J. Fluid Mech. (2008), *vol.* 617, *pp.* 187–205.
doi:10.1017/S0022112008003959 Printed in the United Kingdom

Kelvin waves in the nonlinear shallow water equations on the sphere: nonlinear travelling waves and the corner wave bifurcation

JOHN P. BOYD† AND CHENG ZHOU

Department of Atmospheric, Oceanic and Space Science, University of Michigan, Ann Arbor, MI 48109-2143, USA

(Received 17 August 2007 and in revised form 20 August 2008)

The Kelvin wave is the lowest eigenmode of Laplace's tidal equation and is widely observed in both the ocean and the atmosphere. In this work, we neglect mean currents and instead include the full effects of the Earth's sphericity and the wave dispersion it induces. Through a mix of perturbation theory and numerical computations using a Fourier/Newton iteration/continuation method, we show that for sufficiently small amplitude, there are Kelvin travelling waves (cnoidal waves). As the amplitude increases, the branch of travelling waves terminates in a so-called corner wave with a discontinuous first derivative. All waves larger than the corner wave evolve to fronts and break. The singularity is a point singularity in which only the longitudinal derivative is discontinuous. As we solve the nonlinear shallow water equations on the sphere, with increasing ϵ ('Lamb's parameter'), dispersion weakens, the amplitude of the corner wave decreases rapidly, and the longitudinal profile of the corner wave narrows dramatically.

1. Introduction

The free oscillations of a layer of homogeneous fluid and uniform depth on a rotating, spherical Earth are governed by a trio of nonlinear partial differential equations, the 'Laplace tidal equations', also known as the 'nonlinear shallow water wave equations'. When linearized about a state of rest, these equations have eigenmodes which are commonly called 'Hough' functions. The slowest eastward-travelling wave has been given the special name of the 'Kelvin wave.'

The Kelvin wave has enormous practical importance (Chapman & Lindzen 1970; Andrews *et al.* 1987; Majda 2003). A Kelvin wave is the main oceanic component of the coupled ocean/atmosphere oscillation known as ENSO (El Niño–Southern Oscillation) which combines torrential rains in California with the drought known as the 'Great Dry' in Australia. Kelvin waves are important in the troposphere and the middle atmosphere; they are the primary drivers of the quasi-biennial oscillation (QBO) in the tropical lower stratosphere and the semi-annual oscillation in the tropical upper stratosphere and are also important in Martian atmospheric dynamics (Zurek 1976). Because the Kelvin wave is the lowest mode of the atmosphere, it is the most strongly excited wave from any sort of broadband forcing.

Longuet-Higgins (1968) carried out a long study of the linearized eigenmodes of the shallow water equations on the sphere, known as Hough functions, nearly 40 years ago. The Kelvin mode is the lowest latitudinal Hough function for each longitudinal

† Email address for correspondence: jpboyd@umich.edu

wavenumber s. The nonlinear self-interaction of Kelvin waves has been studied by Boyd (1980, 1984, 1991, 1998), Ripa (1982, 1985), Greatbatch (1985), Marshall & Boyd (1987), Long & Chang (1990), Majda *et al.* (1999), Milewski & Tabak (1999) Fedorov & Melville (2000), Chen & Boyd (2002), and Le Sommer *et al.* (2004). In spite of this work, there are still significant gaps in both linear and nonlinear theories. Some of the linear lacunae have recently been filled by Boyd (2005*d*) and Boyd & Zhou (2008).

The nonlinear shallow water wave equations also describe the baroclinic mode of a two-layer model in the limit that the lower layer depth is infinite, in which case motion is confined to the upper layer, a so-called one-and-a-half-layer model (Gill 1982). (This is a decent first approximation of the ocean, especially in the tropics.) The only modification is that the actual mean depth is replaced by the 'equivalent depth', which is the product of the mean depth with the fractional density difference between the two layers (Pedlosky 1987). As explained in Marshall & Boyd (1987) and other references cited there, the shallow water equations also model the baroclinic modes of a continuously stratified fluid as first observed by G. I. Taylor more than 70 years ago; the main effect of continuous stratification is to slightly weaken the nonlinearity because of coupling between different vertical modes.

In this paper, we extend the nonlinear shallow water equation theory in a couple of ways. First, instead of using the equatorial beta-plane, which corresponds to the limit of a very thin ocean, we explicitly include the effects of the Earth's sphericity and finite depth (i.e. finite 'Lamb's parameter'). In the equatorial beta-plane approximation, travelling waves can be modelled by applying the method of multiple scales to derive the Korteweg–deVries (KdV) equation (with mean currents) or the inviscid Burgers equation (neglecting mean currents) and then invoking the known analytic travelling waves of these models. On the sphere, it is impossible to derive a KdV model. However, a mixture of perturbation theory (for small amplitude) and a continuation/Fourier–Galerkin/Newton iteration algorithm (for larger amplitude) allows us to describe the nonlinear Kelvin wave on the sphere.

The nonlinear Kelvin wave exhibits, with increasing amplitude, the so-called cnoidal/corner/breaking (CCB) scenario. That is, for small amplitude, nonlinear Kelvin waves are steadily travelling periodic waves similar to the cnoidal waves of the KdV equation. The branch of solutions terminates in a wave of finite amplitude which is a 'corner wave' in the sense that there is a slope discontinuity at the crest (Boyd 2003, 2006). The waves whose initial amplitude is higher than that of the corner wave rapidly steepen to an infinite slope ('wave breaking').

Much is known about the CCB scenario as reviewed in Grimshaw *et al.* (1998), Boyd (2003, 2005*a*, *b*) and pioneering work has been done by Stokes (1847), Ostrovsky (1978) and Shrira (1981, 1986). Near-corner waves are described through matched asymptotic expansions in Longuet-Higgins & Fox (1996) and Boyd (2005*c*). Although we shall not compute initial-value solutions here, the statement that large amplitude Kelvin waves break is demonstrated in Boyd (2005*a*, 2006) and Chen & Boyd (2002).

Kelvin breaking has been discussed elsewhere (Boyd 1980, 1998; Fedorov & Melville 2000; Chen & Boyd 2002; Le Sommer, Reznik & Zeitlin 2004). Here we will attempt to describe the travelling waves up to and including the corner wave. This poses a severe numerical challenge because in the corner wave limit, the wave will have a discontinuous x-derivative at the peak of the wave. The convergence rate of Fourier coefficients of functions with a slope discontinuity is only $O(K^{-2})$, where K is the degree of the Fourier coefficients. By employing the Kepler mapping developed in Boyd (2006), the convergence rate can be improved to $O(K^{-4})$.

ϵ	Description	Source
0.012	External mode: Venus	Lindzen (1970)
6.5	External mode: Mars	Zurek (1976)
12	External mode: Earth (7.5 km equivalent depth)	Lindzen (1970)
2.6	Jupiter: simulate Galileo data	Williams (1996)
21.5	Jupiter	Williams (1996)
43.0	Jupiter	Williams (1996)
260	Jupiter	Williams (1996)
2600	Jupiter	Williams & Wilson (1988)
87 000	Ocean: first baroclinic mode (1 m equivalent depth)	Moore & Philander (1977)
$>100\,000$	Ocean: higher baroclinic modes	Moore & Philander (1977)

TABLE 1. Lamb's parameter.

The Kelvin wave depends on two parameters. The zonal wavenumber s is always a positive integer. (The case of $s=0$ is a non-propagating mode which is not relevant here.)

Lamb's parameter ϵ is a non-dimensional mean reciprocal depth which is explicitly

$$\epsilon = \frac{4\Omega^2 a^2}{gH}, \tag{1.1}$$

where Ω is the angular frequency of the Earth's rotation in radians per second; a is the radius of the planet; g is the gravitational constant, which is $9.8\,\mathrm{m\,s^{-2}}$ for Earth; and H is the mean depth of the fluid. As explained in Chapman & Lindzen (1970), Majda (2003) and Marshall & Boyd (1987), Laplace's equations can be profitably employed for continuously stratified (rather than homogeneous) fluids if the depth H is interpreted as the 'equivalent depth' of a given baroclinic mode. Thus, to describe all possible varieties of Kelvin waves in a three-dimensional stratified ocean or atmosphere, one needs to solve Laplace's Tidal equations for a very wide range of ϵ ranging from very small (for the 'barotropic' or 'nearly barotropic' waves) to very large (for high-order baroclinic modes) as illustrated in table 1.

