliptical instabilities grow to ultimately disrupt them. The former scenario confirms earlier work (Kerswell 1999) which highlights the generic instability of inertial waves even at very small amplitudes. The latter represents a first numerical demonstration of two competing elliptical instabilities co-existing in a bounded system.

1. Introduction

It is now well known that a large class of two-dimensional, inviscid flows with patches of elliptical streamlines are subject to strong three-dimensional instabilities (Gledzer *et al.* 1974, 1975; Vladimirov & Tarasov 1985; Vladimirov & Vostretsov 1986; Pierrehumbert 1986; Bayly 1986; Landman & Saffman 1987; Craik 1989; Malkus 1989; Waleffe 1990; Gledzer & Ponomarev 1992). Attention has naturally focused on this 'elliptical' instability as a likely mechanism by which large-scale, two-dimensional coherent structures may break down into small-scale, three-dimensional motions in the transition to turbulence of shear flows (Orszag & Patera 1983; Bayly, Orszag & Herbert 1988; Malkus & Waleffe 1991; Caulfield & Peltier 1999 for a discussion in the mixing layer context). Laboratory experiments of a strained vortex (Malkus 1989) have demonstrated that this elliptical instability ultimately does lead to small-scale disorder as have recent numerical simulations starting from a periodic array of such vortices (Lundgren & Mansour 1996). However, theoretical understanding of this process at present extends no further than the initial instability mechanism which is still being mapped out (Moffatt, Kida & Ohkitani 1994; Bayly, Holm & Lifschitz 1996; Jimenez, Moffatt & Vasco 1996).

The purpose of this paper is to make progress in understanding the secondary, nonlinear phases of the elliptical instability which exist beyond the initial stage of exponential growth. We focus our attention on Malkus's (1989) clever experiments in which he created an elliptical flow by rotating a water-filled flexible cylinder between

two stationary rollers. The elliptical instability manifests itself as the pairwise resonance of two inertial waves coupled by the underlying 'elliptical wave' of deformation in the basic flow. The unstable waves are seen to grow for some time and then dramatically collapse into small-scale disorder. This appears to be a classic example of inertial wave breakdown observed in many other systems where inertial waves are driven to finite amplitude (e.g. Malkus 1968; McEwan 1970; Manasseh 1992, 1994, 1996; Kobine 1995, 1996) and represents a fundamental problem in rapidly rotating fluids. Malkus's experiments seem a particularly clean example of this phenomenon christened 'resonant collapse' by McEwan (1970).

The approach taken here is one of direct numerical simulation (DNS) specifically aimed at investigating the slightly supercritical regime where weakly nonlinear analysis should be applicable. Such analysis already carried out for the experimental set-up of rigid top and bottom boundaries (Waleffe 1989; Kerswell 1992, unpublished) predicts the existence of small-amplitude saturated states. However, one of the more puzzling aspects of Malkus's experiments is the invariable collapse of the growing inertial waves despite every effort to isolate such weakly nonlinear saturated states (Malkus, private communication). This DNS approach also complements a recent study highlighting the linear instability of finite-amplitude, inertial waves (Kerswell 1999) which specifically considers the growing waves seen by Malkus.

Working within Malkus's elliptical cylinder geometry necessarily requires a new Navier–Stokes pseudospectral solver to be developed. We have based this solver on Marcus's time-splitting algorithm which he used with success to examine Taylor–Couette flow (Marcus 1984*a*, *b*). Our non-standard geometry is handled by working in a *non-orthogonal* elliptico-polar coordinate system which effectively converts the elliptical cylinder into a circular cylinder but at the inevitable expense of producing more complicated equations of motion. These equations, however, possess highly desirable features (discussed below) which lead to considerable numerical efficiency compared with other equations produced by more conventional choices of orthogonal coordinate systems such as elliptical cylindrical coordinates, as noted elsewhere (Kerswell & Davey 1996). Once within a standard circular system, we employ Marcus's time-splitting algorithm (with Green's function correction) to step the primitive variables forward in time. The novelty of approach taken here for this geometry makes this work of independent numerical interest.

This paper has two main aims: to establish the existence of a saturated nonlinear state for the elliptical instability (under suitably restricted conditions), and then to examine its stability. The general hypothesis for which we seek evidence here is that the saturated state suggested by weakly nonlinear analyses is never actually *seen* because secondary instabilities of the growing primary inertial waves preempt such equilibration. This would then reiterate the conclusion of recent more speculative work (Kerswell 1999).

The plan of the paper is as follows. Section 2 formulates the problem and introduces the elliptico-polar coordinate system. Section 3 describes how the governing Navier–Stokes equations in this system are stepped forward in time using Marcus's algorithm. A complementary eigenvalue code which examines the linear stability of the elliptical basic state is also presented here. Section 4 collects together the results in five subsections. The first two discuss the elliptical instability mechanism and the results from both the eigenvalue code and a linearized version of the time-stepping code. The emphasis here is to extend the asymptotic linear stability analysis of Waleffe (1989) (valid in the rapid-rotation limit) to finite elliptical distortion, rotation and hence viscosity as appropriate for Malkus's experiments and to test the time-stepping

code for accuracy. The third and fourth subsections present the first nonlinear results concerning equilibration of the elliptical instability and compare them with the predictions of weakly nonlinear analysis for the present case of stress-free top and bottom boundaries which is described in the Appendix. A final fifth subsection addresses secondary stability issues and demonstrates how other competitor elliptical instabilities can become important. Section 5 brings together the various results presented in the paper, discusses their implications and gives some final conclusions.

2. Formulation

The fluid flow under consideration is that produced within a rotating cylinder whose flexible walls are deflected radially inwards as it is rotated by two diametrically opposed rollers aligned with the cylinder's rotation axis and stationary in the laboratory frame (Malkus 1989). This gives rise to an elliptical cylinder of cross-section

$$\frac{x^2}{1+\beta} + \frac{y^2}{1-\beta} = 1 \tag{2.1}$$

(non-dimensionalized by the undistorted cylindrical radius, S†) fixed in the laboratory frame but whose walls are rotating. The ellipticity parameter β measures the elliptical distortion through the cross-section aspect ratio

$$A = \sqrt{\frac{1+\beta}{1-\beta}}. \tag{2.2}$$

An exact solution to the Navier–Stokes equations which satisfies all the relevant boundary conditions at least approximately for $\beta \ll 1$ is the basic elliptical flow

$$\boldsymbol{U} = -\sqrt{\frac{1+\beta}{1-\beta}}\, y\, \hat{\boldsymbol{x}} + \sqrt{\frac{1-\beta}{1+\beta}}\, x\, \hat{\boldsymbol{y}} \tag{2.3}$$

(non-dimensionalized by the magnitude of the cylinder's angular velocity Ω and the cylinder's radius S). Malkus's cylinder has a rigid circular base and a flexible top to accommodate the imposed elliptical distortion. This in general introduces $O(\beta)$ mismatches in the imposed tangential boundary velocity field and that of the solution (2.3). Additionally, since the speed of the cylinder wall is constant whereas that of (2.3) varies as $\sqrt{1-\beta\cos 2\phi}$ ($x = \sqrt{1+\beta}\cos\phi$, $y = \sqrt{1-\beta}\sin\phi$), there is also a small $O(\beta)$ discrepancy in the tangential velocity at the sidewalls. Asymptotically in the rapid rotation limit, we can expect both these discrepancies to have an $O(\beta E^{1/2})$ (E is the Ekman number defined below) and hence negligible effect on the interior flow. Certainly in the experiments there is every reason to suspect that the flow (2.3) is realized (along with weak boundary layers) at low values of the elliptical distortion β (Malkus, private communication) and we assume this henceforth through the imposed boundary conditions used below.

We decompose the total velocity and pressure fields into basic flow and modification components as follows:

$$\boldsymbol{u}_{total} = \boldsymbol{U} + \boldsymbol{u}, \quad p_{total} = P + p, \tag{2.4}$$

† Technically, the circumference of the cylinder remains constant under distortion in the laboratory so that the undistorted cylinder radius is $S\sqrt{1-\beta^2}$ but since $\beta \ll 1$ this distinction is unimportant in what follows.

where P is the pressure field associated with the basic elliptical flow (2.3). The governing Navier–Stokes equations are then

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{U}\cdot\nabla\boldsymbol{u} + \boldsymbol{u}\cdot\nabla\boldsymbol{U} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} + \nabla p = E\nabla^2\boldsymbol{u} \tag{2.5}$$

with the condition for fluid incompressibility

$$\nabla\cdot\boldsymbol{u} = 0. \tag{2.6}$$

The Ekman number $E = \nu/\Omega S^2$ is the usual non-dimensional measure of the fluid's kinematic viscosity ν. Boundary conditions imposed are non-slip on the sidewalls,

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on} \quad \frac{x^2}{1+\beta} + \frac{y^2}{1-\beta} = 1, \tag{2.7}$$

and stress-free on the top ($z = d$) and bottom ($z = 0$) surfaces,

$$\boldsymbol{u}\cdot\hat{\boldsymbol{z}} = 0, \quad \hat{\boldsymbol{z}}\cdot\nabla(\hat{\boldsymbol{z}}\times\boldsymbol{u}) = \boldsymbol{0} \quad \text{on} \quad z = 0, d. \tag{2.8}$$

The latter represent an undesirable but necessary compromise since Malkus's experiments have rigid boundaries top and bottom. Numerically resolving Ekman boundary layers in both radial *and* axial directions at experimental Ekman numbers of $O(10^{-4})$ for a number of inertial waves simultaneously remains prohibitively expensive with current computational facilities (Kerswell & Barenghi 1995). Hence we adopt the least dangerous simplification by relaxing the boundary conditions on the horizontal surfaces as in Kerswell (1999). The problem is then to identify the preferred fluid response to given values of the elliptical distortion β and Ekman number E in a cylinder of height-to-radius ratio d.

In order to work efficiently with spectral representations of the physical fields, the elliptical boundary must be describable by one of the spatial coordinates being constant. With this in mind, a natural transformation to make is to orthogonal elliptic cylindrical coordinates (ζ, η, z) where

$$x = \frac{\sqrt{2}\cosh\zeta}{\sqrt{\cosh 2\zeta_0}}\cos\eta, \quad y = \frac{\sqrt{2}\sinh\zeta}{\sqrt{\cosh 2\zeta_0}}\sin\eta, \tag{2.9}$$

with

$$0 \leqslant \eta < 2\pi, \quad 0 \leqslant \zeta \leqslant \zeta_0 = \frac{1}{2}\log\left(\frac{A+1}{A-1}\right)$$

so that $\zeta = \zeta_0$ defines the sidewall. However, this coordinate system is problematic for two reasons, both of which essentially stem from the fact that, apart from at the boundary, it does not reflect the structure of the basic flow. Curves of constant ζ are confocal ellipses rather than the similar ellipses of the basic flow streamlines and collapse down as $\zeta \to 0$ to a singular line segment joining the foci on the major axis. This inevitably leads to a degradation in any numerical scheme based on this coordinate representation. Secondly and more seriously, any discretization matrices based on spectral expansions of the physical variables will be *dense* matrices due to the awkward form of the transformation. This severely limits the range of parameter space which can be explored computationally.

A far more suitable system proves to be the *non-orthogonal* elliptico-polar coordinates (s, ϕ, z) defined by

$$x = s\sqrt{1+\beta}\cos\phi, \quad y = s\sqrt{1-\beta}\sin\phi \tag{2.10}$$

which maps the elliptical streamlines of (2.3) into circles. Bayly was the first to realize the relevance of this coordinate system to elliptical flow, stimulating the hydrodynamic analysis of a distorted cylinder by Waleffe (1989); Vladimirov & Vostretsov (1986) used this coordinate system independently in the Russian literature. The base vectors

$$\left.\begin{aligned}\tilde{\boldsymbol{s}} &= \sqrt{1+\beta}\cos\phi\,\hat{\boldsymbol{x}} + \sqrt{1-\beta}\sin\phi\,\hat{\boldsymbol{y}},\\ \tilde{\boldsymbol{\phi}} &= -\sqrt{1+\beta}\sin\phi\,\hat{\boldsymbol{x}} + \sqrt{1-\beta}\cos\phi\,\hat{\boldsymbol{y}},\end{aligned}\right\} \tag{2.11}$$

are not orthogonal and are left unnormalized so that the 'cylindrical-polar' expressions

$$\frac{\partial\tilde{\boldsymbol{s}}}{\partial\phi} = \tilde{\boldsymbol{\phi}},\ \frac{\partial\tilde{\boldsymbol{\phi}}}{\partial\phi} = -\tilde{\boldsymbol{s}} \tag{2.12}$$

hold. As a consequence the advective derivative and divergence operators are then identical to the familiar 'cylindrical-polar' expressions:

$$\boldsymbol{u}\cdot\nabla = u_x\frac{\partial}{\partial x} + u_y\frac{\partial}{\partial y} + u_z\frac{\partial}{\partial z} = u\frac{\partial}{\partial s} + \frac{v}{s}\frac{\partial}{\partial\phi} + w\frac{\partial}{\partial z}, \tag{2.13}$$

$$\nabla\cdot\boldsymbol{u} = \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z} = \frac{1}{s}\frac{\partial(su)}{\partial s} + \frac{1}{s}\frac{\partial v}{\partial\phi} + \frac{\partial w}{\partial z}, \tag{2.14}$$

where the new velocity components u, v and w are defined as follows:

$$\boldsymbol{u} = u_x\hat{\boldsymbol{x}} + u_y\hat{\boldsymbol{y}} + u_z\hat{\boldsymbol{z}} = u\tilde{\boldsymbol{s}} + v\tilde{\boldsymbol{\phi}} + w\hat{\boldsymbol{z}}. \tag{2.15}$$

The elliptical boundary is now simply $s = 1$ and the basic state (2.3) transforms to

$$\boldsymbol{U} = s\tilde{\boldsymbol{\phi}}, \tag{2.16}$$

which is now axisymmetric because all the azimuthal (elliptical) variation is hidden in $\tilde{\boldsymbol{\phi}}$. The price paid for this simplification however is the increased complexity of the Navier–Stokes equations caused by the non-orthogonality of the transformation. The complementary set of vectors

$$\tilde{\boldsymbol{l}} = \left[\frac{\cos\phi}{\sqrt{1+\beta}}, \frac{\sin\phi}{\sqrt{1-\beta}}, 0\right],\ \tilde{\boldsymbol{n}} = \left[\frac{-\sin\phi}{\sqrt{1+\beta}}, \frac{\cos\phi}{\sqrt{1-\beta}}, 0\right],\ \hat{\boldsymbol{z}}, \tag{2.17}$$

chosen so that for example $\tilde{\boldsymbol{l}}\cdot\tilde{\boldsymbol{s}} = 1$, $\tilde{\boldsymbol{l}}\cdot\tilde{\boldsymbol{\phi}} = 0$ and $\tilde{\boldsymbol{l}}\cdot\hat{\boldsymbol{z}} = 0$, are used to project out three independent components of the momentum equation. These components are *exactly*

$$\begin{aligned}&\frac{\partial u}{\partial t} - 2v + \frac{\partial u}{\partial\phi} + \boldsymbol{u}\cdot\nabla u - \frac{v^2}{s} + \frac{1-\beta\cos2\phi}{1-\beta^2}\frac{\partial p}{\partial s} + \frac{\beta\sin2\phi}{1-\beta^2}\frac{1}{s}\frac{\partial p}{\partial\phi}\\ &\quad = E\left\{\overline{\nabla}^2 u - \frac{1+\beta\cos2\phi}{1-\beta^2}\frac{1}{s^2}\left[2\frac{\partial v}{\partial\phi} + u\right] - \frac{2\beta\sin2\phi}{1-\beta^2}\left[\frac{1}{s}\frac{\partial v}{\partial s} - \frac{v}{s^2}\right]\right\},\end{aligned} \tag{2.18}$$

$$\begin{aligned}&\frac{\partial v}{\partial t} + 2u + \frac{\partial v}{\partial\phi} + \boldsymbol{u}\cdot\nabla v - \frac{uv}{s} + \frac{\beta\sin2\phi}{1-\beta^2}\frac{\partial p}{\partial s} + \frac{1+\beta\cos2\phi}{1-\beta^2}\frac{1}{s}\frac{\partial p}{\partial\phi}\\ &\quad = E\left\{\overline{\nabla}^2 v + \frac{1+\beta\cos2\phi}{1-\beta^2}\frac{1}{s^2}\left[2\frac{\partial u}{\partial\phi} - v\right] + \frac{2\beta\sin2\phi}{1-\beta^2}\left[\frac{1}{s}\frac{\partial u}{\partial s} - \frac{u}{s^2}\right]\right\}\end{aligned} \tag{2.19}$$

and

$$\frac{\partial w}{\partial t} + \frac{\partial w}{\partial\phi} + \boldsymbol{u}\cdot\nabla w + \frac{\partial p}{\partial z} = E\overline{\nabla}^2 w \tag{2.20}$$

with crucially

$$\nabla \cdot \boldsymbol{u} = 0 \tag{2.21}$$

preserved and where

$$\overline{\nabla}^2 \equiv \frac{1-\beta\cos 2\phi}{1-\beta^2}\frac{\partial^2}{\partial s^2} + \frac{1+\beta\cos 2\phi}{1-\beta^2}\frac{1}{s}\frac{\partial}{\partial s} + \frac{1+\beta\cos 2\phi}{1-\beta^2}\frac{1}{s^2}\frac{\partial^2}{\partial \phi^2}$$
$$+\frac{2\beta\sin 2\phi}{1-\beta^2}\frac{1}{s}\frac{\partial^2}{\partial s\partial\phi} - \frac{2\beta\sin 2\phi}{1-\beta^2}\frac{1}{s^2}\frac{\partial}{\partial\phi} + \frac{\partial^2}{\partial z^2}.$$

The boundary conditions are non-slip on the sides,

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{at} \quad s = 1, \tag{2.22}$$

and stress-free conditions on the top and bottom,

$$\frac{\partial u}{\partial z} = \frac{\partial v}{\partial z} = w = 0 \quad \text{on} \quad z = 0, d. \tag{2.23}$$

Due to the complexity of the equations (2.18)–(2.20), we have chosen to work with the primitive variables, u, v, w and p, of the system. The crucial feature of these equations and ultimately of the coordinate system used is that ϕ only explicitly appears as $\sin 2\phi$ or $\cos 2\phi$. This means that if we expand the primitive variables, u, v, w and p, using Fourier modes in ϕ, only adjacent *even* Fourier modes and separately adjacent *odd* modes will interact through the linear time-stepping operator. (Of course, the nonlinear part of the time-stepping operator, the advective nonlinearity, mixes all the modes up as ever.) This simple coupling in the linear part of the equations allows crucial computational savings to be made which are discussed in the next section. The drawback of time-stepping the Navier–Stokes equations directly is maintaining the fluid's incompressibility or, put another way, time-stepping the pressure. A popular alternative to this is to build incompressibility implicity into the flow representation by using a toroidal-poloidal decomposition for example. This, however, requires taking successive curls of the Navier–Stokes equations and hence leads to even more unwieldy expressions. Moreover, storage becomes less efficient (matrix bandwidth increases – see below) and the code less numerically robust since higher spatial derivatives necessarily appear.

