Shallow water waves on the rotating sphere

D. Müller and J. J. O'Brien
Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, Florida 32306
(Received 8 September 1994)

Analytical solutions of the linear wave equation of shallow waters on the rotating, spherical surface have been found for standing as well as for inertial waves. The spheroidal wave operator is shown to play a central role in this wave equation. Prolate spheroidal angular functions capture the latitude-longitude anisotropy, the east-west asymmetry, and the Yoshida inhomogeneity of wave propagation on the rotating spherical surface. The identification of the fundamental role of the spheroidal wave operator permits an overview over the system's wave number space. Three regimes can be distinguished. High-frequency gravity waves do not experience the Yoshida waveguide and exhibit Coriolis-induced east-west asymmetry. A second, low-frequency regime is exclusively populated by Rossby waves. The familiar β-plane regime provides the appropriate approximation for equatorial, baroclinic, low-frequency waves. Gravity waves on the β-plane exhibit an amplified Coriolis-induced east-west asymmetry. Validity and limitations of approximate dispersion relations are directly tested against numerically calculated solutions of the full eigenvalue problem.

PACS number(s): 47.32.$-$y, 47.35.+i, 92.60.Dj, 92.10.Hm

I. INTRODUCTION

The first theoretical investigation of the large-scale response of the ocean-atmosphere system to external tidal forces was conducted by Laplace [1] in terms of the linearized Euler equations on the rotating sphere. Since then, these equations are known as "the tidal equations," although their significance widely exceeds the tidal problem: the free, linearized Euler equations on the rotating sphere govern the basic small-amplitude dynamics of the shallow fluid envelope of a rotating planet. As such they pose the fundamental linear problem of planetary circulation theory with applications to the atmosphere as well as to the ocean. These applications include short term, large-scale phenomena like the tides in the narrow sense, and long term, large-scale phenomena like the terrestrial climate problem and the circulation of the atmospheres of other planets. The generic dynamics of the overwhelming majority of contemporary numerical circulation models are represented by Laplace's tidal equations.

In contrast to the nonrotating case, Laplace observed that the special form of the rotation vector rendered the three-dimensional (3D) spherical problem nonseparable. With "the tidal equation" he established a 2D approximation, which is separable in spherical coordinates. In this framework the discussed what is now called the Lamb wave under the influence of external tidal forces. Laplace was not able to evaluate the eigensolutions of his equation. A century later, Margules [2] identified asymptotically two classes of solutions: gravity waves and the now called Rossby waves. An extensive investigation of these asymptotics by Hough [3] led to the notion of "Hough functions" for the eigensolutions of Laplace's tidal equations.

The most comprehensive discussion of the dynamics of the "tidal" problem to date has been given by Matsuno [4] in the framework of the β-plane approximation. In this Cartesian approximation of the spherical geometry near the equator Matsuno identified Rossby, mixed Rossby-gravity, as well as gravity waves, including the Kelvin wave. He moreover confirmed Yoshida's hypothesis [5] of an equatorial waveguide. A subsequent numerical study of the spherical problem by Longuet-Higgins [6] provided a quantitative representation of the eigenvalues and eigenfunctions of the tidal equation. These calculations demonstrated that Matsuno's theory was qualitatively complete. However, in spite of its elegance and fundamental role for the terrestrial climate problem, the β-plane approximation is not entirely free from ambiguities. In contrast to Laplace's tidal equation it does not provide a physically meaningful nonrotating limit and postulates, furthermore, an apparently paradoxical east-west asymmetry of gravity of waves. For given mode number and frequency a westward traveling gravity wave has a different phase speed than its eastward counterpart. Such an asymmetry suggests a Doppler-type effect due to a relative motion of medium and observer. However, the "tidal" equations represent a situation where both medium and observer are corotating and at rest relative to each other. In the framework of Doppler theory, these gravity waves should thus be symmetric.

Recently, exact, analytical solutions of the spherical tidal equation have been presented [7] for the special cases of standing waves and inertial waves. These results show that the eigenfunctions of the tidal equation are closely related to spheroidal wave functions and underscore the baroclinic, low-frequency nature of the β-plane approximation. At high frequencies, the β-plane approximation cannot appropriately accommodate the effects of rotation on the dispersion of gravity waves. In this regime, an alternative approximation of the dispersion relation applies. Presently, a complete analytical theory of Laplace's tidal equations is still not available.

The "tidal" equations occupy their fundamental position in planetary circulation theory essentially due to their consistency with respect to geometry, dynamics, and inertial forces, i.e., their covariance with respect to the 2D non-Euclidean geometry of the spherical surface. It is noteworthy that Laplace established his form of the equations some 60 years before Coriolis's work on Newton dynamics in rotating systems and—although spherical trigonometry was known at the time—some 75 years before Riemann's systematic discussion of non-Euclidean geometries. Laplace considered the diagnostics of a barotropic gas. Structurally the same diagnostics hold for shallow water theory in the vertically integrated ("transport") form. Linear, as well as nonlinear shallow water theories in transport form thus preserve the dynamically essential covariance of Laplace's tidal equations. In view of their role for the large-scale circulation, the "tidal" equations are thus primarily the linear shallow water equations on the rotating sphere. However, it has always been questioned whether shallow water theory can provide a sufficiently rich framework to capture the relevant features of the large-scale circulation.

Shallow water theory in its generic form has an unrealistically trivial vertical structure. But, interestingly enough, the incorporation of vertical structure, namely, baroclinicity, into circulation models has been essentially guided by the spirit of boundary layer theory with little regard for covariance requirements. The prototype of the resulting models is the set of the so called "primitive equations," constructed from three basic ingredients: (1) hydrostatics in the vertical, (2) generally a Boussinesq-type formulation of stratification, and (3) Laplace's Coriolis parameter for the representation of inertial forces. As a mixture of 3D and 2D features, the "primitive equations" are strictly noncovariant. The derivation of physically meaningful results from these equations requires a sophisticated set of scaling conditions [8] and in the numerical approach a high degree of sophistication in the selection of appropriate parameter ranges as well as in the evaluation of physically meaningful diagnostics from model results. Access to increasing computer time did not decrease these problems.