When s and ϵ are both small, the Kelvin wave fills the entire globe from pole to pole. When the parameter combination $r \equiv \sqrt{s^2+\epsilon}$ is large compared to one, the Kelvin wave is equatorially trapped, proportional to $\exp(-(1/2)r\mu^2)$, where μ is the sine of latitude. When ϵ is large and much greater than s^2, the Kelvin wave is well approximated by the equatorial beta-plane. When $s \gg \sqrt{\epsilon}$ (and not necessarily large), the velocity potential $\chi \approx \exp(\mathrm{i}s\lambda)\, P_s^s(\mu)$, where P_s^s is the usual associated Legendre function, and the frequency is $\sigma \approx \sqrt{s(s+1)}/\sqrt{\epsilon}$. A new asymptotic approximation is derived in Boyd & Zhou (2008), which fills the gap between these two previously known limits.

In the next section, we apply perturbation theory to analyse travelling waves of small amplitude. Section 3 describes the numerical methods that will be employed. Our perturbative and numerical analysis will explicitly treat only low zonal wavenumbers, $s=1$ and $s=2$. However, the methodologies are completely general. We concentrate on 'small' zonal wavenumbers because, as illustrated in Boyd & Zhou (2008), Kelvin waves of moderate and large s are 'equatorial' rather than 'global' modes and are therefore well modelled by the equatorial beta-plane studies of previous work (Boyd 1998, 2006). The spatial strucure of the corner wave is analysed in § 4. Variations of the phase speed and height of the corner wave are described in following section. Comparisons with observations are then discussed followed by a summary.

2. Nonlinear Kelvin wave on the sphere by perturbation theory

2.1. *Introduction*

In this section, we shall approximate the travelling wave solution of the nonlinear Kelvin wave on the sphere by a perturbation series. First, the nonlinear shallow water equations will be non-dimensionalized. Second, we will simplify by shifting to the velocity variables introduced by Margules and changing the north–south coordinate. Finally, the travelling wave solution with period 2π will be perturbatively derived by expanding the unknowns in a double power series in the wave amplitude and in the square root of Lamb's parameter ϵ.

2.2. *Nonlinear shallow water equations on the sphere*

The nonlinear shallow water equations in spherical coordinates over a flat sea bottom are

$$\frac{\partial u'}{\partial t'} + \frac{u'}{a\sin(\theta)}\frac{\partial u'}{\partial \lambda} - \frac{v'}{a}\frac{\partial u'}{\partial \theta} - \left(2\Omega\cos(\theta) + \frac{u'\cot(\theta)}{a}\right)v' + \frac{g}{a\sin(\theta)}\frac{\partial h'}{\partial \lambda} = 0, \quad (2.1)$$

$$\frac{\partial v'}{\partial t'} + \frac{u'}{a\sin(\theta)}\frac{\partial v'}{\partial \lambda} - \frac{v'}{a}\frac{\partial v'}{\partial \theta} + \left(2\Omega\cos(\theta) + \frac{u'\cot(\theta)}{a}\right)u' - \frac{g}{a}\frac{\partial h'}{\partial \theta} = 0, \quad (2.2)$$

$$\frac{\partial h'}{\partial t'} + \frac{u'}{a\sin(\theta)}\frac{\partial h'}{\partial \lambda} - \frac{v'}{a}\frac{\partial h'}{\partial \theta} + \frac{h'}{a\sin(\theta)}\left(\frac{\partial u'}{\partial \lambda} - \sin(\theta)\frac{\partial v'}{\partial \theta} - v'\cos(\theta)\right) = 0, \quad (2.3)$$

where θ is colatitude; λ is longitude; t' is the dimensional time; u' is the dimensional eastward velocity; v' is the dimensional northward velocity; h' is the total depth of the fluid; and a is the radius of the planet. (Note that following the usual geophysical convention, $v' = -a\,\mathrm{D}\theta/\mathrm{D}t$, where $\mathrm{D}/\mathrm{D}t$ is the total derivative because colatitude θ increases southward.)

It is convenient to non-dimensionalize the variables as follows:

$$t' = (1/[2\Omega])\sqrt{\epsilon}\,t, \quad (2.4)$$

$$u' = 2\Omega a\frac{\hat{u}}{\sin(\theta)}, \qquad v' = 2\Omega a\frac{\hat{v}}{\sin(\theta)}, \quad (2.5)$$

$$h = H\,(1 + \sqrt{\epsilon}\,\phi). \quad (2.6)$$

Note that $\hat{u}$ and $\hat{v}$ are the 'Margules velocities', introduced in the early 20th century by the Austrian meteorologist Max Margules because the $\sin(\theta)$ factors ensure that these modified velocities have the same behaviour at the poles as scalar variables like the height h.

It is common in tidal theory to rescale time by 2Ω, velocities by $2\Omega a$ and the deviation height field by ϵ. We have chosen different scales for time and ϕ for convenience in applying perturbation theory as explained in the next section. We employed the same scalings for the numerical studies to facilitate comparisons between perturbative and numerical results.

For notational simplicity, we omit the accents on the Margules velocities below.

In the meridional direction, we change the coordinate from θ to μ via

$$\cos(\theta) = \mu,\ \sin(\theta) = \sqrt{1-\mu^2},\ \ \frac{\partial}{\partial\theta} \to -\sqrt{1-\mu^2}\frac{\partial}{\partial\mu}. \quad (2.7)$$

We now specialize to travelling waves of the form $u(\lambda - ct, \mu)$ and similarly for the other variables, where c is the non-dimensional phase speed. The usual dimensional

phase speed in units of metres/second is

$$c_{dim} = a\,\frac{2\Omega}{\sqrt{\epsilon}}\,c. \tag{2.8}$$

The eigenmodes of Laplace's tidal equations, which are the linearization of the shallow water equations about a state of rest, are standing waves in latitude and propagate only in the east–west direction, and this is true of the nonlinear travelling waves as well. To explicitly collapse the number of coordinates from three (λ, μ, t) to two, it is convenient to define a new coordinate in a frame of reference that moves with the wave:

$$x \equiv \lambda - ct. \tag{2.9}$$

Then

$$\frac{\partial}{\partial t} \to -c\frac{\partial}{\partial x}, \quad \frac{\partial}{\partial \lambda} \to \frac{\partial}{\partial x}. \tag{2.10}$$

After all these steps and defining $\delta \equiv \sqrt{\epsilon}$, the shallow water equations become, without approximation,

$$(c(1-\mu^2)-\delta u)\frac{\partial u}{\partial x} - \delta(1-\mu^2)v\frac{\partial u}{\partial \mu} - (1-\mu^2)\frac{\partial \phi}{\partial x} + \delta\mu(1-\mu^2)v = 0, \tag{2.11}$$

$$(c(1-\mu^2)-\delta u)\frac{\partial v}{\partial x} - \delta(1-\mu^2)v\frac{\partial v}{\partial \mu} - (1-\mu^2)^2\frac{\partial \phi}{\partial \mu} - \delta\mu\left\{u^2+v^2+(1-\mu^2)u\right\} = 0, \tag{2.12}$$

$$(c(1-\mu^2)-\delta u)\frac{\partial \phi}{\partial x} - \delta(1-\mu^2)v\frac{\partial \phi}{\partial \mu} - (1+\delta\phi)\left(\frac{\partial u}{\partial x} + (1-\mu^2)\frac{\partial v}{\partial \mu}\right) = 0. \tag{2.13}$$

2.3. *Travelling waves by perturbative double expansion*

The travelling wave solution of (2.11)–(2.13) can be approximated by perturbation theory. As is standard in nonlinear wave theory, we assume the wave amplitude is small as measured by a placeholder variable A. Unfortunately, linearizing in A merely generates Laplace's tidal equations whose solutions, the Hough functions, cannot be found analytically except within the limits $\epsilon \to 0$ and $\epsilon \to \infty$. We therefore perform a double expansion on the further assumption that ϵ is sufficiently small. Then u,v,ϕ,c can be expressed as series of A and δ, which is $\sqrt{\epsilon}$:

$$u = \sum_{i=0}^{n}\sum_{j=1}^{m} u_{i,j}\delta^i A^j,\; v = \sum_{i=0}^{n}\sum_{j=1}^{m} v_{i,j}\delta^i A^j,\; \phi = \sum_{i=0}^{n}\sum_{j=1}^{m} \phi_{i,j}\delta^i A^j,\; c = \sum_{i=0}^{n}\sum_{j=0}^{m} c_{i,j}\delta^i A^j. \tag{2.14}$$