3. Numerics

3.1. *The time-stepping code*

We use Marcus's pseudospectral time-splitting algorithm to integrate our governing equations forward in time. A collocation approach is used in s and Galerkin projection over ϕ and z. His algorithm has four stages. The first stage steps the advective and nonlinear terms explicitly using the second-order Adams–Bashforth method. The second updates the pressure field by solving a Poisson equation and the third time-steps the viscous term implicitly using Backward Euler. Finally, a corrective fourth stage applies the full boundary conditions using a Green's function method. Marcus (1984*a*) gives a detailed account of these various stages so here we discuss only the adjustments made for our particular application.

The first and main modification is forced by the reduced symmetry group of an elliptical cylinder compared to a circular cylinder. The rotational symmetry of

Marcus's cylindrical annulus meant that he could treat each Fourier mode in ϕ (see (3.1)) separately except for when the advective nonlinear term is considered. Put another way, all his matrix calculations involved block diagonal matrices so each block could be considered in isolation. Our system of flow within an elliptical cylinder only has a Z_2 symmetry group in its cross-section (generated by reflecting in *both* major and minor axes of the ellipse) rather than the S^1 symmetry group of the axisymmetric cylinder. Consequently, a typical Fourier mode m is coupled through $O(\beta)$ terms to its nearest neighbours $m \pm 2$. This means that for a given axial wavenumber, we must consider all odd m modes together and all even m modes together. In other words, our matrices are partitioned into all odd and all even m and each is now block *tri*diagonal. Algorithms exist to exploit this structure but still the computational effort called for is substantially higher than the fully decoupled situation.

Secondly, we do not have a central core in our cylinder so that the coordinate axis is part of the solution domain. This can lead to numerical problems unless specific efforts are made to desensitize the code to this artificial singularity. We achieve this here by exploiting the representation degeneracy of cylindrical coordinates in which the points $(-s, \phi \pm \pi, z)$ and (s, ϕ, z) are exactly equivalent. This means that each velocity component and scalar pressure function has a definite parity in s determined by whether its corresponding azimuthal wavenumber m is even or odd (see the appendix of Kerswell & Davey 1996). Building the appropriate radial parity into the spectral representation of each field variable not only saves on storage but automatically instils the correct limiting behaviour near the axis. Computationally, we consider the domain $\{-1 \leqslant s \leqslant 1,\ 0 \leqslant \phi < \pi\}$ rather than viewing the interior of the pipe as the region $\{0 \leqslant s \leqslant 1,\ -\pi \leqslant \phi < \pi\}$. The solution in $-1 \leqslant s < 0$ can be constructed from that in $0 < s \leqslant 1$ through the known symmetries and so we need only collocate the equations over the positive zeros of $T_{2N}(s)$ and impose boundary conditions at $s = 1$. Most importantly, this means that the collocation points are at their sparsest near the axis – $O(1/2N)$ spacing – and at their densest – $O(1/4N^2)$ spacing – near the sidewall where Ekman boundary layers must be resolved.

The velocity and pressure fields are represented in the form

$$\begin{bmatrix} u \\ v \\ w \\ p \end{bmatrix} = \sum_{l=0}^{L} \sum_{n=1}^{N} \left\{ \sum_{m=-M,\, m\,\text{odd}}^{M} \begin{bmatrix} u_{lmn}(t)\Theta_{2n}(s)\cos l\alpha z \\ v_{lmn}(t)\Theta_{2n}(s)\cos l\alpha z \\ w_{lmn}(t)\Theta_{2n+1}(s)\sin l\alpha z \\ p_{lmn}(t)T_{2n-1}(s)\cos l\alpha z \end{bmatrix} e^{im\phi} \right.$$

$$\left. + \sum_{m=-M,\, m\,\text{even}}^{M} \begin{bmatrix} u_{lmn}(t)\Theta_{2n+1}(s)\cos l\alpha z \\ v_{lmn}(t)\Theta_{2n+1}(s)\cos l\alpha z \\ w_{lmn}(t)\Theta_{2n}(s)\sin l\alpha z \\ p_{lmn}(t)T_{2n-2}(s)\cos l\alpha z \end{bmatrix} e^{im\phi} \right\} \tag{3.1}$$

where $\boldsymbol{u}_{l-mn} = \boldsymbol{u}^*_{lmn}$ and $p_{l-mn} = p^*_{lmn}$ so that the total velocity and pressure are real and $\alpha = \pi/d$. Here $T_n(s) = \cos(n \cos^{-1} s)$ is the nth Chebyshev Polynomial and

$$\Theta_n(s) \equiv T_n(s) - T_{n-2}(s) \tag{3.2}$$

so that the boundary conditions are built into the spectral functions. The nonlinear term is calculated in physical space and dealiased.

3.2. *Performance*

Theoretically, the numerical error introduced by the time-stepping code over one revolution is $O(E\Delta t, (\Delta t)^2)$ (Marcus 1984*a*) where Δt is the time step. This is confirmed below (see § 4.2) and most importantly stays at this level even over many revolutions. For the Ekman numbers of interest, $2.5 \times 10^{-4} \leqslant E \leqslant 10^{-2}$, the code is essentially second order in time since $10^{-2} \leqslant \Delta t$ (the period of rotation is 2π in our non-dimensional units). Experience indicates that the code can fail for three distinct reasons and always in the same fashion. First, the radial truncation can be too small for the time step and Ekman number, the crucial combination being the small number $E\Delta t$ which appears in the viscous 'fractional' step. Secondly, the time step Δt can be too large so that the numerical errors swamp the integration. Thirdly, the velocity field can become too large for the time step. All lead to blow up of the fluid's divergence which is therefore constantly monitored throughout the time integration. (The precise descriptor we have used is the following:

$$\sum_{j=0}^{4}\sum_{k=0}^{5}\frac{\left|\frac{1}{s}\frac{\partial(su)}{\partial s}+\frac{1}{s}\frac{\partial v}{\partial \phi}+\frac{\partial w}{\partial z}\right|}{\left|\frac{1}{s}\frac{\partial(su)}{\partial s}\right|+\left|\frac{1}{s}\frac{\partial v}{\partial \phi}\right|+\left|\frac{\partial w}{\partial z}\right|}, \tag{3.3}$$

where the indices (j,k) label 30 sampling points ($s = \frac{1}{2}, \phi = 2\pi j/5, z = kd/10$). This number is found to give a quick global measure of the flow solution's divergence and remains $O(10^{-12})$ in a healthy run.)

The radial truncation required for a typical time step of $\Delta t = 0.05$ can vary from $N = 35$ at $E = 10^{-2}$ to $N = 90$ at $E = 2.5 \times 10^{-4}$. This is way above that required for acceptable accuracy and is entirely dictated by the stability requirements of the viscous step since the operator $(1 - E\Delta t\nabla^2)$ must be inverted there. This represents a serious bottleneck in the algorithm which prevents studying Ekman numbers much below $E = 2.5 \times 10^{-4}$.

The existence of a threshold flow amplitude (for a given time step) above which the code breaks down is reminiscent of the Courant–Friedrichs–Lewy (CFL) condition. Technically, this is not directly applicable here because we employ spectral methods and so our numerical domain is in fact global. However, pursuing its general spirit nevertheless yields sensible numbers: the CFL condition takes the form

$$\frac{|v|\Delta t}{\Delta s} \leqslant 1 \tag{3.4}$$

and taking $\Delta t = 0.05$ and $\Delta s = 1/N$, the maximal spacing near the axis, $|v| \leqslant 0.22$ for numerical stability at $N = 90$.

For the runs performed here at $\beta \leqslant 0.1$ and the elliptical instabilities studied, the azimuthal truncation $M = 4$ proves more than sufficient. Axial truncations of $L = 4$ and $L = 8$ are used to consider saturation and stability issues respectively. Typical storage and run times are as follows. For a truncation $(N, M, L) = (60, 4, 4)$ which requires 69 MB, a single revolution takes 4.9 mins on a 200 Mhz Sun Ultrasparc, whereas $(90, 4, 4)$ takes 151 MB and 10.2 mins, and $(90, 4, 8)$ 283 MB and 22.2 mins, all using $\Delta t = 0.05$.

3.3. *Eigenvalue code*

An eigenvalue code was also produced to examine the linear stability of the underlying elliptical flow as a necessary preliminary to a time-stepping approach. This stability

code closely resembles that written to investigate the linear stability of axial flow in an elliptical pipe (Kerswell & Davey 1996). The system there possesses a $Z_2 \times Z_2$ symmetry group generated by reflections in the cross-sectional major and minor axes. This partitions the possible disturbance space into four subspaces which may be then efficiently searched for instability (see expressions (3.1)–(3.4) of Kerswell & Davey 1996). Here, due to the different basic elliptical flow, the symmetry group is only Z_2 generated by the composite transformation of reflections in both axes. The disturbance space then only divides in two: an *odd* and an *even* space where the disturbance fields take the following respective forms:

$$\begin{bmatrix} u \\ v \\ w \\ p \end{bmatrix} = \sum_{n=1}^{N} \sum_{m=-M,\, m\ \mathrm{odd}}^{M} \begin{bmatrix} u_{mn}\Theta_{2n}(s) \\ v_{mn}\Theta_{2n}(s) \\ w_{mn}\Theta_{2n+1}(s) \\ p_{mn}T_{2n-1}(s) \end{bmatrix} \mathrm{e}^{\mathrm{i}m\phi + \mathrm{i}\alpha z + (\mathrm{i}\lambda+\sigma)t}, \tag{3.5}$$

$$\begin{bmatrix} u \\ v \\ w \\ p \end{bmatrix} = \sum_{n=1}^{N} \sum_{m=-M,\, m\ \mathrm{even}}^{M} \begin{bmatrix} u_{mn}\Theta_{2n+1}(s) \\ v_{mn}\Theta_{2n+1}(s) \\ w_{mn}\Theta_{2n}(s) \\ p_{mn}T_{2n-2}(s) \end{bmatrix} \mathrm{e}^{\mathrm{i}m\phi + \mathrm{i}\alpha z + (\mathrm{i}\lambda+\sigma)t}. \tag{3.6}$$

As noted before, these representations implicitly satisfy all boundary conditions.

4. Results

4.1. *Elliptical instability*

The elliptical instability mechanism is most clearly understood by ignoring viscosity in equations (2.18)–(2.20). Linearizing about the basic elliptical flow and moving into the rotating frame leads to the equation

$$\frac{\partial \boldsymbol{u}}{\partial t} + 2\hat{\boldsymbol{k}}\times\boldsymbol{u} + \overline{\nabla} p = \tfrac{1}{2}\beta\left[\mathrm{e}^{2\mathrm{i}(\phi+t)}\mathscr{N}\overline{\nabla} p + \mathrm{e}^{-2\mathrm{i}(\phi+t)}\mathscr{N}^{*}\overline{\nabla} p\right] \tag{4.1}$$

where

$$\mathscr{N} = \begin{bmatrix} 1 & \mathrm{i} & 0 \\ \mathrm{i} & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \overline{\nabla} = \tilde{\boldsymbol{s}}\frac{\partial}{\partial s} + \tilde{\boldsymbol{\phi}}\frac{1}{s}\frac{\partial}{\partial \phi} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z} \tag{4.2}$$

and * denotes complex conjugation. When $\beta = 0$, (4.1) reduces to the inertial wave problem for which a complete set of normal modes or inertial waves are known (see Greenspan 1968). For $0 < \beta \ll 1$, the right-hand side can be viewed as a small coupling term between these neutral inertial modes (Waleffe 1989). The underlying elliptical basic flow is essentially a wave of amplitude β with axial wavenumber 0, azimuthal wavenumber 2 and frequency 2 in the rotating frame. The coupling term can therefore link inertial waves whose frequencies differ by 2, azimuthal wavenumbers differ by 2 and whose axial wavenumbers are equal – familiar triad resonance conditions – to produce resonant growth of the two inertial waves.

Generically, two candidate inertial waves satisfying the axial and azimuthal wavenumber conditions will not be in resonance for a given geometry because the frequency condition is violated. However at certain 'resonant' cylinder heights, their frequencies can be brought into the required relationship and resonance achieved. The effect of reinstating viscosity is to smear out slightly the precise resonant frequency condition so that it only needs to be approximately satisfied (where 'approximately' depends on

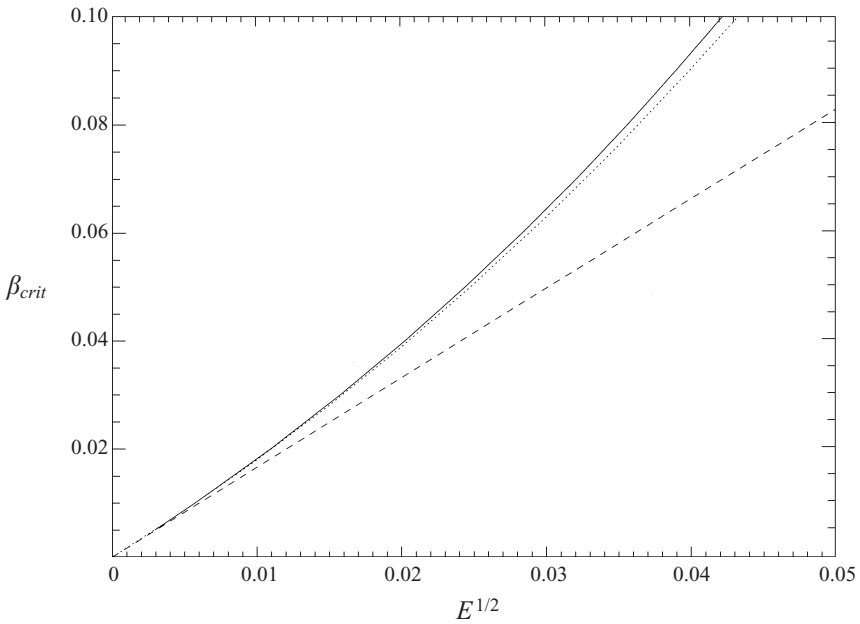

FIGURE 1. Plot of β_{crit} against $E^{1/2}$ for the $m = (-1, 1)$ elliptical instability. The solid line shows the results from the eigenvalue code and the dashed line is the leading asymptotic result $\beta_{crit} = 1.65E^{1/2}$. The dotted line is a modified asymptotic result based upon the more accurate decay rate expression $s_1 = -0.6189(1 + \mathrm{i}) - (k^2 + \alpha^2)E^{1/2}$ with $k = 2.7346$.

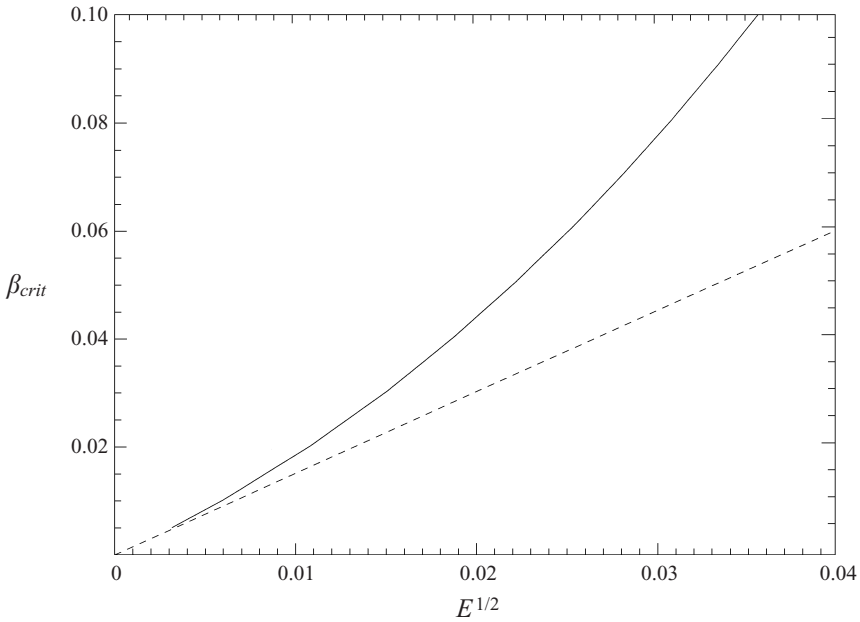

FIGURE 2. Plot of β_{crit} against $E^{1/2}$ for the $m = (0, 2)$ elliptical instability. The solid line shows the results from the eigenvalue code and the dotted line is the leading asymptotic result $\beta_{crit} = 1.48E^{1/2}$.

the size of the viscosity – see below) and to produce a threshold elliptical distortion $\beta_{crit} > 0$ only above which the elliptical instability is manifest. In this paper we choose to focus upon two particular 'resonant' geometries: $d = 3.9796$ in which an $m = 1$ inertial wave resonates with its $m = -1$ complex conjugate, and $d = 2.7009$ where an $m = 0$ wave resonates with an $m = 2$ wave. These instabilities have been the subject of extensive experimental investigation by Malkus. The first represents a typical example of the special subharmonic class of elliptical instability in which an inertial wave interacts with its complex conjugate and therefore offers the simplest situation for theoretical analysis. The second is a generic example of two different inertial waves interacting.

The threshold elliptical distortion, β_{crit}, for the onset of the $m = -1$ and $m = 1$ (hereafter $m = (-1, 1)$) elliptical instability at $d = 3.9796$ as a function of the Ekman number is plotted in figure 1. Here the results of eigenvalue calculations for finite E are compared with the leading-order asymptotic result $\beta_{crit} = 1.65E^{1/2}$ valid as $E \to 0$ (see Appendix equation (A 16), $C_1 = 0.5312$, $\Delta = E^{1/2}s_1^I$, $s_1 = -0.6189(1 + \mathrm{i})$). Figure 2 is the equivalent plot for the $m = 0$ and $m = 2$ (hereafter $m = (0, 2)$) elliptical instability at $d = 2.7009$ where the leading asymptotic result is $\beta_{crit} = 1.48E^{1/2}$. In both cases, the true β_{crit} quickly exceeds the leading-order asymptotic estimates which are based on just the first-order viscous decay rates of the inertial waves due to their boundary layers. Incorporating the next-order contribution due to interior dissipation (through the use of the more accurate expression $s_1 = -0.6189(1 + \mathrm{i}) - (k^2 + \alpha^2)E^{1/2}$ where $k = 2.7346$) in the $m = (-1, 1)$ case gives much better correspondence: see figure 1.

4.2. *Testing: linear code*

To test the main aspects of the time-stepping code, the results of time-stepping the linearized Navier–Stokes equations were compared with the results of the eigenvalue code for the two elliptical instabilities of interest here ($m = (-1, 1)$ at $d = 3.9796$ and $m = (0, 2)$ at $d = 2.7009$). The unstable eigenfunction was used as the initial condition for the time-stepping code in each case and then an estimate of the growth rate made based upon how the energy of the eigenfunction increased with time. Figure 3 shows how the fractional error in these growth rate estimates (as compared to the true eigenvalues) vary with choice of time step over the time period of 1 revolution. This figure contains information from both limiting Ekman numbers of $E = 2.5 \times 10^{-4}$ and $E = 10^{-2}$ for each elliptical instability.

The curves are all consistent with the expected time accuracy of the code which is $O(E\Delta t, (\Delta t)^2)$. When $E = 10^{-2}$, the dominant time-step error appears to be the $O(E\Delta t)$ error in the viscous time step: this is reflected in the straight line curves for both $m = (-1, 1)$ and $m = (0, 2)$. For $E = 2.5 \times 10^{-4}$, the dominant error should be the $O((\Delta t)^2)$ in the advective step and this is seen clearly in the $m = (0, 2)$ curve. The reason that this is not also seen in the $m = (-1, 1)$ curve is that this elliptical instability is *steady* in the laboratory frame in which the numerical code works (in contrast the $m = (0, 2)$ instability has a frequency $\lambda \approx 1$). The advective time-step errors are then presumably much smaller because of this and we see only $O(E\Delta t)$ behaviour. This steadiness, of course, makes this instability particularly amenable to numerical modelling. Figure 4 demonstrates that these time-step errors stabilize over many revolutions, which is especially important as we typically integrate over many hundreds of revolutions.