Covariant shallow water theory, on the other hand, is frequently utilized in numerical studies in the form of multilayer, reduced-gravity models [9]. This form eliminates fast barotropic waves through the coding-effective assumption of a quiescent abyss. This is known to be fairly unrealistic and to "increase" relevance reduced-gravity models had to resort to the equally artificial, but observationally more ambiguous, reference to an abyssal "level of no motion." However, with access to increasing computer time the relevance issue ceased to be one of abyssal dogma and became a matter of comparison to observations. It has now been demonstrated that covariant shallow water models successfully simulate observable large-scale features of the ocean circulation: in simple model versions on seasonal [10] and interannual [11] time scales and in more elaborate multilayer versions even on decadal time scales [12]. This success indicates that shallow water theory takes advantage of increased numerical resolution because its covariance captures the fundamental dynamical nature of the large-scale circulation. Vertical structure has to comply with covariance requirements and for multilayer versions this is the case. Shallow water theory with the covariant inclusion of vertical variability as (prognostic or diagnostic) internal variability of a spatially 2D fluid represents the dynamically appropriate framework for planetary circulation theory.

The lack of a complete analytical theory of the linear shallow water equations on the rotating sphere more than two centuries after their formulation provides a measure of their complexity. This is partly due to the necessarily curvilinear nature of the problem. Partial differential operators in curvilinear coordinates carry coordinate-dependent coefficients, which complicate the derivation of higher-order wave equations from the first order equations of motion. However, these problems are drastically simplified by utilizing tensor analysis in Riemann space, invoking in particular the concept of covariant differentiation. In contrast to the partial derivative, the covariant derivative "automatically" accounts for the coefficient functions so that the usual differentiation rules for products and other combinations of tensors apply to covariant differentiation. The manipulation of differential expressions then almost proceeds as in the case of Cartesian coordinates.

The additional feature to be taken into account is the non-Euclidean nature of the intrinsic geometry of the 2D spherical surface (it may be noted that the "3D spherical" geometry of the "primitive equations" with vertical coordinate $z=r-a$, where r is the radial coordinate of 3D Euclidean, spherical geometry and a the Earth's radius, is also non-Euclidean). For a non-Euclidean geometry the Riemann tensor no longer vanishes. This fourth order tensor is a functional of the metric, measuring the anticommutator of the covariant second order derivatives of a vector. As a consequence of a nontrivial Riemannian, the covariant second order derivatives of a vector no longer commute. This characteristic feature of tensor analysis in non-Euclidean Riemann space becomes relevant for the derivation of higher-order wave equations from the vectorial equations of motion. In the present paper different forms of the wave equation of shallow waters on the rotating sphere will be considered using tensor analysis. For these purposes index notation is quite effective with indices $m,n,\ldots$, running from 1 to 2, subscripts denoting covariant tensor components, superscripts contravariant components, and the semicolon for covariant differentiation. Essentials of tensor analysis can be found in a variety of textbooks [13,14] and the basic formulas for the 2D geometry of the spherical surface in geophysical coordinates (longitude λ, latitude φ) are given in [15] with additional details in the Appendix to this paper.

In this framework it has been shown [7] that the fundamental wave operator of shallow waters on the rotating sphere is the 2D spheroidal wave equation. The separation of this wave equation results in ordinary differential equations with periodic coefficients. Formally similar problems in other branches of fluid dynamics (stability of time periodic flow) or in electrodynamics (wave propagation in an elliptical waveguide) lead to the related

Mathieu equation. While Floquet theory provides a comprehensive analysis of Mathieu functions, a comparable tool for spheroidal wave function does not exist. The major problem stems from the irregular singularity of the spheroidal wave equation at infinity. This prevents the establishment of recurrence relations similar to those for functions of the hypergeometric type [16]. Special case solutions, as obtained in [7], can thus not be generalized by means of some algebra in wave number space.

The identification of the prolate spheroidal wave operator as a key element in the linear shallow water equations on the rotating spherical surface permits nevertheless a qualitative overview of the system's wave number space. The asymptotics of the spheroidal wave operator yield a set of approximations of the dispersion relation for different regimes in wave number space, which essentially link the Margules limits with the β plane. In the paper at hand, the validity and limitations of these approximations are directly tested against numerical computations of the eigenfrequencies of rotating, spherical shallow waters.

The code for the calculation of eigenvalues has been kindly provided by P. Swarztrauber and is extensively discussed in [17]. In contrast to the numerical approach of Longuet-Higgins [6] this code does not require the prior derivation of a wave equation. It rather takes direct advantage of the fact that shallow water theory represents a 2D vector field. Expanding the eigensolutions in spherical vector harmonics, the first order equations of motion are transformed into an infinite, symmetric, homogeneous, pentadiagonal linear algebraic system. The eigenvalues of this system are readily computed numerically. Although the β-plane approximation and the numerical approach have been well established for 30 years, they have—to our knowledge—never been compared in a common format. Besides a direct validity test, such a comparison elucidates the characteristic dispersive features of the β-plane regime within the entire wave number space. In conjunction with the quantitative representation of the complete dispersion relation, the set of analytical approximations obtained here specifies particular dispersion regimes in wave number space. The characteristics of these regimes differentiate the wide range of circulation phenomena, represented by shallow waters on the rotating spherical surface.

II. WAVE EQUATIONS

On the surface of the rotating sphere the linearized shallow water equations read

$$\partial_t r + j^n{}_{;n} = 0 \ , \tag{2.1a}$$

$$\partial_t j_n + \epsilon_{mn} f j^m + c^2 \partial_n r = 0 \ , \tag{2.1b}$$

where r denotes a perturbation of the equilibrium mass per unit area R and the covariant vector $j_n = R v_n = aR(v_\lambda \cos\varphi, v_\varphi)$ the mass flux density near a zero-velocity equilibrium. The constant c denotes the system's intrinsic phase speed and the latitude-dependent scalar $f = 2\Omega \sin\varphi$ is Laplace's Coriolis parameter. The components of the antisymmetric Levi-Cività tensor ϵ_{mn} are given in the Appendix. As necessary for a covariant system, the mass flux density, defined in the divergence term of the continuity equation, equals the momentum density, defined in the prognostic term of the momentum budget. With this identity, the vertically integrated form (2.1) of a 3D, hydrostatic system satisfies Newton's first law in a spatially 2D sense. Correspondingly, Laplace's representation of the Coriolos force refers exclusively to a spatially 2D system.

For the potential vorticity z, defined by

$$Rz = \epsilon^{mn} v_{n;m} - fr/R \ ,$$

the equations of motion (2.1) imply the relation

$$R\partial_t z + f^n v_n = 0 \ , \tag{2.2}$$

where f^n is the contravariant gradient of the Coriolis parameter. The system's energy budget

$$\partial_t Re + s^n{}_{;n} = 0$$

with total energy

$$Re = \tfrac{1}{2} R v^n v_n + c^2 r^2 / 2R$$

and Poynting vector $s_n = c^2 r v_n$ follows by the usual algorithm Newton dynamics from (2.1).