We substitute this expression into (2.11)–(2.13) and collect the coefficients of $\delta^i A^j$ from each equation. Perturbation theory requires all these coefficients should be zero. The lowest order equations, coefficients of $\delta^0 A^1$, are simply the $\epsilon = 0$ limit of Laplace's (linear) tidal equations:

$$c_{0,0}u_{0,1,x} - \phi_{0,1,x} = 0, \tag{2.15}$$

$$c_{0,0}v_{0,1,x} - (1-\mu^2)\phi_{0,1,\mu} = 0, \tag{2.16}$$

$$u_{0,1,x} + (1-\mu^2)v_{0,1,\mu} - c_{0,0}(1-\mu^2)\phi_{0,1,x} = 0, \tag{2.17}$$

where $\phi_{0,1}$ and $v_{0,1}$ can be quickly eliminated, and the equations can be made one single equation of $u_{0,1}(x,\mu)$ and $c_{0,0}$. Furthermore, let $u_{0,1}(x,\mu) = U(\mu)\cos(sx)$, where

s is the zonal wave number. We get

$$(1-\mu^2)\frac{\partial^2 U(\mu)}{\partial \mu^2} - 2\mu\frac{\partial U(\mu)}{\partial \mu} + \left(c_{0,0}^2 s^2 - \frac{s^2}{1-\mu^2}\right) U(\mu) = 0. \tag{2.18}$$

This is Legendre's equation. Since $c_{0,0}$ is unknown, it is an eigenvalue problem. The solutions bounded on the sphere are

$$U(\mu) = P_l^s(\mu),\ l = s, s+1, s+2, \ldots, \quad c_{0,0} = \frac{\sqrt{l(l+1)}}{s}. \tag{2.19}$$

Note that (2.18) also has a negative root for $c_{0,0}$, since the eigenvalue is the 'square' of the phase speed; however, the reduction from three equations to one requires dividing by a factor that is 'zero' for $c_{0,0} = -\sqrt{s(s+1)}/s$. Thus, the 'anti-Kelvin' wave, as this spurious mode has been dubbed, is not a true eigenfunction of the original system of three linearized equations; the only true $l = s$ eigenmode has a positive (eastward) phase speed.

When $l = s$, the solution is the well-known linear Kelvin wave on the sphere:

$$c_{0,0} = \frac{\sqrt{s(s+1)}}{s}, \tag{2.20}$$

$$u_{0,1}(x,\mu) = P_s^s(\mu)\cos(sx), \tag{2.21}$$

$$v_{0,1}(x,\mu) = (1-\mu^2)\frac{\partial P_s^s(\mu)}{\partial \mu}\frac{\sin(sx)}{s}, \tag{2.22}$$

$$\phi_{0,1}(x,\mu) = c_{0,0}P_s^s(\mu)\cos(sx). \tag{2.23}$$

The higher order equations almost have the same form as (2.15), (2.16) and (2.17), except they have forcing terms from lower order solutions on the right-hand sides of these equations. After eliminating $\phi_{i,j}$ and $v_{i,j}$, there is always an inhomogeneous Legendre equation for $u_{i,j}$ with the inhomogeneous part containing $c_{i,j-1}$. By employing the Fredholm solvability condition, which requires the inner product of the eigenfuction and the inhomogeneous term be zero, we can get the eigenvalue $c_{i,j-1}$ and then get the particular solution for $u_{i,j}$. The homogenous solutions can always be absorbed into the lowest order solution because they have the same form as the lowest order solutions. A is a placeholder used to order the variables.

For the $s = 1$ case (i.e. a longitudinal period of 2π), the nonlinear Kelvin wave on the sphere is, to the lowest nonlinear order,

$$c = \sqrt{2} - \frac{1}{4}\delta + \frac{17\sqrt{2}}{320}\delta^2 - \frac{1}{160}\delta^3 + \frac{61\sqrt{2}}{40}\delta^2 A^2 - \frac{83}{40}\delta^3 A^2, \tag{2.24}$$

$$\begin{aligned} u = {} & \sqrt{1-\mu^2}\cos(x)A - \frac{\sqrt{2}}{2}(1+\mu^2)\sqrt{1-\mu^2}\cos(x)\delta A \\ & + \frac{27}{40}(1+\mu^2)\sqrt{1-\mu^2}\cos(x)\delta^2 A + 2\sqrt{2}(1-\mu^2)\cos(2x)\delta A^2, \end{aligned} \tag{2.25}$$

$$\begin{aligned} v = {} & -\mu\sqrt{1-\mu^2}\sin(x)A + \mu\sqrt{2(1-\mu^2)}\sin(x)\delta A \\ & + \frac{\mu}{20}\sqrt{1-\mu^2}(2\mu^2-29)\sin(x)\delta^2 A - 2\sqrt{2}\mu(1-\mu^2)\sin(2x)\delta A^2, \end{aligned} \tag{2.26}$$

$$\phi = \sqrt{2(1-\mu^2)}\cos(x)A - \frac{5}{4}\sqrt{1-\mu^2}\cos(x)\delta A$$
$$-\frac{\sqrt{2}}{320}(64\mu^2-273)\sqrt{1-\mu^2}\cos(x)\delta^2 A + \left(\frac{15}{4}(1-\mu^2)\cos(2x) - \frac{\mu^2}{4}\right)\delta A^2. \quad (2.27)$$

The extension to $s > 1$ is straightforward.

3. Kepler mapping/Galerkin method/Newton continuation method

In this section, we shall calculate the same travelling wave solution derived in the previous section but without the restriction to small wave amplitude and small Lamb's parameter. We shall combine a change of coordinate (Kepler mapping) with a spectral Galerkin method and Newton continuation.

3.1. *Kepler mapping*

To improve the Fourier rate of convergence from second order to fourth order in the degree of the coefficients, we transform the longitudinal coordinate from x to a new stretched coordinate z. The 'Kepler mapping', so named because inverting the transformation requires solving the Kepler equation of celestial mechanics, concentrates high resolution near the discontinuous corner at $x = 0$ while preserving the periodic behaviour in longitude (Boyd 2006). This creates a mild complication because we apply the numerical method to waves of different longitudinal periods $2\pi/s$, where s is the zonal wavenumber of the lowest nonzero longitudinal Fourier component. The form of Kepler mapping has a slightly different form for different s:

$$x = z - \frac{\sin(sz)}{s}, \quad (3.1)$$

$$\frac{\partial}{\partial x} \to \frac{1}{1-\cos(sz)}\frac{\partial}{\partial z}. \quad (3.2)$$

The mapping preserves spatial periodicity, and $x = z$ at $z = n\pi/s$, $n = 0, \pm 1, \pm 2, \ldots$.

3.2. *Galerkin method*

The Galerkin method for discretizing a differential equation demands that when the truncated Fourier series for each unknown is substituted into the shallow water equations to obtain the so-called residual function, the leading terms of the Fourier series of the residual be zero. These constraints are obtained by evaluating the integral inner product of the basis functions with the residual function and demanding that this integral be zero or, equivalently, that each basis function be orthogonal to the residual function. The number of orthogonality conditions is equal to the number of undetermined coefficients in the Fourier series for the unknowns, thus deriving a consistent set of nonlinear algebraic equations for the Fourier coefficients of u, v and ϕ. A full discussion is given in Boyd (2001).

To reduce the number of unknowns by a factor of four, we assume that u and ϕ are symmetric about the equator, and v is antisymmetric about the equator, with $z = 0$ in both cases. (Our success in computing solutions with the assumed symmetries is an *a posteriori* justification for these assumptions.) The domain of the entire globe is $[0, 2\pi]$ in longitude and $[0, \pi]$ in colatitude. But by employing the periodicity and parities of the Kelvin wave, we can reduce the domain by a factor of $2s$ in longitude and half it in latitude. So we only need to calculate the unknowns in the region of $[0, \pi/s]$ in x or z and $[0, \pi/2]$ in θ.

As explained in Boyd (1984), the nonlinear travelling wave eigenproblem has a unique solution only after specification of the longitudinally averaged zonal flow, in

this case, zero. By integrating geostrophic balance around a circle of latitude, one finds that a zero mean for u at each latitude implies that $\phi_{geostrophic}$ must have a zero mean and vice versa. Integrating the 'full' latitudinal momentum equation around a circle of latitude gives $g\,\partial\overline{h}/\partial\theta = \overline{u\partial v/\partial\lambda}/\sin(\theta) - (1/2)\partial\overline{v^2}/\partial\theta + \cot(\theta)\overline{u^2}$, with an overline denoting the zonal mean. As usual, however, such a cyclostrophically balanced flow is negligible Boyd (1976); in our computations, the mean height field $\overline{h}$ was always less than one part in 10 000 relative to the maximum of the wave.

Also by employing the parities of the unknowns, we can halve the number of basis functions used in both directions. In z, the basis functions of u and ϕ are

$$\psi^{(z)}_{sym,0} = 1, \quad \psi^{(z)}_{sym,1} = \cos(sz) + \tfrac{1}{2}, \tag{3.3}$$

$$\psi^{(z)}_{sym,m} = \cos(smz), \qquad m = 2, 3, \ldots, M, \tag{3.4}$$

where the constant $\psi^{(z)}_{sym,0}$ is used only for ϕ, and the basis functions of v are

$$\psi^{(z)}_{asym,m} = \sin(smz), \qquad m = 1, 2, \ldots, M. \tag{3.5}$$

The additive factor of (1/2) in $\psi^{(z)}_{sym,1}$ ensures that all basis functions with $m > 0$ 'individually' have a zero longitudinal mean, despite the change of coordinate from longitude to the Kepler coordinate s, so that the Kelvin wave is not accompanied by a zonal mean flow; note that $\int_0^{\pi} \cos(sz(x))\mathrm{d}x = -\pi/2$ (Boyd 2006).