Based upon these results, the time steps used to study the $m = (-1, 1)$ and $m = (0, 2)$ instabilities are $\Delta t = 0.05$ and $\Delta t = 0.025$ respectively. Certainly in the $m = (-1, 1)$

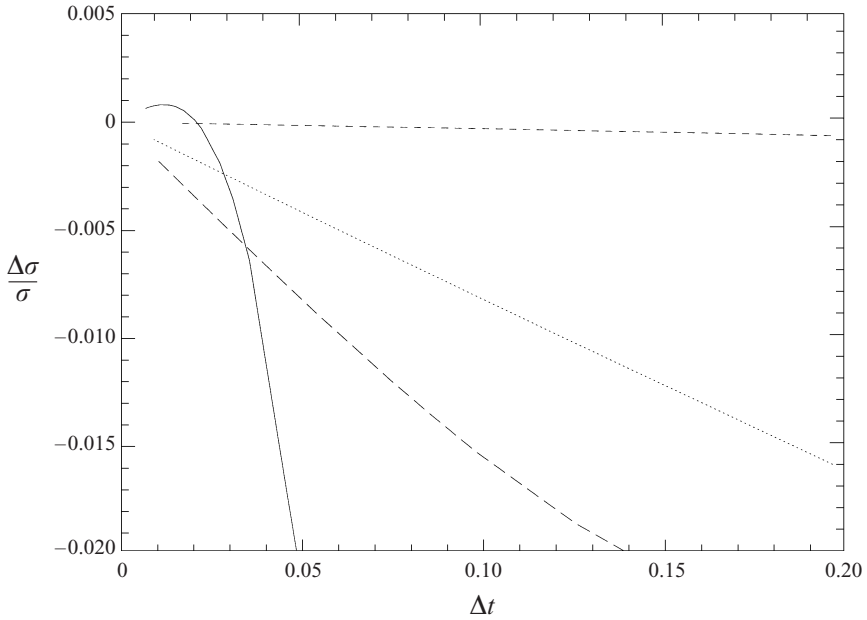

FIGURE 3. Plot of fractional error, $\Delta\sigma/\sigma$, in the growth rate as calculated by time-stepping over 1 revolution verses the time step Δt (1 revolution=2π in non-dimensional units). The (top) dashed and dotted curves correspond to $E = 2.5\times10^{-4}$ and $E = 10^{-2}$ respectively for the $m = (-1, 1)$ instability. The solid and long-dashed curves are the $E = 2.5 \times 10^{-4}$ and $E = 10^{-2}$ versions for the $m = (0, 2)$ instability. The elliptical distortion is $\beta = 0.1$ at $E = 10^{-2}$ and $\beta = 0.035$ at $E = 2.5 \times 10^{-4}$.

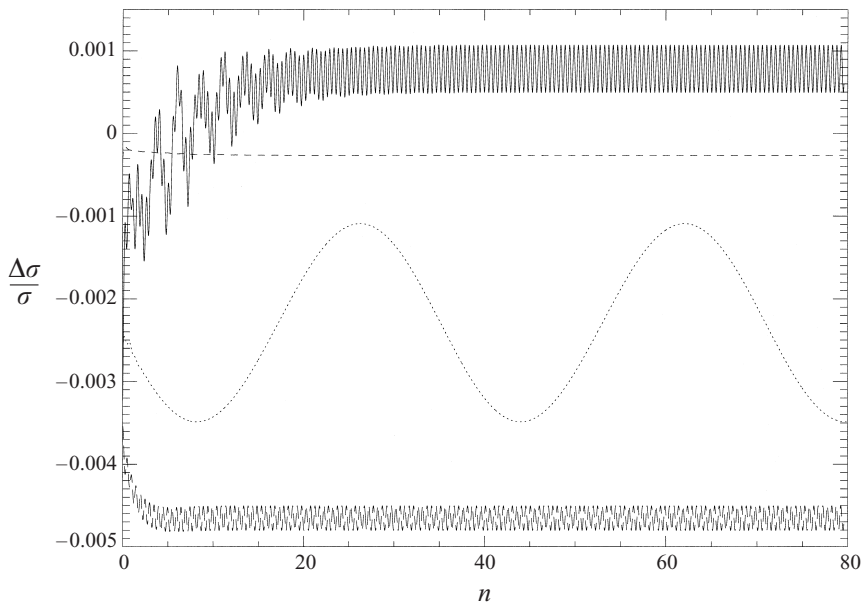

FIGURE 4. Plot of fractional error, $\Delta\sigma/\sigma$, against n, the number of revolutions ($t = 2\pi n$ in non-dimensional units). The (top) solid line is the $m = (0, 2)$ instability at $\beta = 0.035$, $E = 2.5\times10^{-4}$, $\Delta t = 0.025$. The (lowest) long-dashed line is the $m = (0, 2)$ instability at $\beta = 0.1$, $E = 10^{-2}$, $\Delta t = 0.025$. The dashed (second down) line is the $m = (-1, 1)$ instability at $\beta = 0.035$, $E = 2.5\times10^{-4}$, $\Delta t = 0.05$. The dotted (second up) line is the $m = (-1, 1)$ instability at $\beta = 0.1$, $E = 10^{-2}$, $\Delta t = 0.05$.

case, a larger time step seems acceptable and this is true in this situation. However, when nonlinearity is reinstated a large time step seems to restrict how far in amplitude the disturbance can be followed and experience indicated that $\Delta t = 0.05$ was the best choice.

4.3. *Testing: nonlinear code*

Marcus (1984*a*) tested his full nonlinear code by comparing the kinetic energy at a given time calculated from the time-stepped velocity with that determined by directly time-integrating the kinetic energy evolution equation. We have performed similar tests here. In elliptico-polar coordinates, the total kinetic energy is

$$\begin{aligned} K = {} & \frac{1}{2}\left\langle \frac{1+\beta}{1-\beta}y^2 + \frac{1-\beta}{1+\beta}x^2 \right\rangle \\ & + \left\langle s\{v(1-\beta\cos 2\phi) - u\beta\sin 2\phi\}\right\rangle \\ & + \tfrac{1}{2}\left\langle u^2 + v^2 + w^2 + \beta(u^2 - v^2)\cos 2\phi - 2\beta uv\sin 2\phi\right\rangle, \end{aligned} \tag{4.3}$$

where the volume integral is defined as follows:

$$\left\langle \quad \right\rangle := \int_0^d \mathrm{d}z \iint_{\frac{x^2}{1+\beta}+\frac{y^2}{1-\beta}\leqslant 1} \mathrm{d}x\,\mathrm{d}y = \sqrt{1-\beta^2}\int_0^d\int_0^{2\pi}\int_0^1 s\,\mathrm{d}s\,\mathrm{d}\phi\,\mathrm{d}z. \tag{4.4}$$

The first term on the right-hand side of (4.3) is the basic-state energy, the second the interaction contribution and the third commonly referred to as the disturbance energy. The kinetic energy evolution equation is

$$\frac{\partial K}{\partial t} = E\oint_S \boldsymbol{U}\cdot(\hat{\boldsymbol{n}}\cdot\nabla)\boldsymbol{u}\,\mathrm{d}\boldsymbol{S} - E\left\langle|\nabla\times\boldsymbol{u}|^2\right\rangle, \tag{4.5}$$

where

$$\begin{aligned} \oint_S \boldsymbol{U}\cdot(\hat{\boldsymbol{n}}\cdot\nabla)\boldsymbol{u}\,\mathrm{d}\boldsymbol{S} = {} & \frac{1}{\sqrt{1-\beta^2}}\int_0^d\int_0^{2\pi}\Big\{\left(\beta\sin 2\phi - \tfrac{1}{2}\beta^2\sin 4\phi\right)\left[-\frac{\partial u}{\partial s} + \frac{1}{s}\frac{\partial v}{\partial\phi} + \frac{u}{s}\right] \\ & + \tfrac{1}{2}\beta^2(1-\cos 4\phi)\left[-\frac{1}{s}\frac{\partial u}{\partial\phi} + \frac{v}{s}\right] \\ & + \left[1 - 2\beta\cos 2\phi + \tfrac{1}{2}\beta^2(1+\cos 4\phi)\right]\frac{\partial v}{\partial s}\Big\}\,\mathrm{d}\phi\,\mathrm{d}z\Big|_{s=1} \end{aligned}$$

since the top and bottom surfaces are stress-free and

$$\left\langle|\nabla\times\boldsymbol{u}|^2\right\rangle = \sqrt{1-\beta^2}\int_0^d\int_0^{2\pi}\int_0^1(\omega_x^2+\omega_y^2+\omega_z^2)s\,\mathrm{d}s\,\mathrm{d}\phi\,\mathrm{d}z,$$

where the x, y and z components of the vorticity are

$$\begin{aligned} \omega_x &= \frac{1}{\sqrt{1-\beta}}\left(\sin\phi\frac{\partial w}{\partial s} + \cos\phi\frac{1}{s}\frac{\partial w}{\partial\phi}\right) - \sqrt{1-\beta}\left(\sin\phi\frac{\partial u}{\partial z} + \cos\phi\frac{\partial v}{\partial z}\right), \\ \omega_y &= \frac{1}{\sqrt{1+\beta}}\left(\sin\phi\frac{1}{s}\frac{\partial w}{\partial\phi} - \cos\phi\frac{\partial w}{\partial s}\right) - \sqrt{1+\beta}\left(\sin\phi\frac{\partial v}{\partial z} - \cos\phi\frac{\partial u}{\partial z}\right), \\ \omega_z &= \frac{1}{\sqrt{1-\beta^2}}\Big[-\beta\sin 2\phi\left(\frac{\partial u}{\partial s} - \frac{1}{s}\frac{\partial v}{\partial\phi} - \frac{u}{s}\right) + (1-\beta\cos 2\phi)\frac{\partial v}{\partial s} \\ & \quad + (1+\beta\cos 2\phi)\left(-\frac{1}{s}\frac{\partial u}{\partial\phi} + \frac{v}{s}\right)\Big]. \end{aligned}$$

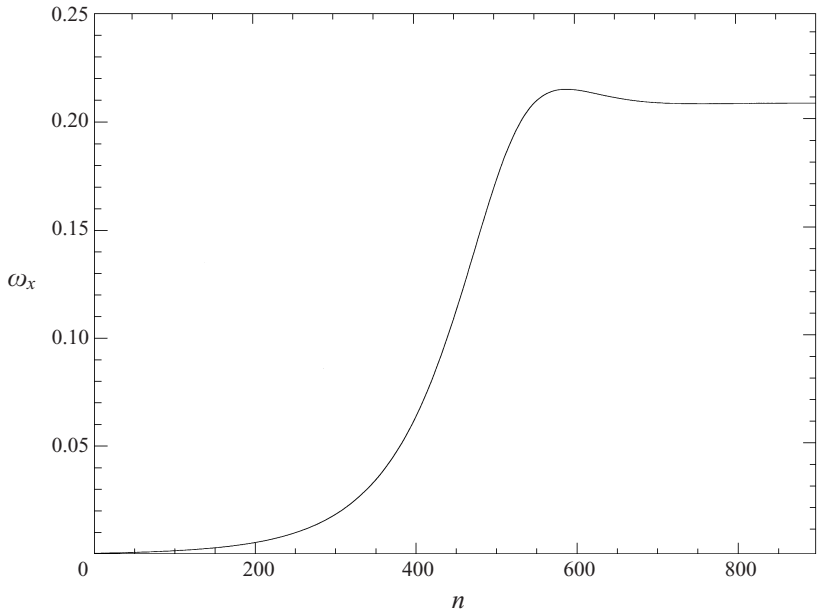

FIGURE 5. The x-component of the vorticity, ω_x (evaluated at $s = 0.5$, $\phi = 0$ and $z = d/4$) plotted against the number of revolutions n for the $m = (-1, 1)$ instability at $\beta = 0.033$, $d = 3.9796$ and $E = 2.5 \times 10^{-4}$ (the basic vorticity is $\boldsymbol{\omega} = 2\hat{\boldsymbol{z}}$). The initial condition is the unstable eigenfunction with amplitude 10^{-4}.

A comparison of the energies calculated via (4.3) and (4.5) after 10 revolutions typically displays an error of 1 part in 10^4 which is consistent with Marcus's findings.

4.4. *Saturation*

In both the $m = (-1, 1)$ and $m = (0, 2)$ instability cases, saturated states were eventually obtained by time-stepping the equations using quite strict modal truncations in m and l. The truncations $M = 4$ and $L = 4$ proved more than sufficient to capture the equilibrated form of the two instabilities. Figures 5 and 6 give typical time evolutions to saturation for the $m = (-1, 1)$ and $m = (0, 2)$ elliptical instabilities respectively. In each case, the x-component of the vorticity, ω_x, evaluated at $(s, \phi, z) = (\frac{1}{2}, 0, \frac{1}{4}d)$ is chosen to indicate the disturbance amplitude (the underlying vorticity is $\boldsymbol{\omega} = 2\hat{\boldsymbol{z}}$). In the $m = (-1, 1)$ case, ω_x always seems to show a small overshoot before settling down to a slightly lower saturated value whereas the amplitude envelope of ω_x evolves monotonically to a limit in the $m = (0, 2)$ case. Figure 7(a–c) shows the saturated velocity field reached in figure 5. The most striking feature of the total velocity field is the wavy rotation axis which is clearly seen in the laboratory (see figure 1 of Kerswell 1999).

The rate at which an elliptical instability can extract energy from the underlying flow depends strongly on its orientation or phase relative to the imposed ellipticity. As a result, the saturation process is quite intricate because the growing instability equilibrates through phase-shifting as well as the usual frequency-detuning and damping mechanisms. Furthermore, our 'simplifying' stress-free top and bottom boundary conditions actually complicate matters here because the mean flow is only generated at second order by the growing instability. In the experimental set-up where all boundary conditions are non-slip, nonlinear horizontal Ekman boundary layers

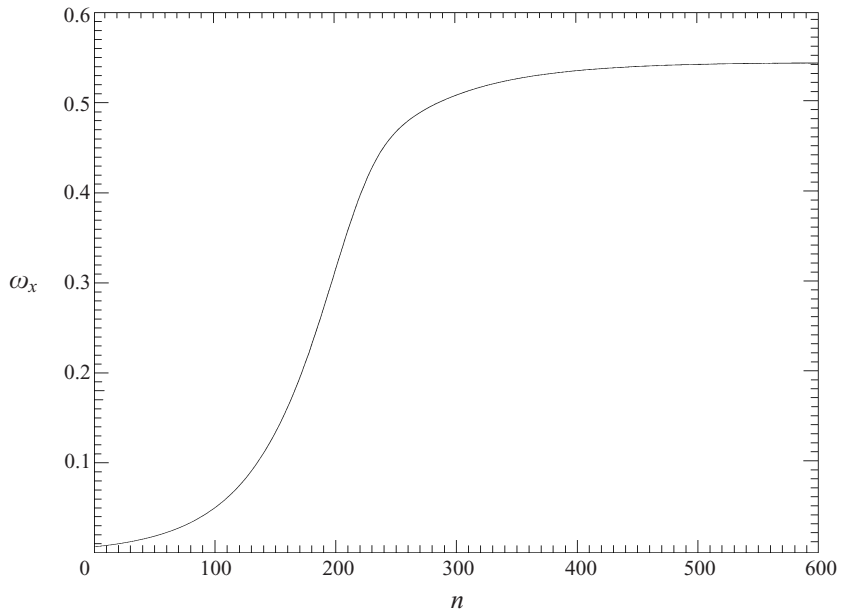

FIGURE 6. The amplitude envelope of the oscillating x-component of the vorticity, $\omega_x(0.5, 0, d/4)$, plotted against the number of revolutions n for the $m = (0, 2)$ instability at $\beta = 0.037$, $d = 2.7009$ and $E = 2.5 \times 10^{-4}$. The initial condition is the unstable eigenfunction with amplitude 10^{-3}.

ensure that frictional damping from an $O(\epsilon^2)$ mean flow can balance the appropriate leading nonlinear terms produced by an $O(\epsilon)$ primary flow. With these layers absent, only the internal friction from an $O(\epsilon^2/E^{1/2})$ mean flow is available for this task. For the elliptically distorted system studied here, in fact, no such large mean flow is required since these leading nonlinear terms are found to vanish identically in the lowest-order phase direction. This necessitates working to second order in the mean flow to capture the leading dynamic balance. A number of interesting features (a second-order linear Ekman boundary layer and a nonlinear Ekman boundary layer amongst other things) arise in the subsequent analysis which is somewhat involved and therefore relegated to the Appendix. Figure 8 compares the weakly nonlinear theoretical results with the actual numerical solutions for the $m = (-1, 1)$ case: the bifurcation is clearly a supercritical pitchfork. The correspondence is good considering that the analysis is based upon a double expansion in $E^{1/2}$ and the inertial wave amplitude ϵ with both quantities numerically small but not 'really' small. Saturated numerical solutions are also obtained in the $m = (0, 2)$ case for $\beta > \beta_{crit}$ so that this bifurcation is supercritical too.

4.5. *Secondary instabilities*

All the equilibrated states reached through numerical computation were essentially weakly nonlinear states in which the initially unstable inertial modes still dominated the velocity structure. As a result, it seemed plausible that the stability characteristics of these dominant inertial modes should largely dictate those of the weakly nonlinear solution. The $m = (-1, 1)$ saturation provides a particularly appealing situation in which to test this hypothesis since only the stability characteristics of the one inertial wave (and its complex conjugate) need be examined. A companion paper (Kerswell 1999) reports on the linear stability of just that inertial mode as a function of its

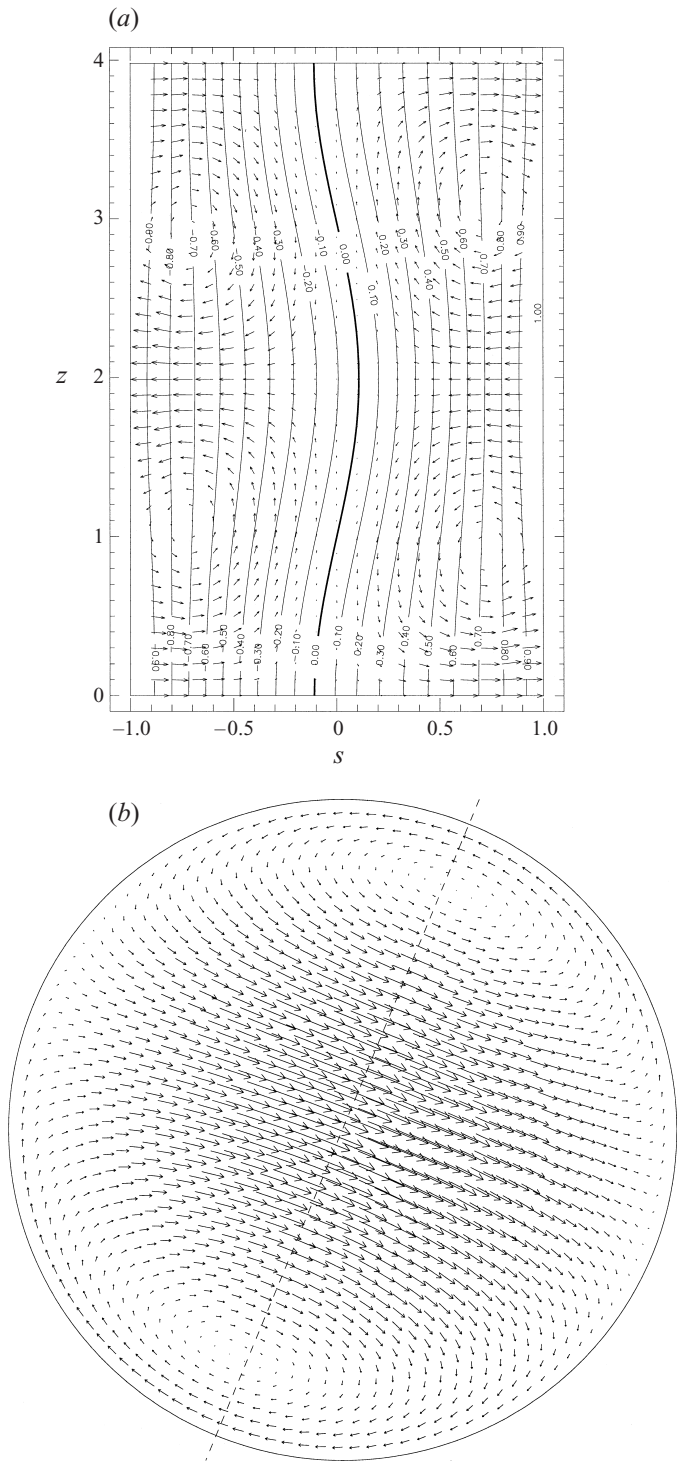

FIGURE 7 (*a*, *b*). For caption see facing page.