One form of the linear wave equation of shallow waters on the rotating sphere is obtained by eliminating the momentum vector from (2.1). To this end take the divergence of (2.1b),

$$(\partial_t^2 - c^2\Delta) r - \epsilon^{mn} (f j_m)_{;n} = 0 \ ,$$

where Δ denotes the 2D Laplacian of a scalar in spherical coordinates as defined in the Appendix. With the repeated use of (2.1b) and $\epsilon^{mn}\epsilon_{am} = -\delta_a^n$ the time derivative of this equation becomes

$$[(\partial_t^2 - c^2\Delta)\partial_t + c^2 \epsilon^{mn} f_n \partial_m] r = (f^2 j^n)_{;n} \ .$$

Evaluation of the right-hand side (RHS) with the help of the mass and potential vorticity budgets leads to

$$[(\partial_t^2 + f^2 - c^2\Delta)\partial_t + c^2 \epsilon^{mn} f_n \partial_m] r = -2R^2 f \partial_t z \ . \tag{2.3}$$

The wave operator on the LHS of (2.3) is the 2D prolate spheroidal wave operator. It is stressed that the spheroidal character of this operator is here not a consequence of underlying 3D spheroidal coordinates. The underlying 3D coordinates are spherical and the spheroidal wave operator appears in (2.3) as a consequence of the nonvanishing Coriolis parameter. To eliminate the potential vorticity from (2.3) consider the operator

$$d^{an} = g^{an}\partial_t + \epsilon^{an} f$$

with

$$d^{an} d_{mn} = (\partial_t^2 + f^2)\delta_m^a \ .$$

Application of this operator to (2.1b) yields

$$(\partial_t^2 + f^2) j^a = -c^2 d^{an} \partial_n r \ . \tag{2.4}$$

After scalar multiplication with the covariant gradient of the Coriolis parameter this can be written with the help of (2.2) as

$$R^2(\partial_t^2+f^2)\partial_t z = c^2 d^{an} f_a \partial_n r \ .$$

Hence, on multiplication of (2.3) with $(\partial_t^2+f^2)$ one obtains

$$(\partial_t^2+f^2)[(\partial_t^2+f^2-c^2\Delta)\partial_t + c^2\epsilon^{mn} f_n \partial_m]r = -c^2 d^{mn}\partial_m f^2 \partial_n r \ . \quad (2.5)$$

This is the linear wave equation of shallow waters on the rotating sphere in terms of the mass perturbation r. While this is not identically the prolate spheroidal wave equation, it is nevertheless obvious that this operator plays a central role in (2.5).

Alternatively, the wave equation may be written in terms of the momentum vector. Taking the time derivative of (2.1b),

$$\partial_t^2 j_n - \epsilon_{mn} f(\epsilon^{am} f j_a + c^2 \partial^m r) + c^2 \partial_n \partial_t r = 0 \ ,$$

and using $\epsilon^{am}\epsilon_{mn} = -\delta^a_n$ as well as the continuity equation in the last term, this becomes

$$(\partial_t^2+f^2) j_n - c^2 j^a{}_{;an} - c^2 \epsilon_{mn} f \partial^m r = 0 \ . \quad (2.6a)$$

To evaluate the gradient of the divergence of the momentum flux, use the definition of potential vorticity

$$\epsilon^{rs} j_{s;r} = R^2 z + fr$$

and define the notation

$$\underline{\Delta} j_n = g^{ab} j_{n;ab} - a^{-2} j_n$$

so that formula (A2) of the Appendix assumes the form

$$j^a{}_{;an} = \underline{\Delta} j_n + \epsilon_{na}\partial^a (R^2 z + fr) \ .$$

Inserting this into (2.6a)

$$(\partial_t^2+f^2-c^2\underline{\Delta}) j_n - c^2 \epsilon_{na}(R^2 \partial^a z + r f^a) = 0$$

and differentiating again with respect to time

$$(\partial_t^2+f^2-c^2\underline{\Delta})\partial_t j_n + c^2 \epsilon_{na}[\partial^a (f^b j_b) + f^a j^b{}_{;b}] = 0 \quad (2.6b)$$

yields a wave equation in terms of the momentum flux vector. In contrast to the temporally fifth order, scalar equation (2.5) this is a temporally third order vector equation. The prolate spheroidal vector operator is seen to play a prominent role in (2.6b). While both (2.5) and (2.6b) exhibit some identifiable structure, they are nevertheless still fairly complex. A more symmetric form of the wave equation can be obtained with the help of certain derivatives of (2.6b).

Consider first the scalar product of (2.6b) with the gradient of the Coriolis parameter. To this end evaluate

$$\Delta(f^n j_n) = g^{ab}(f^n j_n)_{;ab}$$
$$= g^{ab}(f^n j_{n;ab} + 2 f^n{}_{;a} j_{n;b} + f^n{}_{;ab} j_n$$

using

$$f_{n;m} = -f G_{nm}$$

where G_{nm} is the Ricci tensor, defined in the Appendix, and

$$\Delta f = f^n{}_{;n} = -2a^{-2} f \ ,$$

with the result

$$f^n \underline{\Delta} j_n = \Delta(f^n j_n) - (\Delta f) j^n{}_{;n} \ ,$$

or with (2.1a) and (2.2)

$$f^n \underline{\Delta} j_n = (\Delta f)\partial_t r - R^2 \Delta \partial_t z \ .$$

Hence, the scalar product of (2.6b) with the gradient of f becomes

$$R^2[(\partial_t^2+f^2-c^2\Delta)\partial_t + c^2\epsilon^{nm} f_n \partial_m]\partial_t z = -c^2(\Delta f)\partial_t^2 r \ .$$

Together with (2.3) this equation forms a coupled set of prolate spheroidal wave equations for the mass perturbation r and the potential vorticity z. Furthermore, if (2.7) is read in terms of Cartesian coordinates (x,y) with $f=\beta y$, i.e., in particular, $\Delta f = 0$, this equation reduces to the familiar β-plane version of the Matsuno equation. Equation (2.7) is thus the spherical Matsuno equation.

Similarly, one obtains equations for the cross product of the gradient of f with the momentum vector

$$Ru = \epsilon^{nm} f_m j_n$$

from (2.6b). Without derivation the equations

$$[(\partial_t^2+f^2-c^2\Delta)\partial_t + c^2\epsilon^{ab} f_a \partial_b]r = 2R\,\partial_t u \quad (2.8a)$$

and

$$[(\partial_t^2+f^2-c^2\Delta)\partial_t + c^2\epsilon^{ab} f_b \partial_a]u = -R[(\partial_t^2+f^2-c^2\Delta)\partial_t + c^2\epsilon^{ab} f_a \partial_b]fz \quad (2.8b)$$

are mentioned here. Equations (2.3)–(2.8) demonstrate that, in spite of the non-Euclidean geometry, covariant differentiation renders the evaluation of higher-order wave equations from the equations of motion (2.1) fairly straightforward.