In colatitude, the basis functions of u and ϕ are

$$\psi^{(\theta)}_{sym,n} = \cos((2n-2)\theta)\sin(\theta)^s, \qquad n = 1, 2, \ldots, N, \tag{3.6}$$

and the basis functions of v are

$$\psi^{(\theta)}_{asym,n} = \cos((2n-1)\theta)\sin(\theta)^s, \qquad n = 1, 2, \ldots, N. \tag{3.7}$$

The reason for the $\sin^s$ factor is explained in Orszag (1974), Boyd (1978*a*, 2001 Chapter 18, § 8); it suffices to say that the rate of convergence is greatly improved in spherical geometry by multiplying the cosines by $\sin(\theta)^s$.

Finally, we expand u, ϕ and v as

$$u = \sum_{m=1}^{M}\sum_{n=1}^{N} a^u_{mn}\psi^{(z)}_{sym,m}\,\psi^{(\theta)}_{sym,n}, \tag{3.8}$$

$$\phi = \sum_{m=0}^{M}\sum_{n=1}^{N} a^{\phi}_{mn}\psi^{(z)}_{sym,m}\psi^{(\theta)}_{sym,n}, \tag{3.9}$$

$$v = \sum_{m=1}^{M}\sum_{n=1}^{N} a^{v}_{mn}\psi^{(z)}_{asym,m}\psi^{(\theta)}_{asym,n}, \tag{3.10}$$

where a^u_{mn}, a^{ϕ}_{mn} and a^{ϕ}_{mn} are the coefficients u, ϕ and v, whose sizes are $M \times N$, $M \times (N+1)$ and $M \times N$.

The number of points used in z should be no fewer than the number of basis function used in z. So it requires $M_{pts} \geqslant M$. Similarly, $N_{pts} \geqslant N$.

To apply the Galerkin method, we first substitute the (truncated) Fourier series into the nonlinear shallow water equations. The resulting 'residual functions' are just the left-hand sides (LHSs) of (2.11)–(2.13). The residual functions depend on the spectral coefficients $\{a^u_{mn}, a^v_{mn}, a^{\phi}_{mn}\}$ and the phase speed c. We then demand that the residual be orthogonal to a set of test functions when integrated over the domain. In 'mean

weighted residual' methods, the test functions can be very general; Galerkin's method is the special case in which the test functions are the basis functions. Thus, for (2.11) and (2.13), the test functions are $\psi^{(z)}_{asym}$ in z and $\psi^{(\theta)}_{sym}$ in θ. For (2.12) the test functions are $\psi^{(z)}_{sym}$ in z, $\psi^{(\theta)}_{asym}$ in θ.

The Galerkin residual is

$$r_{1,mn} = \int\int \text{LHS}(2.11) \times \psi^{(z)}_{sym,m}(\theta) \times \psi^{(z)}_{asym,n}(z)\,\mathrm{d}z\,\mathrm{d}\theta, \tag{3.11}$$

$$r_{2,mn} = \int\int \text{LHS}(2.12) \times \psi^{(\theta)}_{asym,m}(\theta) \times \psi^{(z)}_{sym,n}(z)\,\mathrm{d}z\,\mathrm{d}\theta, \tag{3.12}$$

$$r_{3,mn} = \int\int \text{LHS}(2.13) \times \psi^{(\theta)}_{sym,m}(\theta) \times \psi^{(z)}_{asym,n}(z)\,\mathrm{d}z\,\mathrm{d}\theta, \tag{3.13}$$

where $m = 1, \ldots, M$, $n = 1, \ldots, N$. The sizes of $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ and $\boldsymbol{r}_3$ are $M \times N$, $M \times (N+1)$ and $M \times N$, so there are in total $3MN + M$ residual elements.

The integrals are numerically approximated using M_{pts} quadrature points in z:

$$z = \frac{(2i-1)}{2M_{pts}}\frac{\pi}{s}, \qquad i = 1, 2, \ldots, M_{pts}; \tag{3.14}$$

so $z \in (0, \pi/s)$, which is also the range of the unmapped coordinate x. We use N_{pts} points in colatitude:

$$\theta = \frac{(2j-1)}{4\,N_{pts}}\,\pi, \qquad j = 1, 2, \ldots, N_{pts}. \tag{3.15}$$

Because the integrands are periodic in both z and θ, the trapezoidal rule converges exponentially fast in M_{pts} and N_{pts} except for the corner wave. Because of parity symmetry, the range of the quadrature points may be halved, just like the number of basis functions, in each coordinate.

3.3. *Newton continuation method*

The $3MN + M + 1$ unknowns are the scalar c plus three vectors of lengths MN, $M(N+1)$ and MN: a^u, a^ϕ and a^v. However, there are only $3MN + M$ values of $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ and $\boldsymbol{r}_3$. So in addition, we require that the value of ϕ at $x = 0$ and latitude $= 0$, or ϕ_{00} for short, is fixed during the iteration. This gives us an additional residual r_4.

We can reshape the coefficient matrices a^u, a^ϕ and a^v to three vectors, stacking these and c together to form a $(3MN + M + 1) \times 1$ vector $\boldsymbol{a}_{(3MN+M+1)\times 1}$. Similarly, we can form a vector $\boldsymbol{r}_{(3MN+M+1)\times 1}$ from $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ and $\boldsymbol{r}_3$ and r_4.

The algebraic system $\boldsymbol{r}(\boldsymbol{a}) = \boldsymbol{0}$ is solved by a Newton/continuation method. The Newton iteration is to iterate the following until the difference between successive iterates is below a (tiny) user-chosen error tolerance:

$$\boldsymbol{a}^{(n+1)} = \boldsymbol{a}^{(n)} - \boldsymbol{J}^{-1}\big(r^{(n)}\big), \tag{3.16}$$

where $\boldsymbol{J}$ is the Jacobian matrix

$$J_{ij} = \frac{\partial r_i}{\partial a_j}, \qquad i = 1, \ldots, 3MN + M + 1,\ \ j = 1, \ldots, 3MN + M + 1, \tag{3.17}$$

and n is the iterative number. Unfortunately, all iterative methods require a 'first guess' or 'initialization'.

Parameter continuation provides the required first guess. To trace a complete branch of solutions, we march from small amplitude (where the initialization is provided by perturbation theory as in the previous section) to large amplitude while keeping all

ϵ	0.01	0.1	0.5	1	3	5	15	30
$s=1$	0.0007	0.002	0.0005	0.0008	0.007	0.001	0.004	0.009
$s=2$	0.001	0.002	0.008	0.009	0.004	0.003	0.003	0.007

TABLE 2. Relative coarse–fine differences in the L_∞ norm.

other parameters fixed. We chose ϕ_{00}, which is the equatorial height at the crest of the wave, $\phi(x=0, \theta=\pi/2)$, as the amplitude parameter (though other choices are possible). The continuation strategy is to march in small steps of the amplitude parameter. The computed solution for the kth value of ϕ_{00} is used as the initialization for Newton's iteration to compute the Fourier coefficients $\boldsymbol{a}$ for the $(k+1)$st value of ϕ_{00}.

As the amplitude of Kelvin wave increases, it evolves to the corner wave, so called because it has a slope discontinuity at the crest (Boyd 2005*a*). The branch of travelling waves ends abruptly at the corner wave: there are no solutions for larger amplitude. (Instead, all waves larger than the corner wave break.) The corner wave is a sort of an anti-bifurcation point in the sense that no additional branches are born at the corner wave, but rather the branch simply dies. So when the amplitude is a little greater than that of the corner wave, Newton's iteration fails.

3.4. *Resolution check*

Two different resolutions were used as a self-consistency check. The coarse solutions used $M=20$ basis functions in longitude, $N=10$ in latitude with $M_{pts}=30$ (longitude) and $N_{pts}=15$ (latitude). All high resolution computations were set at $M=30$, $N=15$, $M_{pts}=40$ and $N_{pts}=20$. The difference between these two results for the corner waves are given in table 2, which catalogues the L_∞ norm of the difference between the coarse and fine grid approximations, divided by ϕ_{00}, for various ϵ and s. We see that coarse resolution gives very decent results: all the coarse–fine relative differences are less than 0.9 %.

4. Spatial structure of the corner wave

We computed steadily propagating Kelvin waves of $s=1$ (longitudinal period of 2π) and $s=2$ (longitudinal period π) for various values of Lamb's parameter ϵ, using the numerical methods described in the previous section.