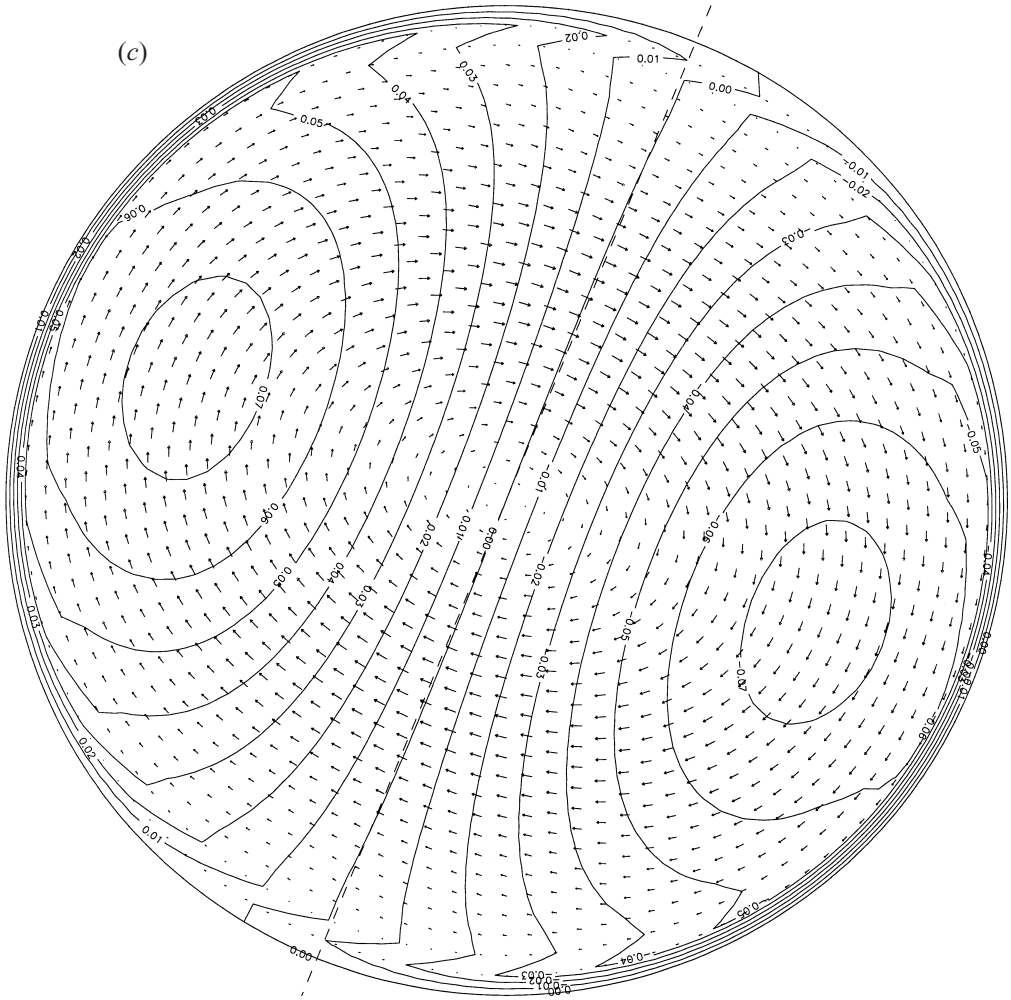

FIGURE 7. Contour and arrow plots of the steady velocity field reached as the saturated endstate of the $m = (-1, 1)$ instability in figure 5 ($\beta = 0.033$, $d = 3.9796$ and $E = 2.5 \times 10^{-4}$). (*a*) A z, s contour/arrow plot of the total velocity. The thick central wavy line represents the new rotation axis of the fluid (compare with figure 1 of Kerswell 1999). The ϕ-plane has been chosen to show the largest bend in this axis: here $\phi = 0.8762\pi$. The numbers represent contour levels on v the azimuthal speed which is just 1 at the sidewalls $s = 1$. The arrows indicate the meridional velocity $u\hat{\boldsymbol{s}} + w\hat{\boldsymbol{z}}$. The longest arrow corresponds to a speed of 0.001582. (*b*) An s, ϕ slice at the equator of the cylinder $z = d/2$ showing the disturbance velocity field (note that w vanishes at this z level). The size of the longest arrow corresponds to a speed of 0.109. The dashed line indicates the meridional slice taken in (*a*). (*c*) An s, ϕ slice at $z = d/4$ of the disturbance velocity field. Contours refer to axial velocity w and Ekman boundary layers are clearly visible. Again the dashed line is the meridional slice taken in (*a*). The longest arrow corresponds to a speed of 0.004015.

amplitude A (defined in equation (2.9) of Kerswell 1999 and equal to $\epsilon/3$ where ϵ is defined in the Appendix) and Ekman number E. According to this analysis at $E = 2.5 \times 10^{-4}$, the inertial mode should become linearly unstable at $A \approx 0.033$ to a pair of inertial modes, $(m, l) = (2, \frac{3}{2})$ and $(3, \frac{1}{2})$, which form a triad with the underlying state $(m, l) = (\pm 1, \pm 1)$ (see table 2, Kerswell 1999). The $m = (-1, 1)$ saturated state corresponding to $\beta = 0.034$ at $E = 2.5 \times 10^{-4}$ (see figure 5) has at its heart the inertial mode $(m, l) = (\pm 1, \pm 1)$ with amplitude $A \approx 0.04$. The unstable

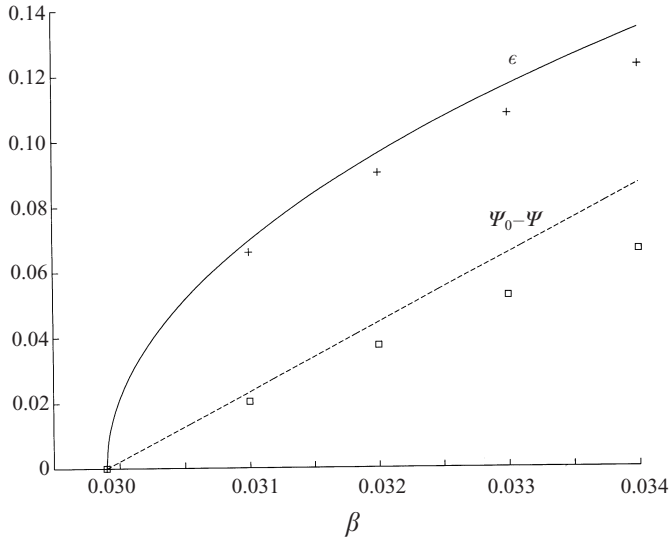

FIGURE 8. A plot of amplitude ϵ and phase $\psi_0 - \psi$ against elliptical distortion β for the saturated $m = (-1, 1)$ elliptical instability at $E = 2.5 \times 10^{-4}$ and $d = 3.9796$. The discrete data points are numerical results and the continuous lines weakly nonlinear theoretical predictions. The ϵ, β graph is standard indicating supercritical pitchfork bifurcation. The $(\psi_0 - \psi), \beta$ plot is less standard. Based upon the expansions, $\psi = \psi_0(E) + \epsilon^2/E^{1/2}\psi_{2-1}$ and $\beta = \beta_0 + \epsilon^2\beta_{20}$, a straight line $\psi_0 - \psi = -\psi_{2-1}/(E^{1/2}\beta_{20})(\beta - \beta_0)$ is predicted as $\epsilon \to 0$. (In the weakly nonlinear results $\beta_0(E) = 0.0299$ and $\psi_0(E) = -0.8595\pi$ have been taken.)

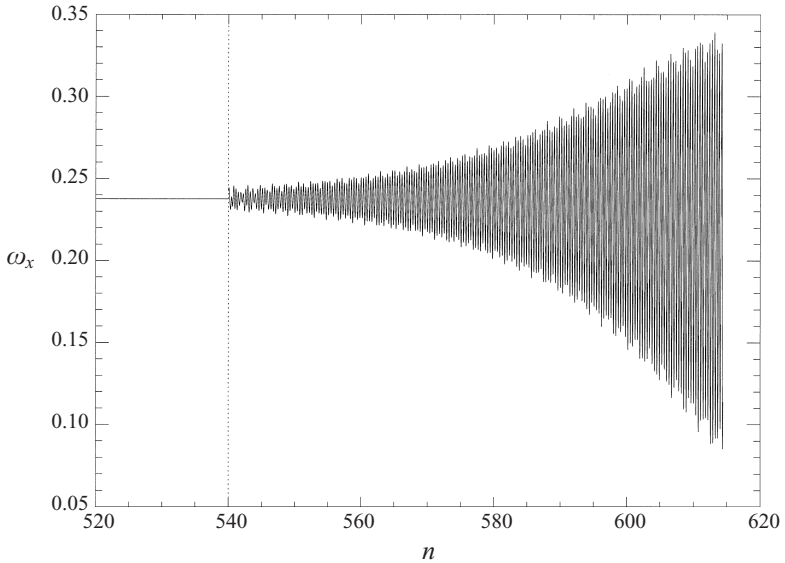

FIGURE 9. The effect of adding the unstable $(2, \frac{3}{2})$–$(3, \frac{1}{2})$ eigenfunction as a small perturbation (amplitude=$O(10^{-4})$) to the saturated $m = (-1, 1)$ state for $\beta = 0.034$ at $E = 2.5 \times 10^{-4}$ (time of addition is 540 revolutions). The code truncation is $(N, M, L) = (90, 4, 8)$ and $\Delta t = 0.05$. The data stop because the code eventually breaks down when the velocity amplitude gets too large.

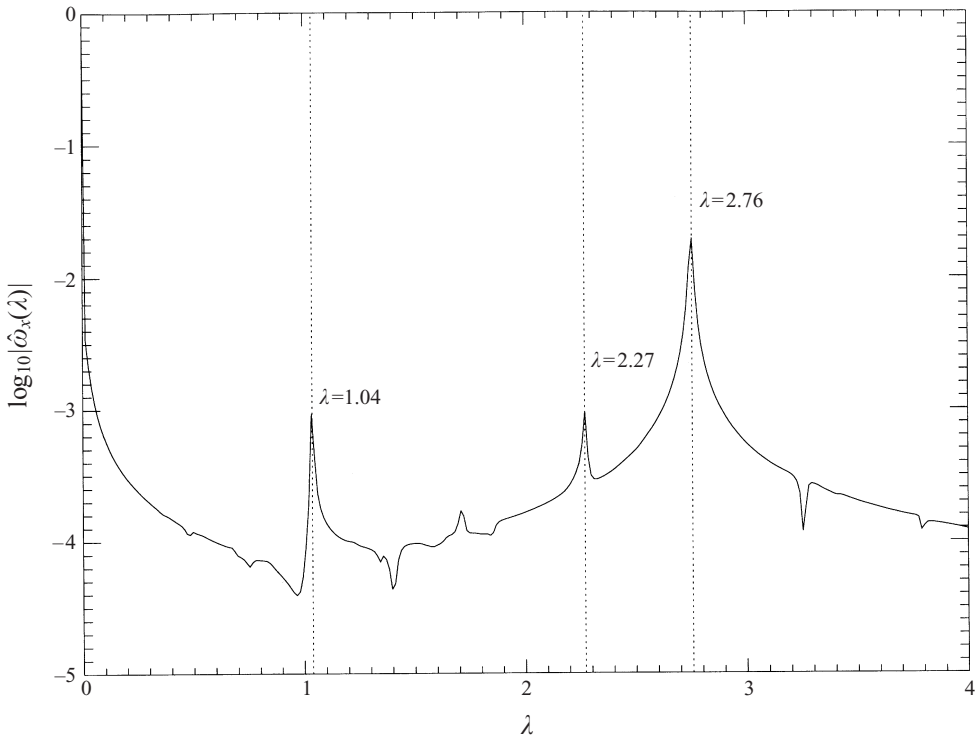

FIGURE 10. A power spectrum of the ω_x data in figure 9. The modulus of the coefficients $\hat{\omega}_x$ in the Fourier series of ω_x for $n \geqslant 540$ are plotted (on a $\log_{10}$ scale) against the corresponding frequency λ. There are three clear peaks: $\lambda = 2.76$ corresponding to the dominant secondary instability; $\lambda = 2.27$ corresponding to a subdominant secondary instability; and $\lambda = 1.04$ corresponding to a competing $m = (0, 2)$ elliptical instability. (The fact that ω_x is not stationary but growing secularly leads to the broadening of the frequency peaks.)

$(2, \frac{3}{2})$–$(3, \frac{1}{2})$ eigenfunction from the linear analysis at $A = 0.04$ was therefore added as a small perturbation (amplitude $= O(10^{-4})$) to the saturated state (the code truncation being necessarily expanded from $(M, L) = (4, 4)$ to $(4, 8)$ to accommodate the half-wavelength disturbance). Figure 9 confirms that the disturbance magnifies in time with a growth rate 2.9×10^{-3} remarkably close to the theoretical prediction of 3.6×10^{-3} based upon an isolated inertial wave basic state.

A power spectrum of the time series data for the growing disturbance in figure 10 indicates three clear peaks. The highest located at a frequency of 2.76 corresponds to the dominant 'triad' instability $(2, \frac{3}{2})$–$(3, \frac{1}{2})$. The lower peak at a frequency of 2.27 coincides with a second 'triad' instability consisting of the modes $(2, \frac{1}{2})$ and $(3, \frac{3}{2})$ which has just become unstable at $A = 0.04$ (frequency 2.28 and growth rate 1.0×10^{-3}; see also figures 5 and 6 of Kerswell (1999) where this instability appears with frequencies 2.32 and 2.35 at $E = 10^{-4}$ and 10^{-5} respectively). The third is at a frequency of 1.04 and since it consists of the two modes $(m, l) = (0, \frac{3}{2})$ and $(2, \frac{3}{2})$ growing together cannot be a triad with the underlying inertial mode. This frequency, in fact, corresponds to the $m = (0, 2)$ elliptical instability whose resonant geometry for $\frac{3}{2}$ axial wavelengths is $d = \frac{3}{2} \times 2.7009 = 4.0514$ and therefore relatively close to $d = 3.9796$. Although the first thought is that the finite-amplitude $m = (-1, 1)$ saturated state must 'tune in' the $m = (0, 2)$ instability, the reality is that the Ekman number is large enough

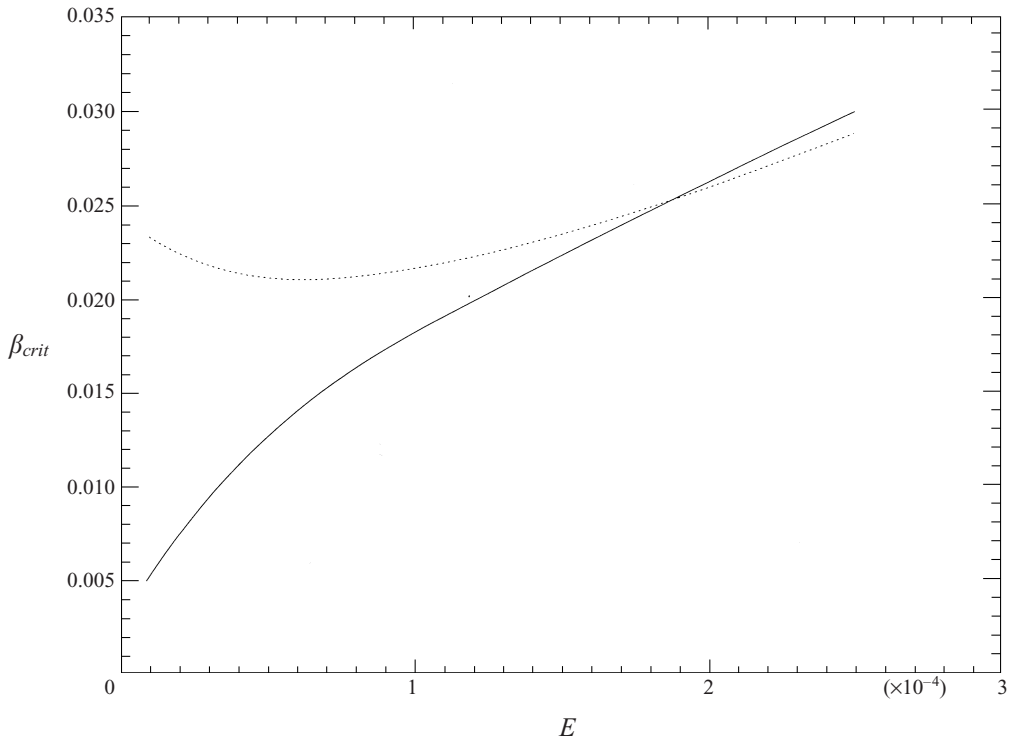

FIGURE 11. A plot of β_{crit} against E for the $m = (-1, 1)$ (solid line) and $m = (0, 2)$ (dotted line) instabilities *at* $d = 3.9796$. For $E \leqslant 1.9 \times 10^{-4}$ the $m = (-1, 1)$ instability appears before the $m = (0, 2)$ instability as it must do eventually in the inviscid limit $E \to 0$.

at $E = 2.5 \times 10^{-4}$ to smear out and shift the instability window around $d = 4.0514$ for the $m = (0, 2)$ elliptical instability to include $d = 3.9796$. In fact, a combination of viscous tuning for the $m = (0, 2)$ elliptical instability and viscous *detuning* for the $m = (-1, 1)$ instability means that the former actually has a marginally *larger* growth rate at 2.96×10^{-3} than the latter's 2.66×10^{-3}! Figure 11 shows that this is purely a result of the Ekman number being relatively large. As soon as the Ekman number is below about 1.9×10^{-4}, and experiments have typically $E = 5 \times 10^{-5}$, the $m = (-1, 1)$ instability reinstates itself as the first and presumably preferred instability.