On the spherical surface, waves propagate in a fundamentally anisotropic environment: only in one direction is the propagation of waves possible, while standing waves form in the perpendicular direction. Due to its topology as an unbounded but finite domain, the spherical surface always acts as a waveguide. On the rotating sphere the propagation direction is the longitudinal direction and it can always be chosen as such in the nonrotating case. This suggests the dependent variables to be of the form

$$e^{-i(\omega t - M\lambda)} F(y)$$

where $y = \sin\varphi$ and $F(y)$ is a latitudinal eigenfunction. To ensure sufficient differentiability on the entire domain, the zonal wave number M should be an integer. For functions of this form the scalar, prolate spheroidal wave operator reduces to the prolate angular operator

$$P = \partial_y(1-y^2)\partial_y - \frac{M^2}{1-y^2} - \alpha^2 y^2 + \nu^2$$

where $\nu = a\omega/c$, while the Lamb parameter $\alpha = 2a\Omega/c$ measures the ratio of the planet's angular velocity to the propagation speed of effective pressure perturbations. For small values of α, prolate angular functions resemble associated Legendre polynomials, while they behave like parabolic cylinder functions for large values of the Lamb parameter. The terrestrial weather and climate systems operate over a wide range of values of this parameter. For atmospheric Lamb waves $\alpha \approx 1$, whereas $\alpha \approx 5$ for barotropic gravity waves in the ocean and $\alpha \approx 300$ for the first baroclinic mode in the ocean. For the purposes of planetary circulation theory, solutions at fixed values of α are therefore of limited interest and the solution of linear shallow waters on the rotating sphere has to be determined for the entire (real and positive) range of the Lamb parameter.

With wave functions of the above form the coupled system (2.3) and (2.7) becomes

$$(P-m)V = -2\alpha yD \;, \tag{2.9a}$$

$$(P+m)D = 2\alpha yV \;, \tag{2.9b}$$

where $D = r/R$ and $V = aRz/c$, while $m = \alpha M/\nu$. For Eqs. (2.8) one obtains correspondingly

$$(P-m)D = -2\alpha U \;,$$

$$(P+m)U = -(P-m)yV \;,$$

with $U = v_\lambda \cos\varphi/c = au/2c\Omega$. The system (2.9) can formally be diagonalized in different ways. For instance,

$$(P+m)y^{-1}(P-m)V = -4\alpha^2 yV \;,$$

$$(P-m)y^{-1}(P+m)D = -4\alpha^2 yD \;,$$

or

$$P(2\alpha y+m)^{-1}P(V+D) = -(2\alpha y-m)(V+D) \;,$$

$$P(2\alpha y-m)^{-1}P(V-D) = -(2\alpha y+m)(V-D) \;.$$

As in both cases the RHS is not readily incorporated into a factorized form of the LHS, these equations do not suggest an obvious strategy to evaluate eigensolutions. Similar expressions are obtained from the U equations. Although the latitudinal eigenfunctions are generally not just prolate angle functions, these equations indicate nevertheless the fundamental role of the prolate angle operator.

III. WAVE PROPAGATION ON THE SURFACE OF THE ROTATING SPHERE

In addition to long gravity waves, rotating sphere shallow waters admit the propagation of Rossby waves. These waves emerge as a consequence of restorative torques generated by the latitudinally varying Coriolis force. The notion of "mixed Rossby-gravity" and "Kelvin" waves does not refer to qualitatively different waves, but to the specific organization of eastern and western branches into modes by the β-plane approximation. Gravity and Rossby waves in rotating shallow waters are dynamically related to sound and Alfvén waves in magnetohydrodynamics.

A large part of the dynamical wealth of planetary circulations is due to the fact that the rotating spherical surface provides an anisotropic and inhomogeneous arena for the propagation of shallow water waves. Besides the basic latitude-longitude anisotropy, rotation effects on the spherical surface introduce an east-west asymmetry into the propagation of waves. However, in a Doppler-type framework gravity waves are expected to be symmetric, as long as observer and medium are at rest relative to each other. In addition to these asymmetries, Yoshida [5] has pointed out that the rotating spherical surface is inhomogeneous with respect to latitude. At low frequencies, latitudinal eigenfunctions behave oscillatorily only in a narrow equatorial belt, the Yoshida waveguide.

The f-plane approximation, successfully employed in the study of internal gravity waves, does not admit Rossby waves, nor does it exhibit anisotropy or inhomogeneity. In contrast, Margules's asymptotics [2] represent the anisotropic, globally homogeneous limits of the system. In the nonrotating case ($\Omega = 0$) one obtains from the wave equations of the previous section

$$r = r_0 P_L^M(\sin\varphi)e^{-i\theta} \;, \tag{3.1a}$$

$$j_n = -j_0 \partial_n P_L^M(\sin\varphi)e^{-i\theta} \;, \tag{3.1b}$$

with $j_0 = ir_0c^2/\omega$, $\theta = \omega t - M\lambda$, and associated Legendre polynominals P_L^M. The dispersion relation of these long gravity waves

$$\nu^2 = L(L+1), \quad -L \le M \le L \;,$$

or alternatively in terms of the mode number $N_0 = L - |M| \ge 0$, counting the zeros of the mass perturbation r in the open interval $y \in (-1,1)$,

$$\nu^2 = N_0(N_0+1) + (2N_0+1)|M| + M^2 \;, \tag{3.1c}$$

is symmetric with respect to the zonal wave number M. On the other hand, for nondivergent flow the limit $c^2 \to \infty$, $r = \text{const}$ of the Matsuno equation (2.7) yields

$$r = \text{const} \;,$$

$$j_n = j_0 \epsilon_{nm} \partial^m P_L^M(\sin\varphi)e^{-i\theta} \;, \tag{3.2b}$$

representing divergence-free Rossby waves with dispersion relation

$$\omega = -2\Omega M/L(L+1), \quad -L \le M < 0 \;,$$

or in terms of the mode number $N_2 = L - |M| \ge 0$, here counting the zeros of j_2 in the open interval $y \in (-1,1)$,

$$\sigma = -M/[N_2(N_2+1) + (2N_2+1)|M| + M^2] \tag{3.2c}$$

with $\sigma = \omega/2\Omega$. Besides latitude-longitude anisotropy and global homogeneity these Rossby waves exhibit a pronounced east-west asymmetry: the propagate westward only. While representing entirely different dynamics, both Margules's limits correspond to the same value of the Lamb parameter, $\alpha = 0$. At this value the rotating spherical surface is globally homogeneous and the latitudinal eigenfunctions are expressed in terms of associated Legendre polynominals.