Figure 1 shows the normalized absolute values of Fourier coefficients of ϕ in the corner wave limit. The rate of convergence is about $O(K^{-4})$, where $K=\sqrt{m^2+n^2}$ is the total degree of the Fourier basis function indexed by m and n. When the amplitude is less than its corner wave limit, the rate of convergence is exponential. At the corner wave limit, the rate slows to $O(K^{-4})$.

It would be helpful if the solution of the discretized partial differential system terminated at the corner wave, but this is not true. The algebraic system has a solution even 'beyond' the corner wave limit because the branching of the solution in a 'finite' dimensional system of polynomial equations cannot simply stop. Instead, when ϕ_{00} is beyond the corner wave limit, the Fourier coefficients cease to converge, so that the finite-dimensional Galerkin-discretized polynomial system no longer yields a good approximation of the solution of the differential equation, which is a system of infinite dimension. The subtleties of identifying the corner wave are discussed in detail in Boyd (2006). However, we will discuss graphical clues to spurious solutions next.

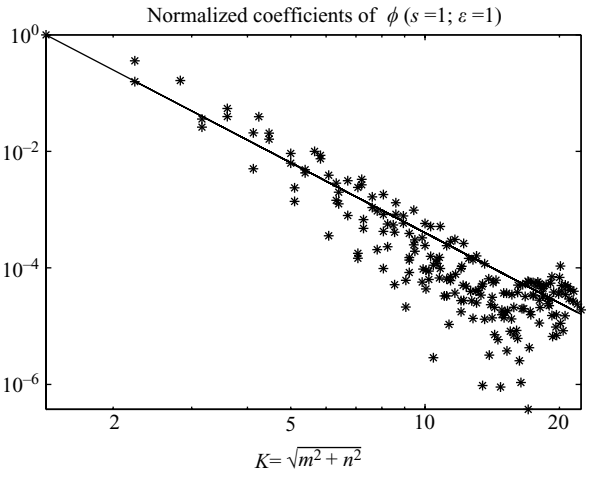

FIGURE 1. Normalized absolute values of Fourier coefficients of ϕ, which is the solution of the corner wave of $s=1$ and $\epsilon=1$ case; m and n are coefficient degree in longitudinal and latitudinal directions respectively. The solid line has a slope proportional to K^{-4}, confirming the predicted fourth-order rate of convergence when the Kepler change-of-coordinate is used. The rate of convergence is good enough to show that the solution of the corner wave is reliable.

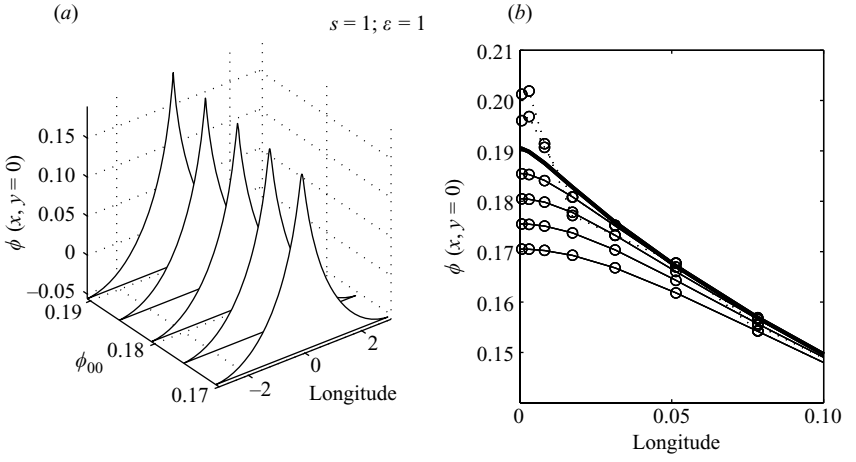

FIGURE 2. Travelling Kelvin wave solutions with $s=1$ and $\epsilon=1$. (a) Equatorial section of ϕ for $\phi_{00}=0.1705, 0.1755, 0.1805, 0.1855$ and 0.1905, respectively; $\phi(x,\theta=\pi/2)$ steepens with the increasing ϕ_{00}. When $\phi_{00}=0.1905$, $\phi(x,\pi/2)$ is the corner wave, discontinuous in its first derivative at the crest. (b) A zoom in plot of $\phi(x,\pi/2)$ with $\phi_{00}=0.1755$, $0.1805, 0.1855, 0.1905, 0.1955, 0.2005$. The heavy curve is for $\phi_{00}=0.1905$. Note that this graph includes two values of ϕ_{00} larger than that of the corner wave (dashed); these are unphysical as indicated by their unphysical oscillations near $x=0$. The interval in longitude is from 0 to 0.1, which is about 1.6% of the total width. (Note that the plot is in the physical longitudinal coordinate x; the circles on each curve show the points of the grid, which is evenly spaced in z but very heavily concentrated in x near $x=0$.) This graph shows that the corner wave is easily distinguished by the eye from near-corner waves in a zoom plot.

Figure 2 shows travelling Kelvin wave at the equator for $s=1$ and $\epsilon=1$ of different amplitudes ϕ_{00}. Figure 2(b) is a zoom plot of the same five solutions illustrated in figure 2(a). All five solutions are very close to the corner wave and seem to rise to

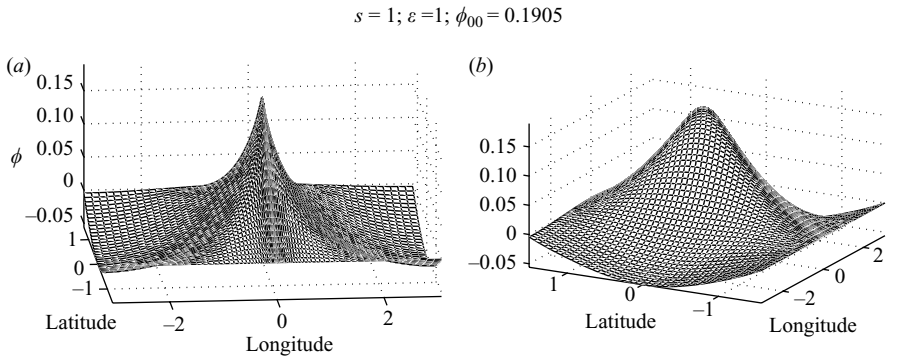

FIGURE 3. For $s = 1$ and $\epsilon = 1$, ϕ of the corner wave; (a) and (b) differ only in viewing angle. The peak value of ϕ_{00} is 0.1905. The comparison shows that only the longitudinal derivative is discontinuous at the peak.

a corner at $x = 0$ in figure 2(a). Graphical magnification by plotting the variable on a small x-interval shows that the slope discontinuities are only optical illusions for four of the five: the lower four solid curves have narrow regions of high curvature very close to the origin at which these waves flatten out to zero slope at $x = 0$. In contrast, when $\phi_{00} = 0.1905$, the corner wave is almost linear all the way to $x = 0$. Thus, the corner wave is easily distinguished by the eye from near-corner waves with a zoom plot (Boyd 2003). The two dashed curves are physically spurious solutions whose amplitudes are larger than that of the corner wave; these are genuine solutions of the system of polynomial equations generated by the Galerkin discretization, but these are not approximations of solutions of the shallow water equations.

Surface mesh plots ϕ for the corner wave limit are shown in figure 3. There is a question: are both components of the gradient of ϕ discontinuous at the peak – or only one? We visually answered this question by plotting ϕ twice from different viewing angles. Figure 3(a) shows that the longitudinal derivative is (at least visually) discontinuous. However, rotating the viewing angle by roughly a quarter turn shows only a smooth, rounded crest: the north–south derivative shows no signs of discontinuity.

Figure 4 displays line graphs that, for two different values of ϵ, make the same point. In each, the solid curve is a longitudinal cross-section at the equator, while the dashed curve shows $\phi(0, y)$. The x-derivative is discontinuous, but the latitudinal derivative is smooth.

Just as for infinitesimal amplitude Kelvin waves, u of the corner wave is graphically indistinguishable from ϕ and so is not plotted. The first derivative of the northward velocity v is everywhere continuous; so v is not plotted.