At $\beta = 0.034$, the secondary instability is too strong to show any signs of equilibrating before the amplitude of the flow solution becomes too large and the code breaks down. However, an analogous run at $\beta = 0.033$ indicates that a new 'secondary' saturated state could be reached. Figure 12 shows the initial exponential growth of the secondary instability arresting at about $n \approx 1100$ before the $m = (0, 2)$ elliptical instability apparently takes over and eventually causes the code to blow up. A plot of the main modal amplitudes as a function of time in figure 13 confirms this interpretation. The seeded $(2, \frac{3}{2})$–$(3, \frac{1}{2})$ secondary instability has the stronger initial growth but appears to saturate at a low level (notice the corresponding decrease in the primary $(1, 1)$ modal strength). Eventually the $m = (0, 2)$ elliptical instability establishes itself once the $(0, \frac{3}{2})$ mode has grown to a sufficient amplitude and then the $(0, \frac{3}{2})$ and $(2, \frac{3}{2})$ modes grow together until the code breaks down. This strongly suggests that for conditions under which the $m = (0, 2)$ instability is *not* coincidentally excited (e.g. smaller E), secondary saturation of the $(2, \frac{3}{2})$–$(3, \frac{1}{2})$ triad instability would be achieved. This

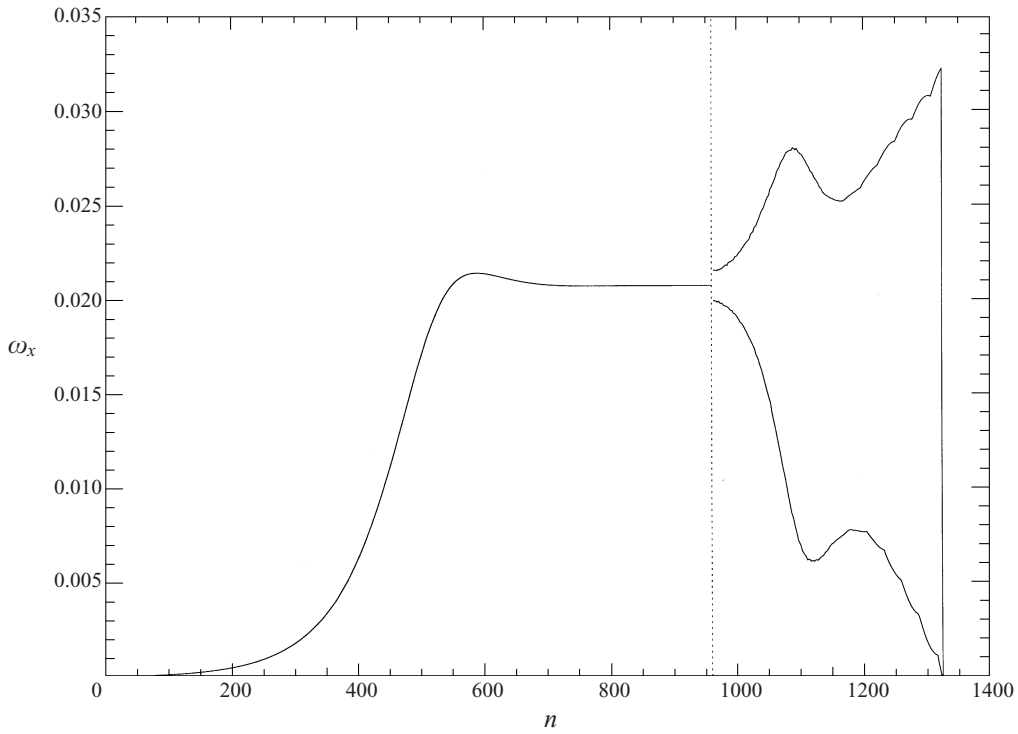

FIGURE 12. The effect of adding the unstable $(2, \frac{3}{2})$–$(3, \frac{1}{2})$ eigenfunction as a small perturbation to the saturated $m = (-1, 1)$ state for $\beta = 0.033$ at $E = 2.5 \times 10^{-4}$ (point of addition indicated by the vertical dotted line). The code truncation is $(N, M, L) = (90, 4, 8)$ and $\Delta t = 0.05$. Only the amplitude envelope of ω_x is shown to avoid solid black filling due to the rapid oscillations compared to the growth timescale. The data shown are again limited by the code breaking down.

secondary Hopf bifurcation therefore appears supercritical as well. Figure 14 shows that considerable nonlinear mixing has occurred by the time the code breaks down. The same primary frequencies are present as in figure 10 but now combinations have become established. Figure 15(*a–c*) gives some idea of how complicated the flow has become just before code breakdown in figure 12.

With the instability of the final $m = (-1, 1)$ saturated state established, the next logical step was to investigate how the $m = (-1, 1)$ instability reacts to the perturbation being added at an earlier stage of its evolution. A run was performed in which the $(2, \frac{3}{2})$–$(3, \frac{1}{2})$ eigenfunction (calculated at $A = 0.04$) was added to the growing $m = (-1, 1)$ instability when its amplitude was $A = 0.002$ ($\beta = 0.034$). A plot of $\omega_x(\frac{1}{2}, 0, \frac{1}{4}d)$ over time (figure 16) clearly shows the steadily growing $m = (-1, 1)$ state carrying a smaller oscillatory instability. The frequency of this signal is evidently different from the initial perturbation's 2.76 added at $n = 200$. A power spectrum of the time series shown in figure 17 confirms that the dominant signal is due to the $m = (0, 2)$ elliptical instability of frequency $\lambda = 1.04$. A modal amplitude verses time plot (figure 18) gives the complete dynamical picture of the evolution showing that the elliptical instability $m = (0, 2)$ actually grows right from the start. The $(3, \frac{1}{2})$ mode initially decays until the $(1, 1)$ mode has reached sufficient amplitude to sustain the $(2, \frac{3}{2})$–$(3, \frac{1}{2})$ triad instability at about $n = 270$. This critical amplitude is lower than would normally be the case because energy is already being fed into the $(2, \frac{3}{2})$ mode

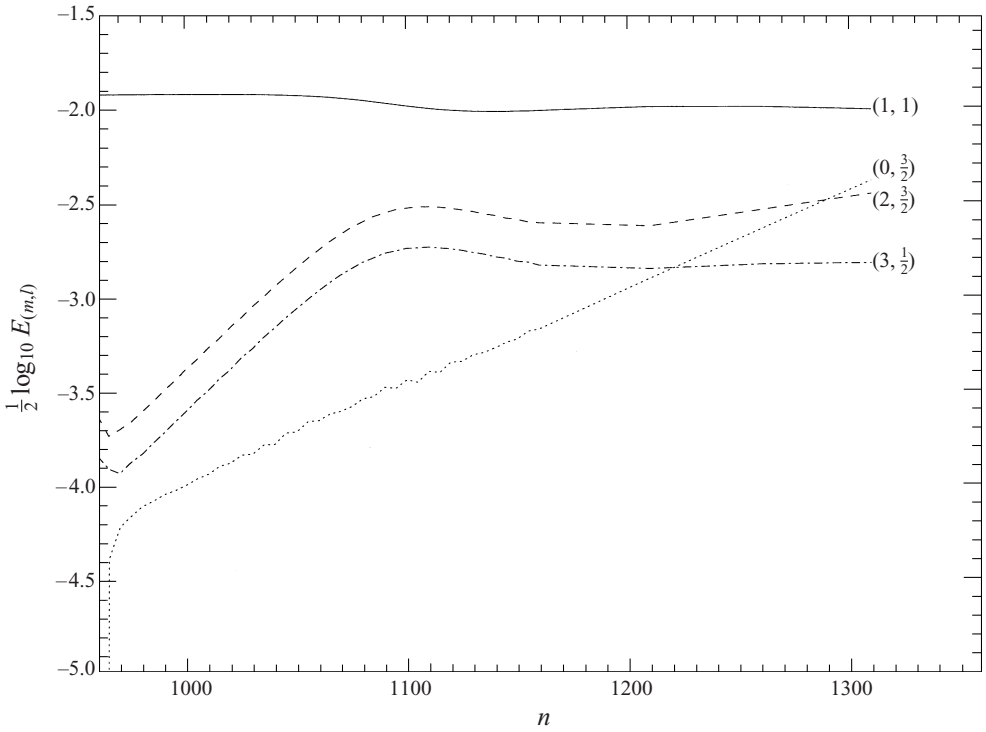

FIGURE 13. A plot of the main modal amplitudes against time for the run shown in figure 12 ($\beta = 0.033$, $E = 2.5 \times 10^{-4}$). The labels are in the form (m, l) so, for example, $(1, 1)$ represents one half of the primary $m = (-1, 1)$ elliptical instability ($(-1, -1)$ is the other). The modes $(2, \frac{3}{2})$ and $(3, \frac{1}{2})$ clearly grow together initially and appear to saturate. Ultimately, however, the elliptical instability modes $(0, \frac{3}{2})$ and $(2, \frac{3}{2})$ disrupt this secondary saturation.

through the elliptical mechanism. This aside, it is clear that at lower Ekman numbers where the $m = (0, 2)$ elliptical instability would not be excited concurrently, the secondary instability would manifest itself as predicted by quasi-static linear analysis of the underlying $(1, 1)$ inertial mode (Kerswell 1999).

Secondary instability in the $m = (0, 2)$ case unfortunately could not be found with $E \geqslant 2.5 \times 10^{-4}$ because the amplitude of the saturated states captured numerically were presumably too small. This was a direct result of the increased numerical difficulty in accurately tracking an oscillatory instability. The smaller time step ($\Delta t = 0.025$) and larger radial truncation ($N = 120$) required invariably caused the code to blow up at smaller disturbance amplitudes than in the $m = (-1, 1)$ case.

5. Discussion

A number of new results have been established in this paper. First, Waleffe's (1989) linear stability analysis for flow within a rotating elliptically distorted cylinder valid in the joint limit of vanishing elliptical distortion and Ekman number ($\beta, E \to 0$) has been confirmed and extended to finite β and E through a two-dimensional eigenvalue code. Secondly, the nonlinear evolutions of two elliptical instabilities, the subharmonic case $m = (-1, 1)$ and a more generic example $m = (0, 2)$, have been numerically simulated using a new Navier–Stokes solver in a non-orthogonal elliptico-polar

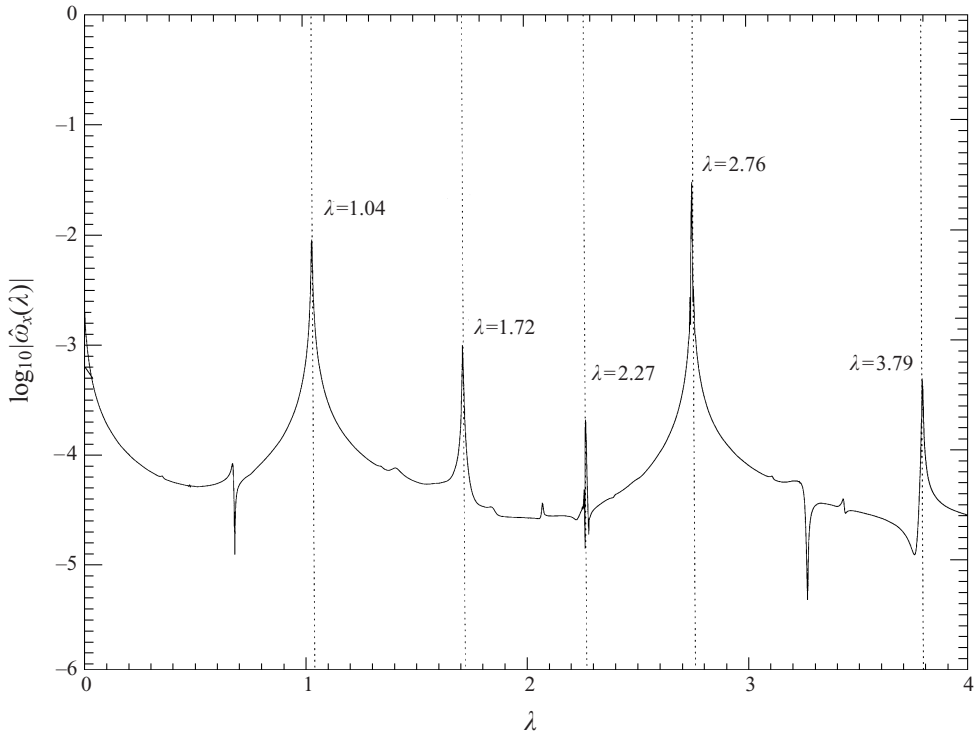

FIGURE 14. A power spectrum for the run shown in figure 12 based upon ω_x data beyond the dotted vertical line. The flow appears more complicated than in figure 10 but this is because nonlinear mixing is more evident. The base frequencies are, however, the same.

coordinate system. Using appropriately restricted truncations, (primary) saturated endstates have been found in both cases consistent with supercritical bifurcation. A rather delicate weakly nonlinear analysis for the $m = (-1, 1)$ instability (with stress-free horizontal boundary conditions) has been developed and shown to capture the main features of the numerical solutions obtained. This analysis has highlighted the existence of an intricate saturation mechanism where a phase-shifting process operates alongside the more usual frequency-detuning and damping processes. Thirdly, the stability of both saturated states has been studied by examining the temporal evolution of flow perturbations. The $m = (-1, 1)$ saturated state has been shown to be linearly unstable at amplitudes consistent with a quasi-static stability analysis of the dominant inertial mode present (Kerswell 1999). Strong evidence has been presented (figure 13) to indicate that this secondary Hopf bifurcation is also supercritical since the secondary instability appears to saturate at a low level leading to a new possibly quasi-periodic state. However the presence of another elliptical instability (coincidentally the $m = (0, 2)$ instability), also excited at $d = 3.9796$ due to the comparatively large value of $E = 2.5 \times 10^{-4}$, ultimately disrupts this (secondary) state leading to quite complicated nonlinear behaviour. For the restricted parameter settings which could be handled by the numerical code, the $m = (0, 2)$ saturated state appeared to be stable.

The numerical model considered here differs in one important aspect from the actual experimental set up: the horizontal boundary conditions. Previous numerical work (Kerswell & Barenghi 1995) directed at predicting when the (primary) elliptical

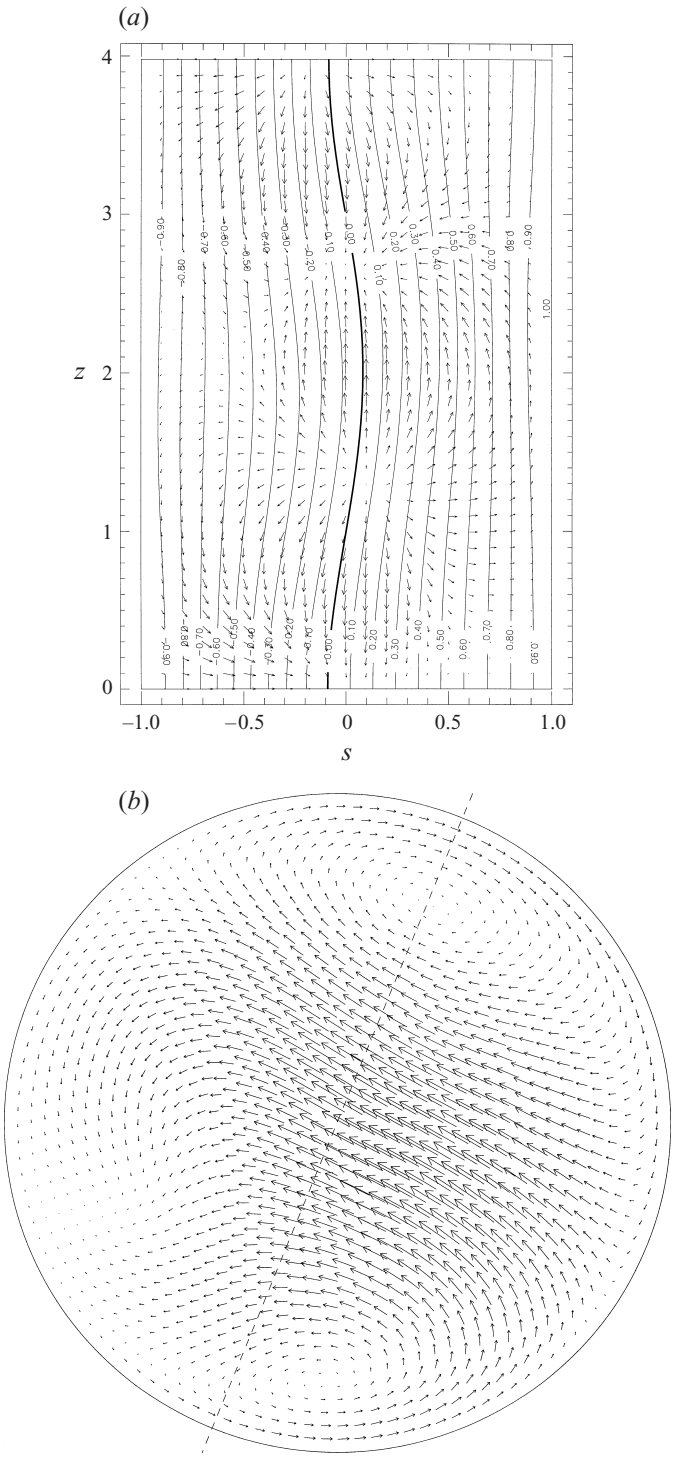

FIGURE 15 (a, b). For caption see facing page.