For finite α, the Cartesian approximation of the spherical geometry near the equator leads to the β-plane approximation. It then follows from the β-plane version of the Matsuno equation (2.7) that the latitudinal eigenfunctions can be approximated in terms of parabolic cylinder functions. β-plane dynamics capture the latitude-longitude anisotropy as well as the east-west asymmetry of low-frequency gravity waves and Rossby waves. As an equatorial approximation, the β plane applies essentially inside the Yoshida waveguide. Margules's asymptotics as well as Matsuno's β plane thus capture primarily homogeneous limits and subdomains.

Exact analytical solutions, which take full account of the latitudinal inhomogeneity for all values of the Lamb parameter, have been derived in special cases [7]. Instead of the diagonalizations of (2.9), less symmetric low-order closures of the coupled system are considered. To obtain a closed form of the Matsuno equation (2.9a) eliminate j_1 between the continuity equation (2.1a) and the one-component of the momentum budget (2.1b) with the result

$$\alpha D = -(\mu+y^2)^{-1}[(1-y^2)\partial_y - my]V ,$$

where $\mu=(M^2-\nu^2)/\nu^2$. Inserting this expression on the RHS of (2.9a), the Matsuno equation becomes

$$(P+m)V=2(\mu+y^2)^{-1}[(1-y^2)y\partial_y+m\mu]V . \quad (3.3)$$

The prolate angular operator maintains its dominating role in this form, while the RHS appears of little advantage for general values of M and ν. However, for strictly standing waves, $M=0$ and $\mu=-1$ so that the Matsuno equation (3.3) assumes the simple form

$$[(1-y^2)\partial_y^2-\alpha^2y^2+\nu^2]V=0 . \quad (3.4)$$

This equation is exactly solved by

$$V=A(1-y^2)^{1/2}S_L^{-1}(y;\alpha) .$$

Here, A is a constant and $S_L^{-1}(y;\alpha)$ the prolate angular function of degree $L\geq 1$ and order (-1), where the negative value has been chosen for correspondence purposes. From this expression, the full eigensolution for standing waves on the rotating spherical surface is found as

$$r=r_0(\partial_\varphi-\tan\varphi)S_L^{-1}e^{-i\omega t} , \quad (3.5a)$$

$$j_1=-j_0\cos\varphi f S_L^{-1}e^{-i\omega t} , \quad (3.5b)$$

$$j_2=ij_0\omega S_L^{-1}e^{-i\omega t} , \quad (3.5c)$$

with $j_0=a^2r_0$ and dispersion relation

$$\nu^2=\epsilon_{\mathrm{PA}}(L,1;\alpha) , \quad (3.5d)$$

where $\epsilon_{\mathrm{PA}}(L,K;\alpha)=\epsilon_{\mathrm{PA}}(L,-K;\alpha)$ denotes the eigenvalue corresponding to the prolate angular function $S_L^K(y;\alpha)$. A closed expression for these eigenvalues does not exist [16,18]. Furthermore, due to the lack of recurrence relations, (3.5) has to be given in terms of the prolate angular function and its first derivatives. This solution describes essentially standing gravity waves, where the β-plane approximation denotes the case $L=1$ as standing mixed Rossby-gravity wave. In contrast to the nonrotating case, standing waves on the rotating spherical surface exhibit a nontrivial j_1 component. The latitude of zero curvature of V determines the width of the Yoshida waveguide, since in the neighborhood of this latitude the eigensolutions are approximated by Airy functions. For standing waves the critical latitude is obtained from (3.4) as

$$\sin\varphi_{\mathrm{cr}}=\pm\omega/2\Omega .$$

This demonstrates that low-frequency waves are concentrated in a narrow belt around the equator, while waves with frequencies higher than the inertial frequency do not detect the Yoshida waveguide at all. The solutions (3.5) thus account for the latitude-longitude anisotropy as well as the latitudinal inhomogeneity of the rotating spherical surface. As standing waves they do not exhibit an east-west asymmetry.

For $\alpha=0$, prolate angular functions reduce to associated Legendre polynomials. In the nonrotating limit, (3.5a) therefore becomes

$$r=r_0(\partial_\varphi-\tan\varphi)P_L^{-1}e^{-i\omega t} .$$

With the familiar recurrence relations for associated Legendre polynominals [18] it is readily seen that

$$\partial_\varphi P_L^{-1}=P_L^0+P_L^{-1}\tan\varphi .$$

In the nonrotating, globally homogeneous case (3.5) thus coincides with the Margules limit (3.1) for $M=0$.

The asymptotics of the prolate angular equation provide an estimate of the domain of validity of the β-plane approximation. Generally, β-plane dynamics are obtained by introducing the Cartesian approximation to the spherical geometry in the vicinity of the equator, formulating the equations of motion (2.1) in these coordinates, and finally deriving the system's wave equation. Conversely, it will here be shown that appropriate approximations of the spherical Matsuno equation lead indeed to the same result on the β plane. The necessary approximations define the conditions of validity of the β-plane approximation in physical and wave number space.

Assuming

$$V=(1-y^2)^{|M|/2}F(\eta)$$

with $\eta=y\sqrt{2\alpha}$ the spherical Matsuno equation (3.3) transforms to

$$[(2\alpha-\eta^2)\partial_\eta^2-2(|M|+1)\eta\partial_\eta-\tfrac{1}{2}\alpha\eta^2+X]F$$
$$=2Y^{-1}[(2\alpha-\eta^2)\eta\partial_\eta-|M|\eta^2+2\alpha\mu m]F ,$$

where $X=\nu^2-M^2+m$ and $Y=2\alpha\mu+\eta^2$. For large $\alpha\gg1$ and $2\alpha\gg\eta^2$, i.e., in the vicinity of the equator, $y^2\ll1$, this is approximated by the equation for parabolic cylinder functions

$$(\partial_\eta^2-\tfrac{1}{4}\eta^2+N_2+\tfrac{1}{2})F\approx0 ,$$

which is the familiar β-plane version of the Matsuno equation with dispersion relation

$$\nu^3-[\alpha(2N_2+1)+M^2]\nu-\alpha M\approx0 , \quad (3.6)$$

where the mode number $N_2=L-|M|$ here counts the zeros of V in the open interval $\eta \in (-\infty, \infty)$. It may be noted that the spectrum of the parabolic cylinder operator is actually continuous. In the approximation of an operator with a discrete spectrum, the mode number emerges here necessarily as an integer, thus selecting those parabolic cylinder functions which can be represented in terms of Hermite polynominals. In contrast, the conventional derivation of the Matsuno equation on the β plane has to confine itself quite arbitrarily to integer mode numbers. Moreover, this derivation of the equatorial β plane can obviously not be generalized to obtain a "midlatitude β plane." Besides formal difficulties to define it as an approximation to the spherical problem, such a concept would necessarily imply unphysical features like a "midlatitude waveguide," midlatitude mixed Rossby-gravity waves, and—independent of coastal trapping—midlatitude Kelvin waves. The β plane is thus the appropriate approximation for equatorial low-frequency waves, if the Lamb parameter is large. As an asymptotic expansion for large α, the β plane excludes a nonrotating limit and approximates the exterior of the Yoshida waveguide by an infinite plane. On Earth, large α refer to baroclinic waves in the ocean. Baroclinic low-frequency dynamics in the equatorial ocean are indeed successfully simulated by reduced-gravity models in β-plane geometry [10,11].