Figure 5 compares $\phi(x, y = 0)$, normalized by dividing by ϕ_{00}, for many different values of ϵ. As ϵ increases, the corner wave becomes increasingly narrow in longitude. This trend is also evident by comparing figure 4(a) and figure 4(b). Dispersion and the height of the corner wave both diminish rapidly as ϵ increases; it is remarkable that the corner wave becomes narrower, more focused in longitude, in this same limit. The latitudinal width, not shown, becomes increasingly narrow as captured by the equatorial beta-plane approximation, $\phi(x, \theta) \sim A(x) \exp(-\sqrt{\epsilon}(\theta - \pi/2)^2)$. However, the latitudinal width is controlled by 'linear' dynamics, whereas the longitudinal focusing is caused entirely by 'nonlinearity': when the amplitude is much smaller

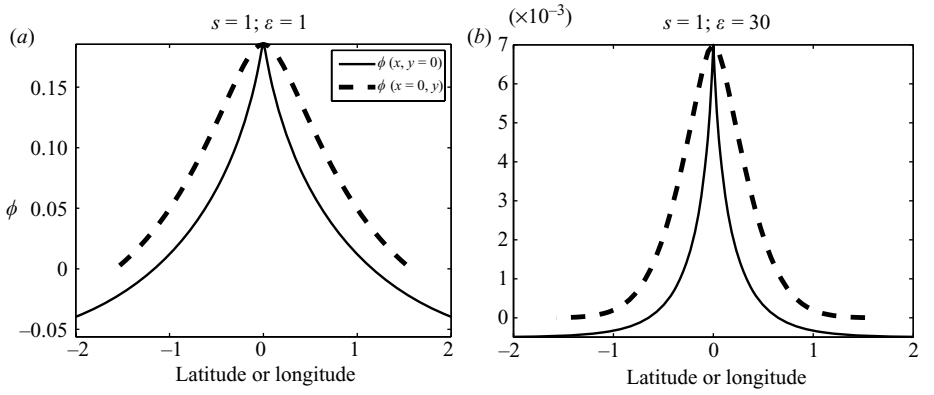

FIGURE 4. Profile of ϕ along the equator (solid) and the profile of ϕ at $x = 0$ as a function of latitude (dashed) for the corner wave for $s = 1$ for two different values of Lamb's parameter: (*a*) $\epsilon = 1$; (*b*) $\epsilon = 30$. The horizontal axis does double duty, being longitude for $\phi(x, \text{latitude} = 0)$ and latitude for $\phi(x = 0, \text{latitude})$. In both panels, the longitudinal derivative (solid) is clearly discontinuous at the crest, whereas the north–south derivative shows not the slightest hint of non-smoothness.

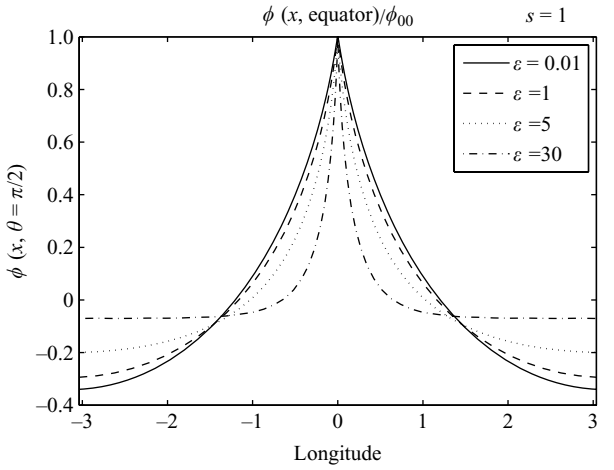

FIGURE 5. Normalized $\phi(x, \theta = \pi/2)$ (longitudinal section at the equator) of the corner wave solution of $s = 1$ and $\epsilon = 0.01, 1, 5, 30$. The profiles of $\phi(x, \theta = \pi/2)$ are scaled by their corresponding maxima, ϕ_{00}. The shape of ϕ at the equator becomes narrower and narrower as ϵ increases.

than the corner wave, the longitudinal structure of the Kelvin mode is approximately $\cos(s\lambda)$.

Another interesting question is how far the slope discontinuity extends from the equator to the poles. To answer this question, we calculated $\mathrm{d}\phi/\mathrm{d}x$. Figure 6 shows ϕ_x at several latitudes, shown on the full longitudinal range in (*a*) and as a zoom plot in (*b*). A finite spectral series must always impose a truncation-dependent smoothing on a discontinuity. Even so, it is clear the slope rapidly diminishes away from the equator. It seems likely that the Kelvin wave is discontinuous only at the equator.

The graphs for $s = 2$ were so similar to those for $s = 1$ that they have been omitted but will appear in Zhou's forthcoming thesis. However, the maximum equatorial

ϵ	0.01	0.1	0.5	1	3	5	15	30
Phase speed c	1.4327	1.3792	1.3008	1.2551	1.1758	1.141	1.082	1.0572
ϕ_{00}	3.5	0.95	0.32	0.19	0.074	0.045	0.0145	0.0071
$(c^2-1)/(3\sqrt{\epsilon})$	3.5088	0.9510	0.3262	0.1918	0.0736	0.0450	0.0147	0.0072
$h_{00} \equiv \phi_{00}\sqrt{\epsilon}$	0.3500	0.3004	0.2263	0.1900	0.1282	0.1006	0.0562	0.0389

TABLE 3. Parameters in the corner wave limit for the $s = 1$ case.

ϵ	0.01	0.1	0.5	1	3	5	15	30
Phase speed	1.2349	1.2169	1.1883	1.1698	1.1327	1.1135	1.0739	1.0537
ϕ_{00}	1.75	0.5	0.192	0.12	0.053	0.036	0.0129	0.0067
$(c^2-1)/(3\sqrt{\epsilon})$	1.7499	0.5069	0.1942	0.1228	0.0545	0.0358	0.0132	0.0067
$h_{00} \equiv \phi_{00}\sqrt{\epsilon}$	0.1750	0.1581	0.1358	0.1200	0.0918	0.0805	0.0500	0.0367

TABLE 4. Parameters in the corner wave limit for the $s = 2$ case.

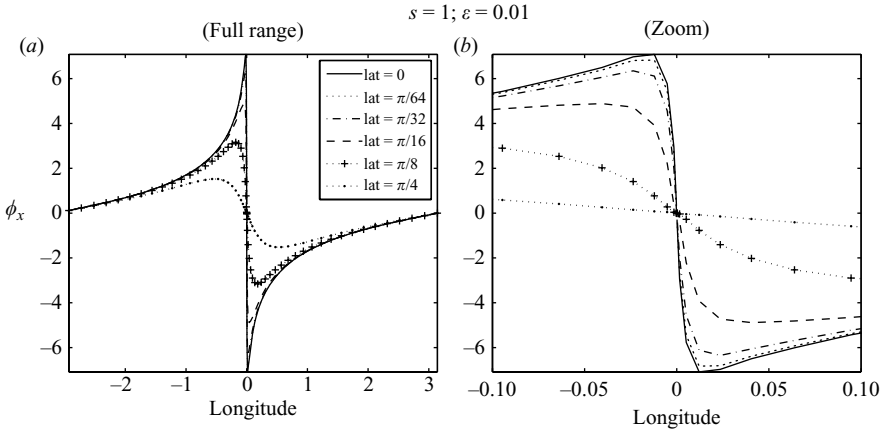

FIGURE 6. The derivative of ϕ of the corner wave solution of the $s = 1$ and $\epsilon = 0.01$ case with respect to the longitude x: (*a*) ϕ_x at latitudes $0, \pi/64, \pi/32, \pi/16, \pi/8, \pi/4$ plotted on the full global domain; (*b*) same as (*a*) but a zoom plot with a much smaller range.

height $\phi_{00}(\epsilon)$ and phase speed $c(\epsilon)$ for the corner wave are discussed for both $s = 1$ and $s = 2$ in the next section.

5. Variations of phase speed and corner height

The parameters of the corner wave for different ϵ are summarized in table 3 ($s = 1$) and table 4 ($s = 2$). From the tables,we can see both ϕ_{00} and phase speed c decrease as ϵ increases. This is as expected: because the dispersion due to the Earth's sphericity decreases rapidly with ϵ (as known from Longuet–Higgins's large ϵ asymptotic expansion of the linear phase speed), it is plausible that nonlinearity will overwhelm dispersion, giving breaking instead of travelling waves, at increasingly lower values of the wave amplitude ϕ_{00} as $\epsilon \to \infty$.

The tables also list the quantity

$$h_{00} \equiv \sqrt{\epsilon}\,\phi_{00}. \tag{5.1}$$

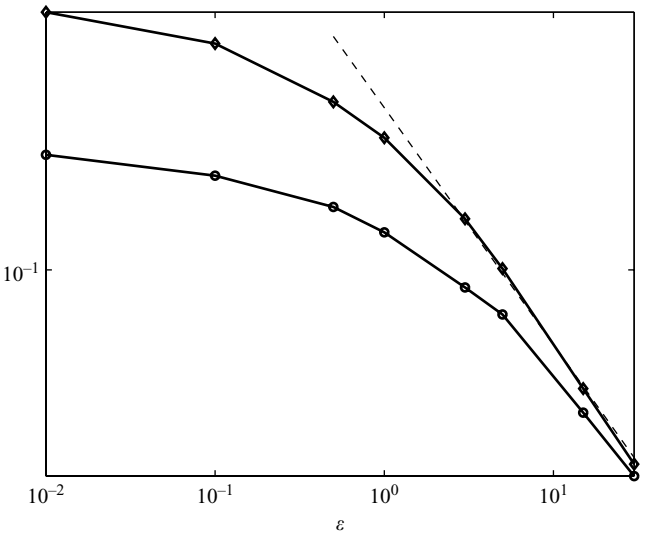

FIGURE 7. Representation of h_{00} of the corner waves of different ϵ for $s=1$ (upper thick curve with diamonds) and $s=2$ (lower thick curve with circles). The dashed line on this log–log plot shows that h_{00} decays asymptotically proportional to $1/\sqrt{\epsilon}$ as $\epsilon \to \infty$.