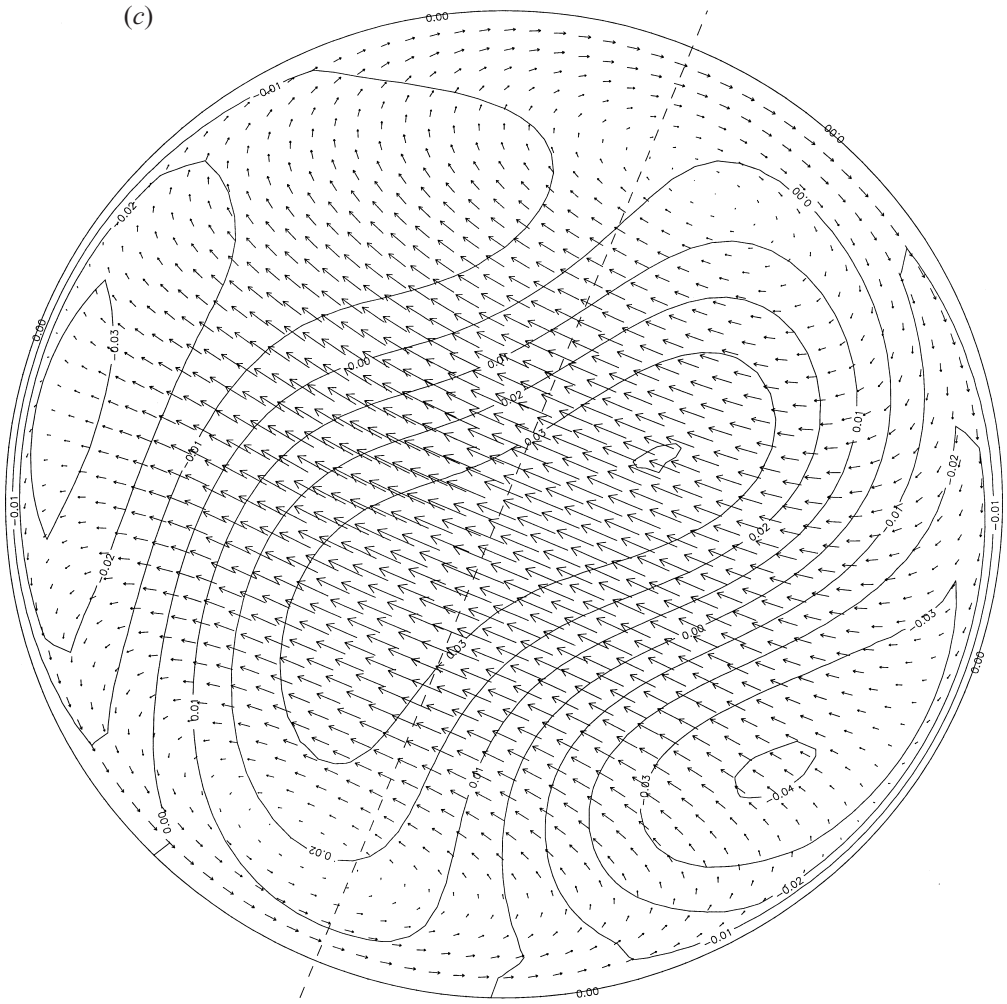

FIGURE 15. The velocity field just before the code breaks down in figure 12 at $n = 1325$. (*a*) The same meridional slice as in figure 7(*a*). The contours indicate azimuthmal velocity and arrows meridional velocity of the total flow (the longest arrow corresponds to a speed 0.0362). (*b*) A slice across the top ($z = d$) of the cylinder showing the horizontal velocities associated with the disturbance (the longest arrow corresponds to a speed of 0.1120). (*c*) A slice across the centre ($z = d/2$) of the cylinder showing the horizontal velocities (arrows) and axial velocity (contours) associated with the disturbance (the longest arrow corresponds to a speed of 0.0847).

instability should emerge in Malkus's (1989) experiments has produced good quantitative agreement with observations (Malkus, private communication). However, the computational overhead incurred by working with realistic non-slip boundary conditions on *all* surfaces at present prevents any further progress beyond this modest starting point. The use of stress-free boundary conditions on the top and bottom cylinder surfaces is absolutely crucial in making the numerics attempted here feasible. This simplification is, in fact, not as dramatic as it first appears. The viscous damping experienced by inertial waves is certainly reduced but only by a fractional amount. The viscous decay rates remain asymptotically $O(E^{1/2})$ due to the sidewall Ekman layers: what is missing is the contribution from the horizontal Ekman layers which

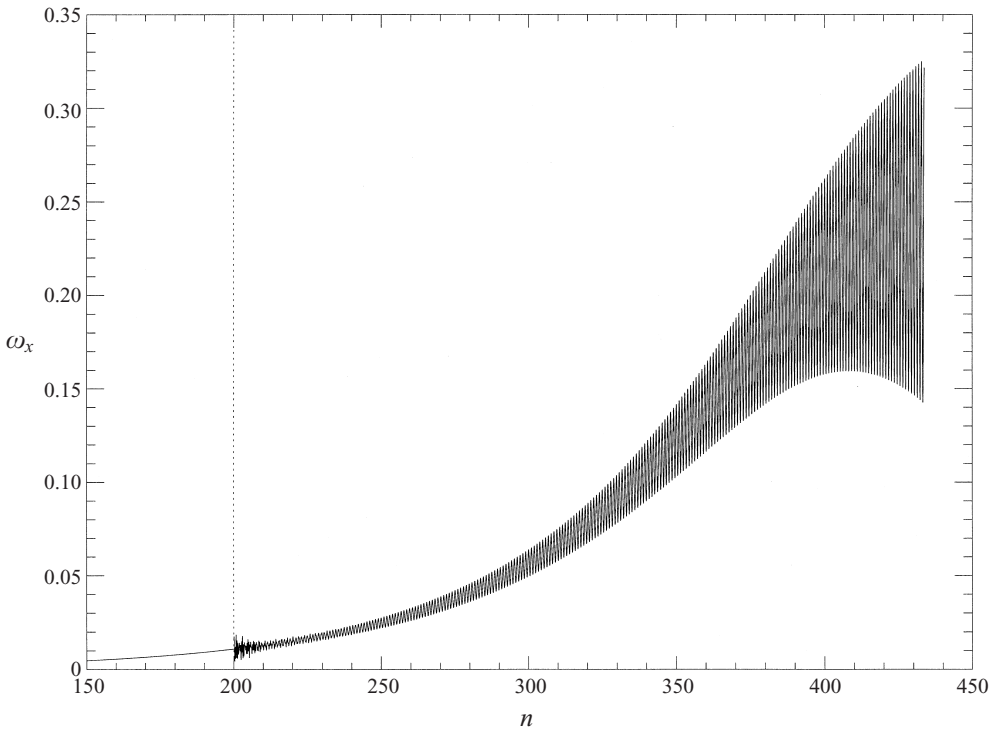

FIGURE 16. A plot of $\omega_x(0.5, 0, d/4)$ against the number of revolutions for $\beta = 0.034$, $E = 2.5 \times 10^{-4}$ and $d = 3.9796$, $(N, M, L) = (90, 4, 8)$ and $\Delta t = 0.05$. The $(2, \frac{3}{2})$–$(3, \frac{1}{2})$ eigenfunction is added as a small perturbation after 200 revolutions (indicated by the vertical dotted line) but is quickly overpowered by the growing $m = (0, 2)$ elliptical instability.

is of the same order. The dynamics of the mean flow, however, appear at first fundamentally different since its horizontal Ekman boundary layers are the dominating influence. What the weakly nonlinear analysis shows is that the loss of these layers is also accompanied by a cancellation of the leading-order nonlinear driving terms for the mean flow. Ultimately, this means that the mean flow is driven at the same order as the non-slip case but through a completely different process. The net result of all this is that we can reasonably expect only quantitative differences between the numerical predictions obtained here and laboratory observations: qualitatively, the two should be equivalent. This is especially true for the secondary instability findings which result from an inviscid triad mechanism: viscous boundary layers act only to determine *when* the instability occurs not *if*.

The main significance of this work is to provide concrete support for the general hypothesis that inertial waves are generically unstable (Kerswell 1999, see also Lifschitz & Fabijonas 1996 and Fabijonas, Holm & Lifschitz 1997 who consider the unbounded Kelvin wave analogues). This has obvious implications for all perturbed rapidly rotating flows in which inertial waves appear either through bifurcation or direct forcing. Here we have focused upon elliptical distortion as our perturbation (see also Aldridge *et al.* 1997; Kerswell & Malkus 1998) but many other physically interesting examples exist such as precession (Malkus 1968; Kerswell 1993; Vanyo *et al.* 1995; Kobine 1995, 1996) and rotation speed modulation (Aldridge & Toomre 1969; Tilgner 1999).

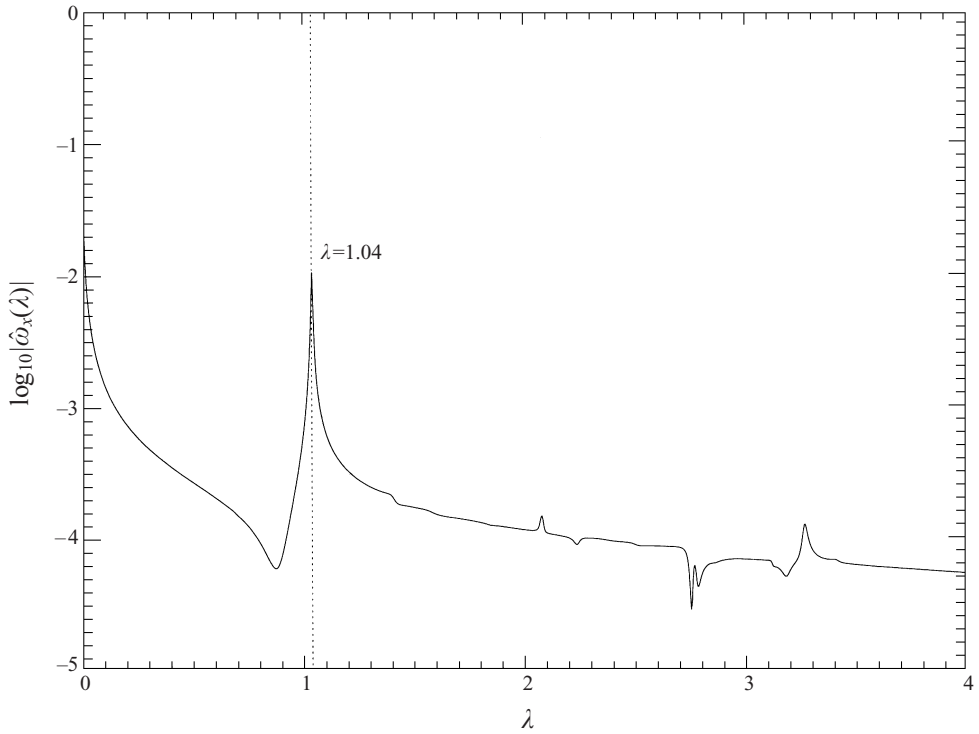

FIGURE 17. A power spectrum of the ω_x time series data shown in figure 16. The modulus of the coefficients $\hat{\omega}_x$ in the Fourier series of ω_x for $n \geqslant 200$ are plotted (on a $\log_{10}$ scale) against the corresponding frequency λ. This plot clearly indicates the dominance of the $m = (0, 2)$ elliptical instability. (The fact that the time series for ω_x is not stationary but evolving secularly leads to broadening of the frequency peaks.)

Beyond this secondary instability, experimental observations seem to show a rapid breakdown to small-scale disorder. This is consistent with the Ruelle–Takens scenario (Ruelle & Takens 1971) in which, as successive bifurcations add new frequencies to the dynamics, the resultant degree-n quasi-periodic motion gives way to a strange attractor at some $n \geqslant 3$ (Newhouse, Ruelle & Takens 1978). For the current dynamical system as well as other inertial wave systems, this transition appears to occur earlier, maybe even at the tertiary bifurcation, rather than later. Establishing this numerically is at present beyond reach but *is* feasible experimentally through time series data analysis (e.g. Gollub & Swinney 1975; Gollub & Benson 1980, and see figure 14 of Kobine 1995). Hopefully, this prospect will help stimulate new experimental studies specifically designed to extract such quantitative data.

The authors gratefully acknowledge the support of The Royal Society and EPSRC through an earmarked studentship for D.M.M. This work has also benefitted from the valuable experimental input of Professor W. V. R. Malkus.

Appendix. Weakly nonlinear analysis for the (−1,1) elliptical instability

The purpose of the weakly nonlinear analysis presented below is to identify an asymptotic relationship between the value of the applied elliptical distortion β, the saturated amplitude ϵ *and* the phase ψ of the primary excited inertial wave at a given

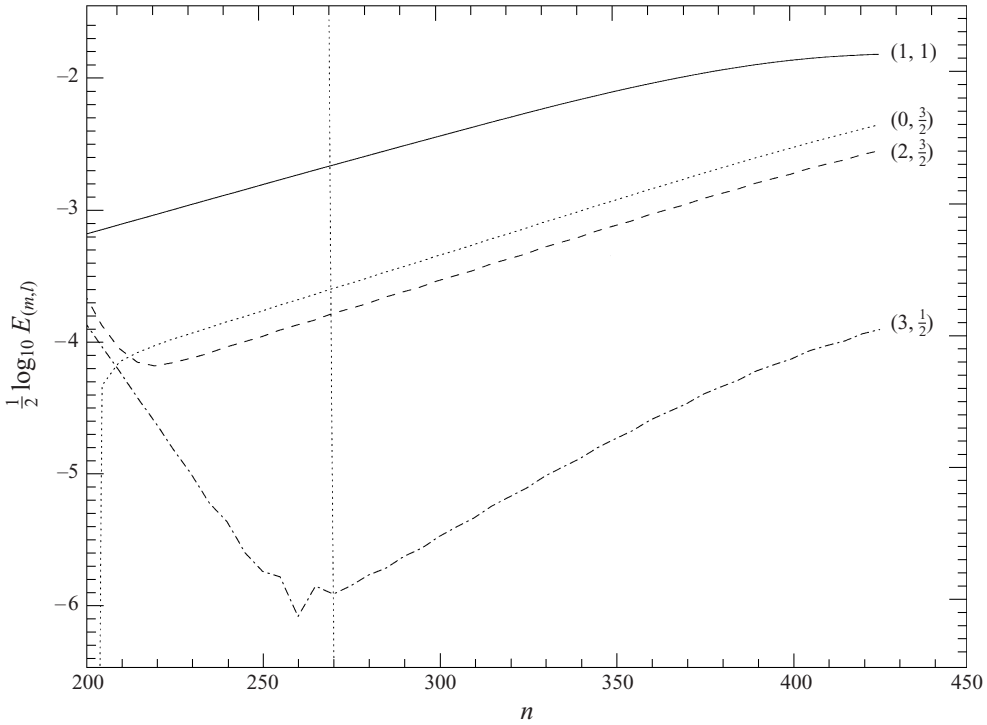

FIGURE 18. A modal amplitude plot against time for the evolution shown in figure 16. This clearly shows the $m = (0, 2)$ elliptical instability growing right from the start and the secondary instability initially decaying before finally growing once the underlying inertial wave hadreached a threshold amplitude.

Ekman number E. In doing so, we will find that no mean flow is generated at the leading order of $O(\epsilon^2/E^{1/2})$. We work in the rotating frame and develop asymptotic expansions for the velocity field $\boldsymbol{u}$, the elliptical distortion β and the phase ψ in terms of ϵ. If the velocity field of the primary inertial wave is represented as $\boldsymbol{Q}$, then by definition

$$\epsilon := \frac{|\langle \boldsymbol{Q}, \boldsymbol{u}\rangle|}{|\langle \boldsymbol{Q}, \boldsymbol{Q}\rangle|} := \left|\iiint \boldsymbol{Q}^* \cdot \boldsymbol{u}\, \mathrm{d}V\right| \left|\iiint |\boldsymbol{Q}|^2\, \mathrm{d}V\right|^{-1} \tag{A 1}$$

and we work with amplitude expansions

$$\left.\begin{aligned} \beta &= \beta_{crit}(E) + \epsilon^2\beta_2(E) + \cdots, \\ A &= A_0(E) + \epsilon^2 A_2(E) + \cdots = \mathrm{e}^{\mathrm{i}(t+\psi(E,\epsilon))} = \mathrm{e}^{\mathrm{i}(t+\psi_0(E)+\epsilon^2\psi_2(E)+\cdots)}, \end{aligned}\right\} \tag{A 2}$$

where A is the phase function for the primary inertial wave and the $O(\epsilon)$ terms vanish in both cases because $\langle \boldsymbol{Q}, \boldsymbol{Q}\cdot\nabla\boldsymbol{Q}\rangle = 0$. Each expansion term can be further expanded as a series in $E^{1/2}$. For example, we will see below that

$$\beta_{crit} = \beta_0(E)E^{1/2} = \beta_{01}E^{1/2} + \beta_{02}E + \cdots, \tag{A 3}$$

$$\psi_0 = \psi_{00} + \psi_{01}E^{1/2} + \cdots\ . \tag{A 4}$$

The expansions for β_2 and ψ_2 can technically start at $O(E^{-1/2})$ and $O(E^{-1})$ respectively

so that

$$\beta_2 = \beta_{2-1}E^{-1/2} + \beta_{20} + \cdots, \tag{A 5}$$

$$\psi_2 = \psi_{2-2}E^{-1} + \psi_{2-1}E^{-1/2} + \psi_{20} + \cdots \tag{A 6}$$

(β_{ln} is associated with $O(\epsilon^l E^{n/2})$). The velocity field is expanded in the two small parameters ϵ and $E^{1/2}$ in the form

$$\begin{aligned}\boldsymbol{u} = \epsilon\big[&A\boldsymbol{u}_{101} + E^{1/2}\left(A\boldsymbol{u}^v_{111} + \beta_0 \mathrm{e}^{2\mathrm{i}t}A^*\boldsymbol{u}^e_{111} + \beta_0\mathrm{e}^{2\mathrm{i}t}A\boldsymbol{u}^e_{113}\right)\\ &+E\left(A\boldsymbol{u}^{vv}_{121} + \beta_0\mathrm{e}^{2\mathrm{i}t}A^*\boldsymbol{u}^{ev}_{121} + \beta_0^2 A\boldsymbol{u}^{ee}_{121} + \cdots\right) + \cdots\big]\\ &+\epsilon^2\big[\left(\bar{v}_0(s)E^{-1/2} + \bar{v}_1(s) + \cdots\right)\hat{\boldsymbol{\phi}} + |A|^2\boldsymbol{u}_{200} + A^2\boldsymbol{u}_{202} + \cdots + E\boldsymbol{u}_{220} + \cdots\big]\\ &+\epsilon^3\big[\boldsymbol{u}_{3-11}E^{-1/2} + \boldsymbol{u}_{301} + \cdots\big]\cdots + \text{c.c.}\end{aligned} \tag{A 7}$$

We will establish that the leading mean flow $\bar{v}_0(s)\hat{\boldsymbol{\phi}}$ vanishes so that, in fact, $\beta_{2-1} = \psi_{2-2} = 0$. The notation used throughout is that the velocity field $\boldsymbol{u}_{lnm}$ occurs at $O(\epsilon^l E^{n/2})$ and has azimuthal wavenumber m, so in particular $\boldsymbol{u}_{101} = \boldsymbol{Q}$. The superscripts, e and v, are used to help distinguish between flows generated by ellipticity or viscous effects. For example, the flow fields $\boldsymbol{u}^v_{111}$ and $\boldsymbol{u}^{vv}_{121}$ represent first- and second-order viscous corrections to the primary inertial wave in a cylinder. The flows $\boldsymbol{u}^e_{111}$ and $\boldsymbol{u}^e_{113}$ represent the first-order elliptical corrections, and $\boldsymbol{u}^{ee}_{121}$ the relevant part of the second. The flow $\boldsymbol{u}^{ev}_{121}$ is generated both as the leading viscous correction of $\boldsymbol{u}^e_{111}$ and as the leading elliptical correction to $\boldsymbol{u}^v_{111}$. There is no *a priori* assumption of relative scaling between ϵ and $E^{1/2}$. The physical processes associated with these parameters are very different and are left to organise the expansion. At the order to which we work, it is sufficient to consider the equation

$$\frac{\partial \boldsymbol{u}}{\partial t} + 2\hat{\boldsymbol{k}}\times\boldsymbol{u} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} + \nabla p = \tfrac{1}{2}\beta\left[\mathrm{e}^{2\mathrm{i}(\phi+t)}\mathscr{N}\nabla p + \mathrm{e}^{-2\mathrm{i}(\phi+t)}\mathscr{N}^*\nabla p\right] + E\nabla^2\boldsymbol{u} \tag{A 8}$$

(where $\mathscr{N}$ is defined in (4.2)) and all operators can be taken as the usual ones in cylindrical coordinates. In other words, we need only retain the 'ellipticity' in the pressure term and neglect it in the viscous diffusion term. (In deriving (A 8) from (2.18)–(2.20), we have also rescaled as follows:

$$[w, z, p] \to [\sqrt{1-\beta^2}w, \sqrt{1-\beta^2}z, (1-\beta^2)p],$$

although since $\beta^2 = O(E)$ this is not important in what follows.)† The weakly nonlinear analysis presented below is centred upon the determination of four solvability conditions. The threshold elliptical distortion, β_{crit}, is obtained via the solvability condition at $O(\epsilon E^{1/2})$. The $O(\epsilon^2/E^{1/2})$ and $O(\epsilon^2)$ mean flows, $\bar{v}_0\hat{\boldsymbol{\phi}}$ and $\bar{v}_1\hat{\boldsymbol{\phi}}$, emerge through solvability conditions at $O(\epsilon^2E^{1/2})$ and $O(\epsilon^2E)$ (the former vanishing identically) and the key saturation parameters β_2 and ψ_2 are obtained by applying a solvability condition at $O(\epsilon^3)$.