A second exact solution in terms of prolate angular functions is readily found from a low-order closure of (2.9b). Evaluation of the two-component of (2.4) yields

$$\alpha V = -(\sigma^2 - y^2)^{-1}[(1-y^2)\partial_y + my]D ,$$

where $\sigma = \nu/\alpha$. With this representation of V, Eq. (2.9b) becomes

$$(P-m)D = -2(\sigma^2-y^2)^{-1}[(1-y^2)y\partial_y + m\sigma^2]D . \quad (3.7a)$$

The structure of (3.7a) is similar to the spherical Matsuno equation (3.3). A particularly simple form of (3.7a) is obtained for inertial waves with $\sigma = 1$:

$$\left[(1-y^2)\partial_y^2 - \frac{M(M-2)}{1-y^2} - \alpha^2 y^2 + \alpha^2 - M\right] D = 0 . \quad (3.7b)$$

This is exactly solved by

$$D = A(1-y^2)^{1/2} S_L^{M-1}(y;\alpha) ,$$

where A denotes a constant amplitude and $S_L^{M-1}(y;\alpha)$ the prolate angular function of degree $L \geq 0$ and order $(M-1)$ with $1-L \leq M \leq L+1$. The eigensolutions of linear shallow waters for inertial waves on the rotating sphere are found from this expression as

$$r = r_0 \cos\varphi S_L^{M-1} e^{-i\theta} , \quad (3.8a)$$

$$j_1 = j_0 \left[\sin\varphi \partial_\varphi + \frac{M-1}{\cos\varphi} + \cos\varphi\right] S_L^{M-1} e^{-i\theta} , \quad (3.8b)$$

$$j_2 = -ij_0(\cos\varphi)^{-1}[\partial_\varphi + (M-1)\tan\varphi] S_L^{M-1} e^{-i\theta} , \quad (3.8c)$$

with $j_0 = c^2 r_0 / 2\Omega$ and dispersion relation

$$\nu^2 = \alpha^2 = \epsilon_{\rm PA}(L, M-1;\alpha) + M . \quad (3.8d)$$

The lack of recurrence relations leads again to a representation of (3.8) in terms of the prolate angular function and its first derivatives. This solution describes gravity waves at the inertial frequency $\omega = 2\Omega$, including the Kelvin wave for $M = L+1 > 0$ and the gravity branch of the mixed Rossby-gravity wave for $M = L > 0$. It accounts not only for the latitude-longitude anisotropy and latitudinal inhomogeneity, but also for the east-west asymmetry. As indicated by (3.8d), $\nu(M) \neq \nu(-M)$.

The cosine of latitude in the denominator of the momentum flux of inertial waves raises the question of regularity at the poles for this solution. Near the poles the prolate angular function behaves like [16]

$$S_L^{M-1}(y;\alpha) = (1-y^2)^{|M|/2} F(L, M-1;\alpha) ,$$

where F generally does not vanish. This is sufficient to ensure the polar regularity of (3.8), except for $M=0$ and 2. For $M=2$, tabulated values of the spheroidal wave function [16] show that $S_L^1(y=\pm 1;\alpha)=0$. However, it is not obvious that this value is approached sufficiently fast for all L and α. At $M=0$, (3.8) should not only be regular at the poles, but also coincide with (3.5) at $\nu = \alpha$. With regard to the dispersion relations (3.5d) and (3.8d) this correspondence is obvious. Regularity and correspondence of the solutions can in this case even be checked directly, since $S_L^{\pm 1}(y;\alpha^2 = \epsilon_{\rm PA})$ is known in closed form [16]. For even L

$$S_L^{-1}(y;\alpha^2 = \epsilon_{\rm PA}) = A \sin(\alpha \sin\varphi)/\cos\varphi$$

with constant A is a regular prolate angular function, if α is an even multiple of $\pi/2$. A similar expression holds for odd L [16]. Thus, if α is for instance an even multiple of $\pi/2$, the solutions (3.5) and (3.8) become

$$r \sim \cos(\alpha \sin\varphi) ,$$

$$j_1 \sim \sin\varphi \sin(\alpha \sin\varphi) ,$$

$$j_2 \sim \sin(\alpha \sin\varphi)/\cos\varphi .$$

This demonstrates that at $M=0$ regular and common solutions (3.5) and (3.8) exact, although not for arbitrary values of the Lamb parameter. Hence standing waves at $\omega = 2\Omega$ only occur if α assumes certain specific values. It is furthermore noted that for $M = \frac{1}{2}$ and $\frac{3}{2}$ Eq. (3.7b) has exact solutions in terms of Mathieu functions. However, at these values of M the full wave function is not continuous for all longitudes $\lambda \in [0, 2\pi]$.

Exact solutions in terms of prolate angular functions capture the latitude-longitude anisotropy and the east-west asymmetry as well as the latitudinal inhomogeneity of the rotating spherical surface. In contrast to associated Legendre polynomials or parabolic cylinder functions they are capable of representing an equatorial waveguide on the finite spherical surface, where the inhomogeneity-scale—i.e., the width of the Yoshida waveguide—is controlled by the Lamb parameter. In the corresponding limits the prolate angular functions reproduced Margules's asymptotics and the β-plane approximation, respectively. At sufficiently high frequencies, gravity waves are no longer confined by an equatorial waveguide, but propagate on a globally homogeneous, though anisotropic spherical surface. From an alternative viewpoint, the Lamb parameter appears as the remnant of a vertical wave number in a vertically integrated, strictly horizontal system and the dependence of eigenfrequencies and eigenfunctions on this "wave number" as a rotation-induced "vertical dispersion." In a two-layer model, the barotropic eigenfrequency $\omega_1(L,M;\alpha_1)$ and the baroclinic eigenfrequency $\omega_2(L,M;\alpha_2)$ at the same L and M belong to different eigenfunctions $S_L^M(y;\alpha_1)$ and $S_L^M(y;\alpha_2)$, respectively. In the nonrotating limit these eigenfunctions reduce to the same associated Legendre polynominal $P_L^M(y)$ and the system becomes strictly nondispersive. There is, nevertheless, a pronounced difference between the Lamb parameter and a wave number in the proper sense: for given α the prolate angular functions form a complete, orthonormal set of basis functions, while this is not the case for given L and M. The Lamb parameter provides the link between covariant, linear shallow waters and vertical variability that maintains consistency with respect to Laplace's representation of inertial forces on the rotating spherical surface. Obviously, this vertical variability has to appear as internal variability of a strictly 2D fluid. The representation of vertical variability in terms of a set of Lamb parameters *and* a vertical wave number is not only redundant, but will generally be inconsistent. In practice, prolate angular functions are the appropriate set of basis functions for a rotating, shallow few-layer fluid on the global domain.