This gives the maximum perturbative height of the Kelvin corner wave relative to the mean depth H; that is the maximum perturbative height is $h_{00}\,H$ in meters. We have listed this quantity because it decreases more slowly with increase in ϵ than does ϕ_{00}.

Figure 7 compares the equatorial height of the corner wave versus ϵ for both $s=1$ and $s=2$. The results are very similar for the two wavenumbers. As ϵ increases, the dispersion due to the sphericity of the Earth decays very rapidly. Consequently, the height h_{00} of the corner wave diminishes very rapidly, too. On a log–log plot, a power law asymptotes to a straight line; the dashed line here suggests that $h_{00} \approx 0.2/\sqrt{\epsilon}$ for both wavenumbers. The graph suggests that the maximum height of the corner wave is independent of zonal wavenumber s in the equatorial beta-plane limit that $\epsilon \to \infty$.

The tables also show an interesting empirical relationship between the phase speed and the maximum height of the corner wave:

$$\phi_{00} \approx \frac{c^2-1}{3\sqrt{\epsilon}}. \tag{5.2}$$

By matching discontinuities in the x-derivatives of u and ϕ, we can derive the diagnostic relationship $(c-u_{00}\sqrt{\epsilon})^2 = 1 + h_{00}$ at the crest of the corner wave. (We thank a reviewer for suggesting this. Unfortunately, it is not possible to extend this further: the rest of our study is based on perturbation series and computations).

6. Limitations of theory

Our paper is a comprehensive study of the full parameter space for flow without mean currents, clouds or vertical propagation, aimed at illuminating the reality that Kelvin waves sometimes break and sometimes don't and have something or other as a boundary between these two regimes as seen in numerical models that include mean flow as well. We have resisted making comparisons between our results and observations because of the limitations of our theory. Equatorially trapped Kelvin waves are strongly affected by the strong mean currents in the tropical ocean: the

alternating jets known as the South Equatorial Current, North Equatorial Current, North Equatorial Counter-Current and Equatorial Undercurrent. We have shown here that the dispersion created by the Earth's sphericity increasingly weakens as ϵ increases. For the large values of ϵ relevant to the tropical ocean, as catalogued in table 1, spherical effects will be overwhelmed by the much stronger dispersion induced by these alternating jets. A goal of future work is to therefore incorporate mean flow into models of travelling Kelvin waves.

This is hard within the shallow water model because of critical latitude effects (Boyd & Christidis 1982; Natarov & Boyd 2001), which are known to weakly destabilize the Kelvin waves. The Equatorial Undercurrent, which has strong vertical shear and (as its name implies) does not extend to the surface, can only be incorporated into a three-space-dimensional model, a task orders of magnitude more difficult than the two-dimensional travelling wave models considered here.

Mean flow, though not negligible in the troposphere, is relatively less important to atmospheric Kelvin waves than the oceanic because the rather small values of ϵ relevant to the atmosphere imply much stronger dispersion due to the Earth's sphericity. However, tropospheric Kelvin waves are both excited by cumulus convection and simultaneously organize and transport this convection. Parameterizing cumulus convection in general circulation models is still a major research frontier.

Another difficulty in atmospheric theory is that the air is not a vertically confined layer like the sea but rather a semi-infinite layer with free vertical propagation. This is a serious technical complication because as the atmospheric density decreases with height, the amplitude of vertically propagating waves must increase. Eventually, all inviscid, vertically propagating waves must break. The breaking of tropospheric gravity waves is a major source of damping for larger scale motions in the middle atmosphere. Kelvin waves are also dissipated by rather strong radiative damping in the mesosphere. Thus, there is competition between growth (due to decreasing density) and decay (due to damping). Water waves experience a similar energy growth when propagating into water of decreasing depth. (Similarly, ocean equatorial Kelvin waves propagate through variable depth because of the deepening of the main thermocline to the west in the Pacific Ocean (Long & Chang 1990). Grimshaw (1970, 1971, 1979) showed that the soliton slowly adjusts to the changing depth while shedding a shelf to conserve total mass and energy. Thirty-seven years later, he and his collaborators have progressed to a model with both a gradual slope and dissipation (El, Grimshaw & Kamchatnov 2007); however, a two-dimensional model without Coriolis force is much simpler than the shallow water equations on a rotating sphere, and parameterized gravity wave breaking and radiative dissipation are more complicated than the Chezy bottom friction of their paper.

Mean currents strongly modulate the propagation of Kelvin waves. Indeed, this is the heart of the Lindzen–Holton theory of the QBO in the tropical lower stratosphere (Andrews *et al.* 1987): Kelvin and Yanai waves in turn force the mean flow to reverse sign, but the QBO strongly modulates the upward penetration of both wave species, so that they switch roles with each QBO cycle. Existing theories use parameterizations of wave breaking and other drastic simplifications; the length scales are so short that it is only recently that general circulation models (GCMs), metaphorically the supercomputer-hogging aircraft carriers of climate research, have been able to even crudely capture the QBO. Kelvin waves also are a major driving force in the semi-annual wind oscillation in the tropical upper stratosphere and are in turn controlled by the ever-changing mean current (Boyd 1978*b*, *c*). There is thus a large gap between the idealizations of our computations and the real ocean and atmosphere, but this is

not uncommon in geophysical fluid dynamics (e.g. Lorenz and Krishnamurthy 1987; Boyd 1994, 1995).

7. Summary and conclusions

The computations confirm the results of simplified models and equatorial beta-plane computations: the travelling waves of the Kelvin mode terminate in a corner wave of finite height. The amplitude of the corner wave diminishes very rapidly with ϵ when the mean flow is neglected. In the real ocean or atmosphere, our results for large ϵ are quantitatively suspect because the very weak dispersion due to spherical geometry would likely be overwhelmed by the stronger dispersion due to the mean zonal currents.

As ϵ increases, the longitudinal profile of the corner wave becomes very narrow, whereas the corner waves for small ϵ span the whole equator.

In two space dimensions, slope discontinuities may take the form of a 'cone' (with discontinuities in both x and y derivatives at the peak), a 'crease' with a 'curve' or 'line' of discontinuous slope, extending away from the equator into both hemispheres or a 'point' singularity in which only one derivative is discontinuous and that too only at a single point. All previous studies of corner waves have been limited to one horizontal dimension and therefore furnish no guidance. Although it is impossible to prove theorems through inexact numerical computations, our graphs strongly suggest that the third possibility is true of the Kelvin corner wave: the height and velocity fields are singular only at the peak and only through a discontinuity in the direction of propagation, longitude.

Although we performed detailed computations only for zonal wavenumbers $s = 1$ and $s = 2$, there was so little qualitative difference that it appears that these conclusions are independent of zonal wavenumber s at least for small s. As illustrated in Boyd & Zhou (2008), Kelvin waves of moderate and large s are 'equatorially trapped.' Therefore, short Kelvin waves are well described by the equatorial beta-plane theory and computations in Boyd (1998, 2006).

Our computations cannot exclude the possibility that there may be nonlinear Kelvin branches which are not continguous with small-amplitude, linear Kelvin waves. This is not a difficulty peculiar to Kelvin waves but rather is a generic worry when computing the roots of any system of nonlinear algebraic or transcendental equations, whether resulting from the discretization of travelling waves or not; the peril of the 'missed solution branch' is ubiquitous. However, no such additional branches have been detected in numerous initial-value experiments: all Kelvin modes bigger than the corresponding corner wave break.

This work was supported by the National Science Foundation through grants OCE OCE 0451951 and ATM 0723440. We thank the four reviewers for their very detailed comments.

REFERENCES

Andrews, D. G., Holton, J. R. & Leovy, C. B. 1987 *Middle Atmospheric Dynamics*. Advances in Geophysics, Vol. 40. Academic.

Boyd, J. P. 1976 The noninteraction of waves with the zonally averaged flow on a spherical earth and the interrelationships of eddy fluxes of heat, energy, and momentum. *J. Atmos. Sci.* **33**, 2285–2291.