A.1. $O(\epsilon)$: *inertial wave problem*

The leading-order problem is merely the Poincaré eigenproblem for the primary inertial wave. In the regime of interest (smallest β_0), the inertial wave frequency, λ,

† Technically, the one place that this rescaling does raise its head is in producing an extra term $-\nabla p_{101}$ on the right-hand side of equation (A 21). However this is harmlessly absorbed into ∇p^{ee}_{121} and can be suppressed.

must be tuned to be within $O(E^{1/2})$ of 1 for linear instability to occur. Thus taking $\lambda = 1 + O(E^{1/2})$, the $O(\epsilon)$ problem is

$$\mathrm{i}\boldsymbol{u}_{101} + 2\hat{\boldsymbol{k}}\times\boldsymbol{u}_{101} + \nabla p_{101} = \boldsymbol{0} \tag{A 9}$$

with $\nabla\cdot\boldsymbol{u}_{101} = 0$ and $\boldsymbol{u}_{101}\cdot\hat{\boldsymbol{n}} = 0$ on $s = 1$. The inertial wave solution is

$$\boldsymbol{u}_{101} = \begin{bmatrix} \mathrm{i}u_0(s)\cos\alpha z \\ v_0(s)\cos\alpha z \\ \mathrm{i}w_0(s)\sin\alpha z \end{bmatrix}\mathrm{e}^{\mathrm{i}\phi} := \frac{\mathrm{e}^{\mathrm{i}\phi}}{2(4-\lambda^2)}\begin{bmatrix} \mathrm{i}\{(2+\lambda)J_0(ks) + (2-\lambda)J_2(ks)\}\cos\alpha z \\ -\{(2+\lambda)J_0(ks) - (2-\lambda)J_2(ks)\}\cos\alpha z \\ \frac{2\mathrm{i}\lambda k}{\alpha}J_1(ks)\sin\alpha z \end{bmatrix}, \tag{A 10}$$

$$p_{101} = -\frac{J_1(ks)}{k}\cos\alpha z\,\mathrm{e}^{\mathrm{i}\phi} \tag{A 11}$$

with the dispersion relation

$$\frac{k^2}{\alpha^2} = \frac{4-\lambda^2}{\lambda^2}$$

and eigencondition

$$(2+\lambda)J_0(k) + (2-\lambda)J_2(k) = 0. \tag{A 12}$$

The cylinder height-to-radius ratio $d = 3.9796$ gives a solution of (A 12) with $\lambda = 1$, radial wavenumber $k = 2.7346$ and $\alpha = 2\pi/d$ so that a complete wavelength is contained with the cylinder.

A.2. $O(\epsilon E^{1/2})$: *leading elliptical and viscous corrections*

Both the ellipticity and the viscosity contribute to the $O(\epsilon E^{1/2})$ balance. The $\mathrm{e}^{\mathrm{i}\phi}$ component is

$$\frac{\partial}{\partial t}\left[\beta_0 A_0^*\mathrm{e}^{2\mathrm{i}t}\boldsymbol{u}^e_{111} + A_0\boldsymbol{u}^v_{111}\right] + 2\hat{\boldsymbol{k}}\times\left[\beta_0 A_0^*\mathrm{e}^{2\mathrm{i}t}\boldsymbol{u}^e_{111} + A_0\boldsymbol{u}^v_{111}\right] + \nabla\left[\beta_0 A_0^*\mathrm{e}^{2\mathrm{i}t}p^e_{111} + A_0 p^v_{111}\right]$$
$$= \tfrac{1}{2}\beta_0 A_0^*\mathrm{e}^{2\mathrm{i}(\phi+t)}\mathscr{N}\nabla p^*_{101} - \frac{\mathrm{i}(1-\lambda)}{E^{1/2}}A_0\boldsymbol{u}_{101} \tag{A 13}$$

with $\nabla\cdot\boldsymbol{u}^e_{111} = \nabla\cdot\boldsymbol{u}^v_{111} = 0$ and boundary conditions on $s=1$, $\boldsymbol{u}^e_{111}\cdot\hat{\boldsymbol{n}} = 0$ and $\boldsymbol{u}^v_{111}\cdot\hat{\boldsymbol{n}} = F_1(\boldsymbol{u}_{101})$ where $F_1(\boldsymbol{u}_{101})$ is the (linear) Ekman pumping due to $\boldsymbol{u}_{101}$ (Greenspan 1968). At this point, it is simpler to treat β_0 (and ψ_0) as a whole rather than splitting it into its β_{01} and β_{02} parts and solving the same problem twice. The $\mathrm{e}^{3\mathrm{i}\phi}$ component is

$$3\mathrm{i}\boldsymbol{u}^e_{113} + 2\hat{\boldsymbol{k}}\times\boldsymbol{u}^e_{113} + \nabla p^e_{113} = \tfrac{1}{2}\mathrm{e}^{2\mathrm{i}\phi}\mathscr{N}\nabla p_{101} \tag{A 14}$$

with $\nabla\cdot\boldsymbol{u}^e_{113} = 0$ and $\boldsymbol{u}^e_{113}\cdot\hat{\boldsymbol{n}} = 0$ on $s = 1$. Constructing $\langle\boldsymbol{u}_{101}, (\mathrm{A}\,13)\rangle$ leads to a complex solvability condition,

$$\mathrm{i}(\lambda-1) + E^{1/2}s_1 = -\mathrm{i}\beta_0 E^{1/2}C_1\mathrm{e}^{-2\mathrm{i}\psi_0} \tag{A 15}$$

which correctly applied only at leading order gives

$$\beta_0 E^{1/2}C_1 = \beta_{01}E^{1/2}C_1 = \sqrt{\Delta^2 + s_1^{R^2}E}, \tag{A 16}$$

and

$$\tan 2\psi_{00} = \frac{E^{1/2}s_1^R}{\Delta} \quad \text{with} \quad \sin 2\psi_{00} = -s_1^R/(\beta_{01}C_1) > 0, \tag{A 17}$$

where $\Delta = \lambda + E^{1/2}s_1^I - 1$ is the frequency detuning ($\lambda - 1$ is the geometrical detuning and $E^{1/2}s_1^I$ the viscous detuning),

$$\mathrm{i}C_1 := \frac{\langle \boldsymbol{u}_{101}, \frac{1}{2}\mathrm{e}^{2\mathrm{i}\phi}\mathscr{N}\nabla p^*_{101}\rangle}{\langle \boldsymbol{u}_{101}, \boldsymbol{u}_{101}\rangle} = \left.\frac{9}{16 + 8/(1+k^2)}\right|_{\text{when } \lambda=1} \tag{A 18}$$

is the (first) elliptical coupling coefficient and

$$s_1 = s_1^R + \mathrm{i}s_1^I := \frac{-\oint p^*_{101}F(\boldsymbol{u}_{101})\,\mathrm{d}S}{\langle \boldsymbol{u}_{101}, \boldsymbol{u}_{101}\rangle} = \left.\frac{-3(1+\mathrm{i})(1+\alpha^2)}{2\sqrt{2}(1+2\alpha^2)}\right|_{\text{when } \lambda=1} \tag{A 19}$$

is the (first) complex viscous frequency shift (Kerswell & Barenghi 1995, equation (2.12)). The expression (A 16) gives the critical distortion in terms of the viscous decay rate s_1^R and detuning Δ in the asymptotic limit $E \to 0$. The interest here is more on the control of viscosity rather than detuning so we have assumed $\Delta = O(E^{1/2})$ which explains our initial $\beta_{crit} = O(E^{1/2})$ assumption. At perfect tuning, $\Delta = 0$, the phase angle is $\pi/4$ indicating that the inertial mode 'sits' in the direction of maximal vortex stretching (Waleffe 1990). The critical distortion is correspondingly minimized at $\beta_0 C_1 = |s_1^R|$ which underlines the fundamental balance achieved between ellipticity and viscosity. (It is important when calculating the velocity fields $\boldsymbol{u}^v_{111}$ and $\boldsymbol{u}^e_{111}$ to renormalize them so that $\langle \boldsymbol{u}_{101}, \boldsymbol{u}^e_{111}\rangle = \langle \boldsymbol{u}_{101}, \boldsymbol{u}^v_{111}\rangle = 0$. This enforces the definition of ϵ.)

A.3. *$O(\epsilon E)$: secondary elliptical and viscous corrections*

At this order, we must consider the second-order viscous correction $\boldsymbol{u}^{vv}_{121}$, the second-order elliptical correction $\boldsymbol{u}^{ee}_{121}$, and the hybrid second-order flow $\boldsymbol{u}^{ev}_{121}$. The respective problems to be solved are as follows. The problem for $\boldsymbol{u}^{vv}_{121}$ is

$$\mathrm{i}\boldsymbol{u}^{vv}_{121} + 2\hat{\boldsymbol{k}}\times\boldsymbol{u}^{vv}_{121} + \nabla p^{vv}_{121} = \nabla^2\boldsymbol{u}_{101} - s_2^v\boldsymbol{u}_{101} \tag{A 20}$$

with $\nabla\cdot\boldsymbol{u}^{vv}_{121} = 0$ and $\boldsymbol{u}^{vv}_{121}\cdot\hat{\boldsymbol{n}} = F_2(\boldsymbol{u}_{101}, \boldsymbol{u}^v_{111})$ on $s = 1$ where F_2 is the *second*-order (linear) Ekman pumping. The coefficient s_2^v represents the second-order viscous frequency shift present in the absence of ellipticity. The problem for $\boldsymbol{u}^{ee}_{121}$ is

$$\mathrm{i}\boldsymbol{u}^{ee}_{121} + 2\hat{\boldsymbol{k}}\times\boldsymbol{u}^{ee}_{121} + \nabla p^{ee}_{121} = \tfrac{1}{2}\left[\mathrm{e}^{2\mathrm{i}\phi}\mathscr{N}\nabla p^{e*}_{111} + \mathrm{e}^{-2\mathrm{i}\phi}\mathscr{N}^*\nabla p^e_{113}\right] - \mathrm{i}C_2\boldsymbol{u}_{101} \tag{A 21}$$

with $\nabla\cdot\boldsymbol{u}^{ee}_{121} = 0$ and $\boldsymbol{u}^{ee}_{121}\cdot\hat{\boldsymbol{n}} = 0$ on $s = 1$. The coefficient $\beta_0^2 C_2$ is the second-order elliptical coupling coefficient. The problem for $\boldsymbol{u}^{ev}_{121}$ is

$$\mathrm{i}\boldsymbol{u}^{ev}_{121} + 2\hat{\boldsymbol{k}}\times\boldsymbol{u}^{ev}_{121} + \nabla p^{ev}_{121} = \tfrac{1}{2}\mathrm{e}^{2\mathrm{i}\phi}\mathscr{N}\nabla p^{v*}_{111} - s_2^e\boldsymbol{u}_{101} \tag{A 22}$$

with $\nabla\cdot\boldsymbol{u}^{ev}_{121} = 0$ and $\boldsymbol{u}^{ev}_{121} = F_1(\boldsymbol{u}^e_{111})$ on $s = 1$. Here $\beta_0 s_2^e$ is the second-order viscous frequency shift to $\boldsymbol{u}_{101}$ due to the first-order viscous correction to the elliptical flow $\boldsymbol{u}^e_{111}$ and the elliptical correction to the viscous flow $\boldsymbol{u}^v_{111}$. These coefficients mean that the higher-order version of (A 15) is

$$\mathrm{i}(\lambda - 1) + E^{1/2}s_1 + \mathrm{i}\beta_0 E^{1/2}C_1\mathrm{e}^{-2\mathrm{i}\psi_0} = -E\left(s_2^v + \beta_0 s_2^e\mathrm{e}^{-2\mathrm{i}\psi_0} + \mathrm{i}\beta_0^2 C_2\right). \tag{A 23}$$

For the situation studied in the main text, we have $\lambda = 1$, and the coefficients

$$s_1 = -0.618933(1+\mathrm{i}), \quad C_1 = 0.531174, \quad s_2^v = -10.3212,$$

$$s_2^e = -0.0886034 + 0.069463\mathrm{i}, \quad C_2 = 0.171714.$$

Since $E = 2.5 \times 10^{-4}$ is relatively large, the leading-order asymptotic predictions of $\beta_{crit} = 0.0261$ and $\psi_0 = -0.875\pi$ are a little distant from the numerical values of $\beta_{crit} = 0.0299$ and $\psi_0 = -0.8595\pi$. A next-order calculation using (A 23) produces much closer values,

$$\beta_{crit} = 0.0295, \quad \psi_0 = -0.855\pi. \tag{A 24}$$

A.4. $O(\epsilon^2)$: *nonlinear corrections*

The leading nonlinear corrections $\boldsymbol{u}_{200}$ and $\boldsymbol{u}_{202}$ are defined by the inhomogeneous system at $O(\epsilon^2)$ forced by the quadratic nonlinearity of the Navier–Stokes equation. The problems are

$$\frac{\partial \boldsymbol{u}_{200}}{\partial t} + 2\hat{\boldsymbol{k}}\times\boldsymbol{u}_{200} + \nabla p_{200} = -\boldsymbol{u}_{101}\cdot\nabla\boldsymbol{u}^*_{101} - \boldsymbol{u}^*_{101}\cdot\nabla\boldsymbol{u}_{101}, \tag{A 25}$$

$$\frac{\partial \boldsymbol{u}_{202}}{\partial t} + 2\hat{\boldsymbol{k}}\times\boldsymbol{u}_{202} + \nabla p_{202} = -\boldsymbol{u}_{101}\cdot\nabla\boldsymbol{u}_{101}, \tag{A 26}$$

with $\nabla\cdot\boldsymbol{u}_{200} = \nabla\cdot\boldsymbol{u}_{202} = 0$ and $\boldsymbol{u}_{200}\cdot\hat{\boldsymbol{n}} = \boldsymbol{u}_{202}\cdot\hat{\boldsymbol{n}} = 0$ on $s = 1$. The solutions are

$$\boldsymbol{u}_{202} = \begin{bmatrix} \mathrm{i}u_{202}(s) \\ v_{202}(s) \\ 0 \end{bmatrix} \mathrm{e}^{2\mathrm{i}\phi} := \frac{1}{9s}\begin{bmatrix} \mathrm{i}J_1^2(ks) - \mathrm{i}s^2J_1^2(k) \\ s^2J_1^2(k) - sJ_1(ks)\dfrac{\mathrm{d}J_1(ks)}{\mathrm{d}s} \\ 0 \end{bmatrix}\mathrm{e}^{2\mathrm{i}\phi}$$

$$p_{202} = -\tfrac{1}{3}J_1^2(ks)\mathrm{e}^{2\mathrm{i}\phi} - \tfrac{1}{2}\boldsymbol{u}^2_{101}, \tag{A 27}$$

$$\boldsymbol{u}_{200} = v_{200}(s)\cos 2\alpha z\hat{\boldsymbol{\phi}} := \frac{1}{9s}\left\{J_2^2(ks) - 3J_1^2(ks)\right\}\cos 2\alpha z\hat{\boldsymbol{\phi}},$$

$$p_{200} = \tfrac{1}{18}\left\{9J_0^2(ks) - J_2^2(ks)\right\}\cos 2\alpha z - \boldsymbol{u}_{101}\cdot\boldsymbol{u}^*_{101}. \tag{A 28}$$

No mean flow is *directly* generated at this order in keeping with the well-known general result that the interaction of any two inertial waves cannot drive geostrophic flows (Greenspan 1969).

A.5. *Calculation of mean flow*

Mean flow can first appear at $O(\epsilon^2E^{1/2})$ in the momentum equation driven by the nonlinear interactions of $\boldsymbol{u}_{101}$ with $\boldsymbol{u}^v_{111}$ and $\boldsymbol{u}^e_{111}$ together with an 'elliptical' contribution from the pressure p_{202}. In a cylinder with rigid lid and bottom, these nonlinear terms would be balanced by horizontal Ekman layers corresponding to an $O(\epsilon^2)$ mean flow. In an infinite cylinder, only the weak internal friction of the mean flow acts as a balancing agency and as a result an $O(\epsilon^2/E^{1/2})$ flow can be driven. Letting $\bar{v}_0(s) = |A|^2g_0(s) + \beta_0A^2\mathrm{e}^{-2\mathrm{i}t}h_0(s) + \beta_0A^{*2}\mathrm{e}^{2\mathrm{i}t}h^*_0(s)$, we have the problems

$$\frac{1}{2\pi d}\int_0^d\int_0^{2\pi}\hat{\boldsymbol{\phi}}\cdot\left(\boldsymbol{u}_{101}\cdot\nabla\boldsymbol{u}^{v*}_{111} + \boldsymbol{u}^{v*}_{111}\cdot\nabla\boldsymbol{u}_{101} + \text{c.c.}\right)\mathrm{d}\phi\,\mathrm{d}z = \left(\nabla^2 - \frac{1}{s^2}\right)g_0,$$

$$\frac{1}{2\pi d}\int_0^d\int_0^{2\pi}\hat{\boldsymbol{\phi}}\cdot\left(\boldsymbol{u}_{101}\cdot\nabla\boldsymbol{u}^{e*}_{111} + \boldsymbol{u}^{e*}_{111}\cdot\nabla\boldsymbol{u}_{101} - \tfrac{1}{2}\mathrm{e}^{-2\mathrm{i}\phi}\mathscr{N}^*\nabla p_{202}\right)\mathrm{d}\phi\,\mathrm{d}z = \left(\nabla^2 - \frac{1}{s^2}\right)h_0,$$

to be solved subject to the homogeneous boundary conditions $g_0(0) = h_0(0) = g_0(1) = h_0(1) = 0$.