IV. THE DISPERSION RELATION

While exact eigensolutions could here only be obtained in special cases, the identification of the role of the prolate spheroidal wave operator is nevertheless sufficient for a qualitative overview of the system's wave number space. The asymptotics of this operator yield a set of approximations of the dispersion relation, which link the Margules limits with the β plane. These approximations are here tested against numerical solutions of the full eigenvalue problem of shallow waters on the rotating spherical surface.

Benchmarks for these approximations are provided by the exact dispersion relations (3.5d) and (3.8d) as well as the Margules limits (3.1c) and (3.2c). The prolate angular eigenvalues are explicitly known [16,18] in terms of the power series

$$\epsilon_{PA}(L,K;\alpha)=\Lambda^2+\alpha^2(2\Lambda^2-2K^2-1)/(4\Lambda^2-3)+\cdot$$

with $\Lambda^2=L(L+1)$, or for large α in terms of the asymptotic expansion

$$\epsilon_{PA}(L,K;\alpha)=\alpha q-2^{-2}(2p+3-4K^2)$$
$$-(16\alpha)^{-1}q(p+3-8K^2)- \qquad (4.1b)$$

with $q=2N+1$, $p=N(N+1)$, and $N=L-|K|$. Higher-order terms are given in [16,18] and the optimal matching of both expressions depends on L, K, and α. The dispersion relations (3.5d) with $K=-1$ and (3.8d) with $K=M-1$ can be evaluated with (4.1) to an arbitrary degree of accuracy. For wave numbers $M=-1,0,1$ Fig. 1 shows these eigenfrequencies in comparison to numerically computed eigenfrequencies as a function of the Lamb parameter α. Similar figures of numerical solutions of the eigenvalue problem have been obtained by Longuet-Higgins [6], who shows $\sigma=\nu/\alpha$ vs $1/\alpha$. For all three wave numbers the intersections of the dashed line $\nu=\alpha$ with the solid lines have been calculated with (3.8d), while the solid lines represent numerical computations. These computations use the Swarztrauber-Kasahara code [17] to solve the equations of motion (2.1) numerically. In Fig. 1(b), the solid lines represent in fact two calculations: numerical solutions of the eigenvalue problem and (3.5d). Both results obviously coincide. For $\alpha=0$ the numerical code [17] reproduces the Margules limits (3.1c) and (3.2c).

Figure 1(a) clearly shows low-frequency Rossby waves, well separated by the mixed Rossby-gravity wave from higher-frequency gravity modes. The standing waves in Fig. 1(b) are gravity waves, where the β-plane approximation denotes the lowest mode as mixed Rossby-gravity wave. Figure 1(c) depicts eastward traveling gravity waves with the distinguished behavior of the Kelvin mode ($N_0=0$) for large values of α. In the β-plane approximation, the mode $N_0=1$ is considered the gravity branch of the mixed Rossby-gravity wave. Figure 1 indicates a pronounced difference in the behavior of ν for small and large values of α, corresponding to the approximation of prolate angular functions in terms of associated Legendre polynomials or parabolic cylinder functions. The transition region is approximately given by $1\le\alpha\le10$, where these approximations converge poorly in comparison to the left and right boundary regions of the diagrams. Under terrestrial conditions, this transition range is occupied by Lamb waves, atmospheric gravity, and Rossby waves as well as barotropic waves in the ocean.

For inertial waves (3.8), the east-west asymmetry can be evaluated explicitly. According to (4.1b) the mode number in this case is given as $N_0=L-|M-1|$, counting the numbers of zeros of the mass perturbation r in the open interval $y\in(-1,1)$. For given Lamb parameter and mode number, (3.8d) has two solutions and with

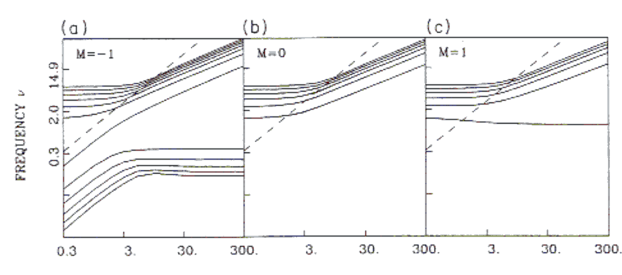

FIG. 1. Nondimensional frequency $\nu=a\omega/c$ versus Lamb parameter $0.3\le\alpha=2a\Omega/c\le300$ in logarithmic scales for mode numbers $0\le N_0\le5$. Dashed line: $\nu=\alpha$. (a) Zonal wave number $M=-1$. Solid lines: numerical computations. (b) Zonal wave number $M=0$. Solid lines: numerical computations and (3.5d). (c) Zonal wave number $M=1$. Solid lines: numerical computations.

(4.1b) one obtains to $O(\alpha^0)$

$$M_{E/W}=\tfrac{1}{2}\pm\sqrt{\alpha^2-\alpha q+\tfrac{1}{12}p}$$

such that

$$|M_W|=M_E-1<M_E \tag{4.2}$$

This inequality is maintained if higher-order terms of (4.1b) are included or (4.1a) is used for small α with $L=N_0+|M-1|$. The inequality (4.2) states that a corotating observer finds a smaller phase speed for eastward traveling inertial waves than for their westward counterparts at the same frequency and mode number.