Boyd, J. P. 1978*a* The choice of spectral functions on a sphere for boundary and eigenvalue problems: a comparison of Chebyshev, Fourier and associated Legendre expansions. *Mon. Weath. Rev.* **106**, 1184–1191.

Boyd, J. P. 1978*b* The effects of latitudinal shear on equatorial waves. Part I. theory and methods. *J. Atmos. Sci.* **35**, 2236–2258.

Boyd, J. P. 1978*c* The effects of latitudinal shear on equatorial waves. Part II. applications to the atmosphere. *J. Atmos. Sci.* **35**, 2259–2267.

Boyd, J. P. 1980 The nonlinear equatorial Kelvin wave. *J. Phys. Oceanogr.* **10**, 1–11.

Boyd, J. P. 1984 Equatorial solitary waves. Part IV. Kelvin solitons in a shear flow. *Dyn. Atmos. Oceans* **8**, 173–184.

Boyd, J. P. 1991 Nonlinear equatorial waves. In *Nonlinear Topics of Ocean Physics: Fermi Summer School, Course LIX* (ed. A. R. Osborne), pp. 51–97. North-Holland.

Boyd, J. P. 1994 The slow manifold of a five mode model. *J. Atmos. Sci.* **51**, 1057–1064.

Boyd, J. P. 1995 Eight definitions of the slow manifold: seiches, pseudoseiches and exponential smallness. *Dyn. Atmos. Oceans* **22**, 49–75.

Boyd, J. P. 1998 High order models for the nonlinear shallow water wave equations on the equatorial beta-plane with application to Kelvin wave frontogenesis. *Dyn. Atmos. Oceans* **28** (2), 69–91.

Boyd, J. P. 2001 *Chebyshev and Fourier Spectral Methods*. Dover.

Boyd, J. P. 2003 A Legendre pseudospectral method for computing travelling waves with corners (slope discontinuities) in one space dimension with application to Whitham's equation family. *J. Comput. Phys.* **189** (5), 98–110.

Boyd, J. P. 2005*a* The Cnoidal Wave/Corner Wave/Breaking Wave Scenario: a one-sided infinite-dimension bifurcation. *Math. Comput. Sim.* **69** (3–4), 235–242.

Boyd, J. P. 2005*b* Near-corner waves of the Camassa–Holm equation. *Phys. Lett.* A **336**, 342–348.

Boyd, J. P. 2005*c* Ostrovsky and Hunter's generic wave equation for weakly dispersive short waves: analytical and numerical study of the parabolic and paraboloidal travelling waves (corner and near-corner waves). *Eur. J. Appl. Maths* **16**(1), 65–81.

Boyd, J. P. 2005*d* The short-wave limit of linear equatorial Kelvin waves in a shear flow. *J. Phys. Oceanogr.* **35**(6), 1138–1142.

Boyd, J. P. 2006 Fourier pseudospectral method with Kepler mapping for travelling waves with discontinuous slope: application to corner waves of the Ostrovsky–Hunter equation and equatorial Kelvin waves in the four-mode approximation. *Appl. Math. Comput.* **177**(1), 289–299.

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

Boyd, J. P. & Zhou, C. 2008 Uniform asymptotics for the linear Kelvin wave in spherical geometry. *J. Atmos. Sci.* **65** (2), 655–660.

Chapman, S. & Lindzen, R. S. 1970 *Atmospheric Tides*. D. Reidel.

Chen, G. Y. & Boyd, J. P. 2002 Nonlinear wave packets of equatorial Kelvin waves. *Geophys. Astrophys. Fluid Dyn.* **96**(5), 357–379.

El, G. A., Grimshaw, R. H. J. & Kamchatnov, A. M. 2007 Evolution of solitary waves and undular bores in shallow-water flows over a gradual slope with bottom friction. *J. Fluid Mech.* **585**, 213–244.

Fedorov, A. V. & Melville, W. K. 2000 Kelvin fronts on the equatorial thermocline. *J. Phys. Oceanogr.* **30**(7), 1692–1705.

Gill, A. E. 1982 *Atmosphere–Ocean Dynamics*. International Geophysics, Vol. 30. Academic.

Greatbatch, R. J. 1985 Kelvin wave fronts, Rossby solitary waves and nonlinear spinup of the equatorial oceans. *J. Geophys. Res.* **90**, 9097–9107.

Grimshaw, R. 1970 Solitary wave in water of variable depth. *J. Fluid Mech.* **42**, 639–.

Grimshaw, R. 1971 Solitary wave in water of variable depth. Part 2. *J. Fluid Mech.* **46**, 611–.

Grimshaw, R. 1979 Slowly varying solitary waves. Part 1. Korteweg-deVries. *Proc. R. Soc. Lond.* A **368**(1734), 359–375.

Grimshaw, R. H. J., Ostrovsky, L. A., Shrira, V. I. & Stepanyants, Y. A. 1998 Long nonlinear surface and internal gravity waves in a rotating ocean. *Surv. Geophys.* **19**, 289–338.

Le Sommer, J., Reznik, G. M. & Zeitlin, V. 2004 Nonlinear geostrophic adjustment of long-wave disturbances in the shallow-water model on the equatorial beta-plane. *J. Fluid Mech.* **515**, 135–170.

Lindzen, R. S. 1970 The application of terrestial atmospheric tidal theory to Venus and Mars. *J. Atmos. Sci.* **27** (4), 536–549.

Long, B. & Chang, P. 1990 Propagation of an equatorial Kelvin wave in a varying thermocline. *J. Phys. Oceanogr.* **20**, 1826–1841.

Longuet-Higgins, M. S. 1968 The eigenfunctions of Laplace's tidal equation over a sphere. *Phil. Trans. R. Soc. Lond.* A **262**, 511–607.

Longuet-Higgins, M. S. & Fox, M. J. H. 1996 Asymptotic theory for the almost-highest wave. *J. Fluid Mech.* **317**, 1–19.

Lorenz, E. N. & Krishnamurthy, V. 1987 On the nonexistence of a slow manifold. *J. Atmos. Sci.* **44**, 2940–2950.

Majda, A. 2003 *Introduction to PDEs and Waves for the Atmosphere and Ocean*. Courant Lecture Notes, Vol. 9. American Mathematical Society.

Majda, A. J., Rosales, R. R., Tabak, E. G. & Turner, C. V. 1999 Interaction of large-scale equatorial waves and dispersion of Kelvin waves through topographic resonances. *J. Atmos. Sci.* **56**(24), 4118–4133.

Marshall, H. G. & Boyd, J. P. 1987 Solitons in a continuously stratified equatorial ocean. *J. Phys. Oceanogr.* **17**, 1016–1031.

Milewski, P. A. & Tabak, E. G. 1999 A reduced model for nonlinear dispersive waves in a rotating environment. *Geophys. Astrophys. Fluid Dyn.* **90**, 139–159.

Moore, D. W. & Philander, S. G. H. 1977 Modelling of the tropical oceanic circulation. In *The Sea* (ed. E. D. Goldberg), vol. 6, pp. 319–361. Wiley.

Natarov, A. & Boyd, J. P. 2001 Beyond-all-orders instability in the equatorial Kelvin wave. *Dyn. Atmos. Oceans* **33**(3), 181–200.

Orszag, S. A. 1974 Fourier series on spheres. *Mon. Weath. Rev.* **102**, 56–75.

Ostrovsky, L. A. 1978 Nonlinear internal waves in a rotation ocean. *Oceanology* **18**, 181–191.

Pedlosky, J. 1987 *Geophysical Fluid Dynamics*. 2nd ed. Springer.

Ripa, P. 1982 Nonlinear wave–wave interactions in a one-layer reduced-gravity model on the equatorial beta-plane. *J. Phys. Oceanogr.* **12**, 97–111.

Ripa, P. 1985 Nonlinear effects in the propagation of Kelvin pulses across the Pacific Ocean. In *Advances in Nonlinear Waves* (ed. L. Debnath), pp. 43–56. Pitman.

Shrira, V. I. 1981 Propagation of long nonlinear waves in a layer of rotating fluid. *Ivestiya Atm. Oceanic Phys.* **17**(1), 76–81.

Shrira, V. I. 1986 On long strongly nonlinear waves in a rotating ocean. *Ivestiya Atm. Oceanic Phys.* **22**(4), 298–305.

Stokes, G. G. 1847 On the theory of oscillatory waves. *Camb. Trans.* **8**, 441–473.

Williams, G. P. 1996 Jovian dynamics. Part I. Vortex stability, structure and genesis. *J. Atmos. Sci.* **53** (18), 2685–2734.

Williams, G. P. & Wilson, R. J. 1988 The stability and genesis of Rossby vortices. *J. Atmos. Sci.* **45**, 207–249.

Zurek, R. W. 1976 Diurnal tide in Martian atmosphere. *J. Atmos. Sci.* **33** (2), 321–337.