At next order $O(\epsilon^2)$, we look for a solution of similar form $\bar{v}_1 = |A|^2g_1(s) +$

$\beta_0 A^2 \mathrm{e}^{-2\mathrm{i}t} h_1(s) + \beta_0 A^{*2} \mathrm{e}^{2\mathrm{i}t} h_1^*(s)$. The calculation of $g_1(s)$ involves solving

$$\frac{1}{2\pi d}\int_0^d\int_0^{2\pi} \hat{\boldsymbol{\phi}}\cdot[\,\boldsymbol{u}_{101}\cdot\nabla\boldsymbol{u}_{121}^{vv*} + \boldsymbol{u}_{121}^{vv*}\cdot\nabla\boldsymbol{u}_{101} + \boldsymbol{u}_{111}^{v}\cdot\nabla\boldsymbol{u}_{111}^{v*} + \text{c.c.}\,]\,\mathrm{d}\phi\,\mathrm{d}z\left(\nabla^2 - \frac{1}{s^2}\right)g_1. \tag{A 29}$$

The boundary condition on g_1 at the sidewall is modified by a boundary layer in which $\boldsymbol{u}_{101}$ interacts with its boundary correction $\tilde{\boldsymbol{u}}_{101}$ over a distance of $O(E^{1/2})$ from $s = 1$. Solving the boundary layer problem,

$$\frac{1}{2\pi d}\int_0^d\int_0^{2\pi} \hat{\boldsymbol{\phi}}\cdot[\,\boldsymbol{u}_{101}\cdot\nabla\tilde{\boldsymbol{u}}_{101} + \tilde{\boldsymbol{u}}_{101}\cdot\nabla\boldsymbol{u}_{101} + \tilde{\boldsymbol{u}}_{101}\cdot\nabla\tilde{\boldsymbol{u}}_{101}^{*} + \text{c.c.}\,]\,\mathrm{d}\phi\,\mathrm{d}z = \frac{\partial^2}{\partial\xi^2}\tilde{g}_1, \tag{A 30}$$

where ξ is the stretched boundary layer variable $(1-s)/E^{1/2}$ and $\tilde{g}_1(\xi)$ is the boundary correction to $g_1(s)$, leads to the modified boundary condition

$$g_1(1) = -\tfrac{3}{2}v_0(v_0 + \alpha w_0)|_{s=1}$$

(see (A 10) for definitions of v_0 and w_0) along with $g_1(0) = 0$. Due to phase coherence, nonlinear interactions between pure elliptical flows vanish identically, i.e.

$$\frac{1}{2\pi d}\int_0^d\int_0^{2\pi} \hat{\boldsymbol{\phi}}\cdot[\,\boldsymbol{u}_{101}\cdot\nabla\boldsymbol{u}_{121}^{ee*} + \boldsymbol{u}_{121}^{ee*}\cdot\nabla\boldsymbol{u}_{101} + \boldsymbol{u}_{111}^{e}\cdot\nabla\boldsymbol{u}_{111}^{e*} + \boldsymbol{u}_{113}^{e}\cdot\nabla\boldsymbol{u}_{113}^{e*} + \text{c.c.}\,]\,\mathrm{d}\phi\,\mathrm{d}z = 0$$

which just leaves a straightforward calculation for $h_1(s)$:

$$\frac{1}{2\pi d}\int_0^d\int_0^{2\pi} \hat{\boldsymbol{\phi}}\cdot[\,\boldsymbol{u}_{101}\cdot\nabla\boldsymbol{u}_{121}^{ev*} + \boldsymbol{u}_{121}^{ev*}\cdot\nabla\boldsymbol{u}_{101} + \boldsymbol{u}_{111}^{v}\cdot\nabla\boldsymbol{u}_{111}^{e*} + \boldsymbol{u}_{111}^{e*}\cdot\nabla\boldsymbol{u}_{111}^{v}\,]\,\mathrm{d}\phi\,\mathrm{d}z = \left(\nabla^2 - \frac{1}{s^2}\right)h_1 \tag{A 31}$$

where $h_1(0) = h_1(1) = 0$.

A.6. *Saturation*

With an $O(\epsilon^2/E^{1/2})$ mean flow the leading saturation balance is potentially derived by the solvability condition on the $O(\epsilon^3/E^{1/2})$ momentum equation for $\boldsymbol{u}_{3-11}$. This gives

$$\mathrm{i}(\beta_2 A_0^* + \beta_0 E^{1/2} A_2^*)C_1\mathrm{e}^{2\mathrm{i}t} + [\mathrm{i}(\lambda - 1) + s_1 E^{1/2}]A_2 = -\frac{1}{E^{1/2}}N(\boldsymbol{u}_{101}, \bar{v}_0\hat{\boldsymbol{\phi}}) \tag{A 32}$$

where we define

$$N(\boldsymbol{u},\boldsymbol{v}) := \frac{\langle \boldsymbol{u}_{101}, \boldsymbol{u}\times\nabla\times\boldsymbol{v} + \boldsymbol{v}\times\nabla\times\boldsymbol{u}\rangle}{\langle \boldsymbol{u}_{101}, \boldsymbol{u}_{101}\rangle}. \tag{A 33}$$

The right-hand side is

$$-A|A|^2 N(\boldsymbol{u}_{101}, g_0\hat{\boldsymbol{\phi}}) - \beta_0 A\left[A^2\mathrm{e}^{-2\mathrm{i}t}N(\boldsymbol{u}_{101}, h_0\hat{\boldsymbol{\phi}}) + A^{*2}\mathrm{e}^{2\mathrm{i}t}N(\boldsymbol{u}_{101}, h_0^*\hat{\boldsymbol{\phi}})\right]$$

which reduces to

$$-A\left[\sigma_1 + 2\mathrm{i}\beta_0\sigma_2\sin 2\psi_0\right],$$

where $N(\boldsymbol{u}_{101}, g_0\hat{\boldsymbol{\phi}}) = \sigma_1 = 1.010\times10^{-2}\mathrm{i}$ and $N(\boldsymbol{u}_{101}, h_0\hat{\boldsymbol{\phi}}) = -N(\boldsymbol{u}_{101}, h_0^*\hat{\boldsymbol{\phi}}) = \sigma_2 = -4.334\times10^{-3}$. Substituting in the leading result, $\sin 2\psi_0 = -s_1^R/(\beta_0 C_1)$, causes this whole expression to vanish identically so that the mean flow, $\bar{v}_0$, is strictly zero at this

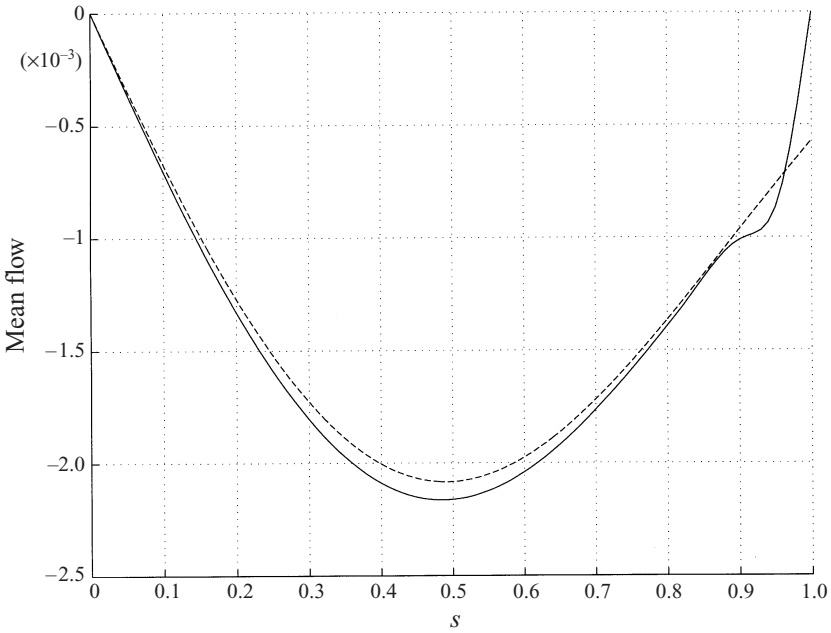

FIGURE 19. A comparison of the mean flow at saturation ($\beta = 0.031$, $E = 2.5 \times 10^{-4}, d = 3.9796$) for the $m = (-1, 1)$ elliptical instability obtained from the numerics (solid line) and the weakly nonlinear prediction (dashed) using $\epsilon = 0.06613$ and $\psi_0 = 0.855\pi$ (plotted against radius s).

order. This is a non-trivial result which in its 'simplest form'

$$\hat{\boldsymbol{\phi}} \cdot \left\{ -\mathrm{i}\boldsymbol{u}_{101} \times \nabla \times \left[\mathrm{e}^{-2\mathrm{i}\phi} \mathscr{N}^* \nabla p_{101} \right] - \mathscr{N}^* \nabla p_{202} \right\} \equiv 0 \tag{A 34}$$

in our non-orthogonal elliptico-polar coordinates, means that, in fact, $h_0(s) = -\mathrm{i}C_1/(2s_1^R) \times g_0(s)$. In terms of β_2 and ψ_2, the leading coefficients β_{2-1} and ψ_{2-2} vanish so that we must determine β_{20} and ψ_{2-1} at next order to capture saturation.

The solvability condition on the $O(\epsilon^3)$ equation for $\boldsymbol{u}_{301}$ is

$$\mathrm{i}(\beta_2 A_0^* + \beta_0 E^{1/2} A_2^*) C_1 \mathrm{e}^{2\mathrm{i}t} + \mathrm{i}(\lambda - 1) A_2 + s_1 E^{1/2} A_2 = -N(\boldsymbol{u}, \boldsymbol{u}), \tag{A 35}$$

where $A_2 = \mathrm{i}\psi_2 \mathrm{e}^{\mathrm{i}(t+\psi_0)}$ so that this represents a complex equation for β_{20} and ψ_{2-1}. The right-hand side of (A 35) reduces to

$$N(\boldsymbol{u}, \boldsymbol{u}) = \frac{A_0}{E^{1/2}} \left[\sigma_1 + 2\mathrm{i}\beta_0 \sigma_2 \sin 2\psi_0 \right] + A_0 |A_0|^2 \{ \sigma_3 + \sigma_4 + \sigma_5 \} + \beta_0 A_0^3 \mathrm{e}^{-2\mathrm{i}t} \sigma_6 - \beta_0 A_0^* |A_0|^2 \mathrm{e}^{2\mathrm{i}t} \sigma_6^*, \tag{A 36}$$

where $\sigma_i \in \mathbb{C}$ are defined as follows:

$$\begin{aligned} N(\boldsymbol{u}_{101}^*, \boldsymbol{u}_{202}) &= \sigma_3 = 5.685 \times 10^{-2}\mathrm{i}, \\ N(\boldsymbol{u}_{101}, \boldsymbol{u}_{200}) &= \sigma_4 = -8.679 \times 10^{-2}\mathrm{i}, \\ N(\boldsymbol{u}_{101}, g_1 \hat{\boldsymbol{\phi}}) &= \sigma_5 = 5.717 \times 10^{-3}\mathrm{i}, \\ -N(\boldsymbol{u}_{101}, h_1^* \hat{\boldsymbol{\phi}})^* = N(\boldsymbol{u}_{101}, h_1 \hat{\boldsymbol{\phi}}) &= \sigma_6 = 4.026 \times 10^{-3} - 3.870 \times 10^{-3}\mathrm{i} \end{aligned}$$

and the first term on the right is carried over from the previous order. The left-hand

side of (A 35) can be simplified using the leading-order balance (A 15) to give

$$2\beta_0 C_1 \psi_{2-1} + \mathrm{i}\beta_{20} = -\frac{\mathrm{e}^{2\mathrm{i}\psi_0}}{E^{1/2}}\left[\sigma_1 + 2\mathrm{i}\beta_0\sigma_2 \sin 2\psi_0\right] - \mathrm{e}^{2\mathrm{i}\psi_0}\left\{\sigma_3 + \sigma_4 + \sigma_5\right\} - \beta_0 \mathrm{e}^{4\mathrm{i}\psi_0}\sigma_6 + \beta_0\sigma_6^* \quad \text{(A 37)}$$

and ultimately

$$\beta_{20} = 0.2272, \qquad \psi_{2-1} = -0.0760. \quad \text{(A 38)}$$

The bifurcation is therefore a supercritical pitchfork bifurcation which is confirmed by the numerical computations. To compare the theoretical predictions (A 38) with the numerics, we take the numerically calculated critical values of 0.0299 for β_{crit} and -0.8595π for ψ_0 to produce figure 8. This eliminates the discrepancy in the critical point prediction which arises through $E^{1/2}$ being relatively large and instead focuses attention on the ϵ expansion procedure. Figure 19 shows that the predicted mean flow matches the numerical solution well except, of course, for the boundary layer which is suppressed in the weakly nonlinear theory. Since the mean flow only emerges at second order, it is very sensitive to the phase angle ψ because of the delicate leading balance which exists. Taking $\psi_0 = -0.853\pi$ rather than the asymptotic value -0.855π actually used in figure 19 causes the theoretical and numerical mean flows to lie on top of each other for much of the interior.

REFERENCES

Aldridge, K., Seyed-Mahmoud, B., Henderson, G. & Wijngaarden, W. van 1997 Elliptical instability of the Earth's fluid core. *Phys. Earth Planet. Int.* **103**, 365–374.

Aldridge, K. D. & Toomre, A. 1969 Axisymmetrc inertial oscillations of a fluid in a rotating spherical container. *J. Fluid Mech.* **37**, 307–323.

Bayly, B. J. 1986 Three-dimensional instability of elliptical flow. *Phys. Rev. Lett.* **57**, 2160–2163.

Bayly, B. J., Holm, D. D. & Lifschitz, A. 1996 Three-dimensional stability of eliiptical vortex columns in external strain flows. *Phil. Trans. R. Soc. Lond.* **354**, 895–926.

Bayly, B. J., Orszag, S. A. & Herbert, T. 1988 Instability mechanisms in shear-flow transition. *Ann. Rev. Fluid Mech.* **20**, 359–391.

Caulfield, C. P. & Peltier, W. R. 1999 The anatomy of the mixing transition in homogeneous and stratified free shear layers. *J. Fluid Mech.* (submitted).

Craik, A. D. D. 1989 The stability of unbounded two- and three-dimensional flows subject to body forces: some exact solutions. *J. Fluid Mech.* **198**, 275–292.

Fabijonas, B., Holm, D. D. & Lifschitz, A. 1997 Secondary instabilities of flows with elliptic streamlines. *Phys. Rev. Lett.* **78**, 1900–1903.

Gledzer, E. B., Dolzhanskiy, F. V., Obukhov, A. M. & Ponomarev, V. M. 1975 An experimental and theoretical study of the stability of motion of a liquid in an elliptical cylinder. *Isv. Atmos. Ocean. Phys.* **11**, 617–622.

Gledzer, E. B., Novibov, Yu. V., Obukhov, A. M. & Chusov, M. A. 1974 An investigation of the stability of liquid flows in a three-axis ellipsoid. *Isv. Atmos. Ocean. Phys.* **10**, 69–71.

Gledzer, E. B. & Ponomarev, V. M. 1992 Instability of bounded flows with elliptical streamlines. *J. Fluid Mech.* **240**, 1–30.

Gollub, J. P. & Benson, S. V. 1980 Many routes to turbulent convection. *J. Fluid Mech.* **100**, 449–470.

Gollub, J. P. & Swinney, H. L. 1975 Onset of turbulence in a rotating fluid. *Phys. Rev. Lett.* **35**, 927–930.

Greenspan, H. P. 1968 *The Theory of Rotating Fluids.* Cambridge University Press (reprinted Breukelen Press 1990).

Greenspan, H. P. 1969 On the non-linear interaction of inertial modes. *J. Fluid Mech.* **36**, 257–264.

Jimenez, J., Moffatt, H. K. & Vasco, C. 1996 The structure of the vortices in freely decaying two-dimensional turbulence *J. Fluid Mech.* **313**, 209–222.

Kerswell, R. R. 1993 The instability of precessing flow. *Geophys. Astrophys. Fluid Dyn.* **72**, 107–144.
Kerswell, R. R. 1999 Secondary instabilities in rapidly rotating fluids: inertial wave breakdown. *J. Fluid Mech.* **382**, 283–306.
Kerswell, R. R. & Barenghi, C. F. 1995 On the viscous decay rates of inertial waves in a rotating circular cylinder. *J. Fluid Mech.* **285**, 203–214.
Kerswell, R. R. & Davey, A. 1996 On the linear instability of elliptic pipe flow. *J. Fluid Mech.* **316**, 307–324.
Kerswell, R. R. & Malkus, W. V. R. 1998 Tidal instability as the source for Io's magnetic signature. *Geophys. Res. Lett.* **25**, 603–606.
Kobine, J. J. 1995 Inertial wave dynamics in a rotating and precessing cylinder. *J. Fluid Mech.* **303**, 233–252.
Kobine, J. J. 1996 Azimuthal flow associated with inertial wave resonance in a precessing cylinder. *J. Fluid Mech.* **319**, 387–406.
Landman, M. J. & Saffman, P. G. 1987 The three-dimensional instability of strained vortices in a viscous fluid. *Phys. Fluids.* **30**, 2339–2342.
Lifschitz, A. & Fabijonas, B. 1996 A new class of instabilities of rotating fluids. *Phys. Fluids* **8**, 2239–2241.
Lundgren, T. S. & Mansour, N. N. 1996 Transition to turbulence in an elliptic vortex. *J. Fluid Mech.* **307**, 43–62.
Malkus, W. V. R. 1968 Precession of the Earth as the cause of geomagnetism. *Science* **160**, 259–264.
Malkus, W. V. R. 1989 An experimental study of the global instabilities due to the tidal (elliptical) distortion of a rotating elastic cylinder. *Geophys. Astrophys. Fluid Dyn.* **48**, 123–134.
Malkus, W. V. R. & Waleffe, F. A. 1991 Transition from order to disorder in elliptical flow: a direct path to shear flow turbulence. In *Advances in Turbulence 3* (ed. A. V. Johansson & P. H. Alfredsson), pp. 197–203. Springer.
Manasseh, R. 1992 Breakdown regimes of inertia waves in a precessing cylinder. *J. Fluid Mech.* **243**, 261–296.
Manasseh, R. 1994 Distortions of inertia waves in a rotating fluid cylinder forced near its fundamental mode resonance. *J. Fluid Mech.* **265**, 345–370.
Manasseh, R. 1996 Nonlinear behaviour of contained inertia waves. *J. Fluid Mech.* **315**, 151–173.
Marcus, P. S. 1984*a* Simulation of Taylor–Couette Flow. Part 1. Numerical methods and comparison with experiment. *J. Fluid Mech.* **146**, 45–64.
Marcus, P. S. 1984*b* Simulation of Taylor–Couette Flow. Part 2. Numerical results for wavy-vortex flow with one travelling wave. *J. Fluid Mech.* **146**, 65–113.
McEwan, A. D. 1970 Inertial oscillations in a rotating fluid cylinder. *J. Fluid Mech.* **40**, 603–640.
Moffatt, H. K., Kida, S. & Ohkitani, K. 1994 Stretched vortices – the sinews of turbulence; large-Reynolds-number asymptotics. *J. Fluid Mech.* **259**, 241–264.
Newhouse, S., Ruelle, D. & Takens, F. 1978 Occurence of Strange Axiom A Attractors near quasiperiodic flows on T^m, $m \geqslant 3$. *Commun. Math. Phys.* **64**, 35–40.
Orszag, S. A. & Patera, A. T. 1983 Secondary instability of wall-bounded shear flows. *J. Fluid Mech.* **128**, 347–385.
Pierrehumbert, R. T. 1986 Universal short-wave instability of two-dimensional eddies in an inviscid fluid. *Phys. Rev. Lett.* **57**, 2157–2159.
Ruelle, D. & Takens, F. 1971 On the nature of turbulence. *Commun. Math. Phys.* **20**, 167–192.
Tilgner, A. 1999 Driven inertial oscillations in spherical shells. *Phys. Rev.* E **59**, 1789–1794.
Vanyo, J. P., Wilde, P., Cardin, P. & Olson, P. 1995 Experiments on precessing flows in the Earth's liquid core *Geophys. J. Intl* **121**, 136–142.
Vladimirov, V. A. & Tarasov, V. 1985 Resonance instability of the flows with closed streamlines. In *Laminar-Turbulent Transition; IUTAM Symposium Novosibirsk* (*1984*) (ed. V. V. Kozlov), pp 717–722. Springer.
Vladimirov, V. A. & Vostretsov, D. 1986 Instability of steady flows with constant vorticity in vessels of elliptic cross-section. *Prikl. Matem. Mekhan.* **50**(3), 367–377. (Transl. in *J. Appl. Math. Mech.* **50**(3), 369–377.)
Waleffe, F. A. 1989 The 3D instability of a strained vortex and its relation to turbulence. PhD thesis, MIT.
Waleffe, F. A. 1990 On the three-dimensional instability of a strained vortex. *Phys. Fluids* A **2**, 76–80.