To evaluate the physics of this asymmetry, assume a Doppler-type approach and consider the effect of relative motion between medium and observer on the rotating sphere. The stationary, zonal flow

$$R=\text{const},\quad U_n=a^2U(\cos^2\varphi,0)$$

with constant U is an exact solution of the fully nonlinear shallow water equations on the rotating spherical surface. It is uncritical in the present context that this solution generally requires some external time-independent forcing for arbitrary values of U. This stationary flow is nondivergent and exhibits the shear

$$U_n;_m=\epsilon_{mn}U\sin\varphi \ .$$

Linearization of the shallow water equations around this equilibrium leads to

$$(\partial_t+U\partial_\lambda)r+j^n;_n=0\ ,$$

$$(\partial_t+U\partial_\lambda)j_n+\epsilon_{mn}Fj^m+c^2\partial_n r=0\ ,$$

where $F=2(\Omega+U)\sin\varphi$. On the spherical surface, the stationary mean flow is seen to introduce two effects in comparison to (2.1): first, the partial time derivative is replaced by the substantial derivative following the mean flow. This is the Doppler effect in the familiar sense. Secondly, the shear of the zonal flow modifies the Coriolis parameter. As a stationary zonal flow on a spherical surface will generally be a shear flow, this additional shear effect will always appear for a medium moving relative to the observer. For the above system it is in particular obvious that the shear of a westward flow with $U=-\Omega$ (i.e., a medium at rest with respect to a nonrotating observer) completely compensates the Coriolis effect. Conversely, this implies that the effect of the Coriolis term on the dispersion of gravity waves on the rotating spherical surface is physically equivalent to the shear effect of a stationary, eastward zonal flow on the dispersion of gravity waves on the nonrotating sphere, as seen by an observer moving with the flow.

For general values of U the dispersion relation of the above system

$$\omega_*=\omega_*(L,M;\alpha_*$$

has the same form as the dispersion relation of (2.1), where $\omega_*=\omega-UM$ is the frequency seen by an observer moving with the flow and $\alpha_*=2a(\Omega+U)/c$. In the case of a westward current with $U=-\Omega$ the corotating observer will find gravity waves with

$$\nu=\sqrt{L(L+1)}-\tfrac{1}{2}\alpha M\ ,$$

i.e., the eastward propagation of gravity waves with $M>2\sqrt{L(L+1)}/\alpha$ is inhibited, while stationary perturbations ($\nu=0$) can propagate with $M=2\sqrt{L(L+1)}/\alpha$ (provided, of course, the RHS of this equation assumes an integer value). Rossby waves do not propagate on this flow. On the other hand, for a westward current with $U=-2\Omega$ an observer moving with the flow will see eastward propagating Rossby waves only. In the nondivergent case, they satisfy the dispersion relation

$$\omega_*=2\Omega M/L(L+1),\quad 0<M\le L\ ,$$

while an observer rotating with the sphere still sees westward traveling Rossby waves

$$\omega=2\Omega M[1-L(L+1)]/L(L+1)\ .$$

These considerations demonstrate that the Lamb parameter captures the Coriolis effect, as well as shear effects of stationary, zonal flows on the dispersion of shallow water waves on the rotating spherical surface. Zonal currents on the surface of the sphere necessarily exhibit shear. In addition to the usual Doppler effect this shear introduces an east-west asymmetry into the propagation of waves that is still seen by an observer moving with the flow. The east-west asymmetry introduced by the Coriolis term is physically closely related to this shear-induced Doppler effect.

The same east-west asymmetry is exhibited by gravity waves with $\nu\gg\alpha$. In this frequency domain the denominator on the RHS of (3.7a) can be Taylor expanded. For (3.7a) this expansion yields an approximate prolate angular equation with the eigenvalue relation

$$\nu^3-\epsilon_{\mathrm{PA}}(L,M;\alpha)\nu+\alpha M\approx 0\ .$$

The gravity modes which satisfy the defining inequality are thus given as the roots

$$\nu_g\approx 2\nu_0\cos\gamma>\alpha \tag{4.3}$$

of this dispersion relation with $\nu_0^2=\epsilon_{\mathrm{PA}}/3$ and

$$\gamma=\tfrac{1}{3}\arccos(-\alpha M/2\nu_0^3)\ .$$

The domain of validity of (4.3) in Fig. 1 is the upper left corner above the dashed line. For $\alpha=5$ and mode numbers $2\le N_0=L-|M|\le 10$ these gravity modes are shown in Fig. 2 as a function of the zonal wave number M in comparison to numerically computed eigenfrequencies of the full system. For all mode numbers, (4.3) has been calculated with the power series expansion (4.1a) for ϵ_{PA} to $O(\alpha^2)$. Discrepancies appear at low N_0 and small wave numbers M. For sufficiently high frequencies, however, (4.3) provides a good approximation to the eigenfrequencies of long gravity waves on the rotating spherical surface. In agreement with (4.2) the dispersion diagram exhibits east-west asymmetry: westward branches are steeper than their eastward counterparts. For $\alpha=0$, (4.3) reduces to the Margules limit (3.1c) and the east- and westward branches become symmetrical with respect to

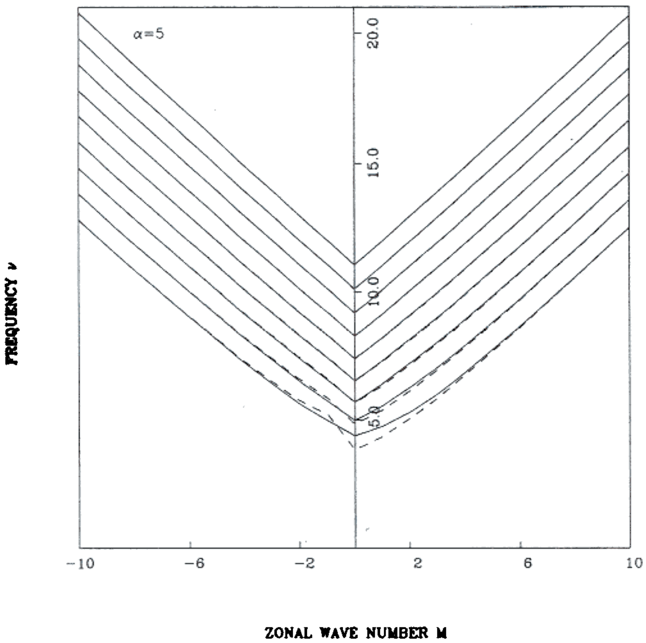

FIG. 2. Normal east-west asymmetry of gravity waves: westward branches are steeper than eastward branches. Nondimensional gravity frequencies $\nu = a\omega/c$ versus zonal wave number M for modes $2 \le N_0 \le 10$ and Lamb parameter $\alpha = 5$. Solid lines: (4.3), dashed lines: numerical computations.

$M = 0$. For increasing α discrepancies expand to higher mode numbers. However, with the inclusion of higher-order terms of (4.1a) it can always be achieved that agreement at $\nu = \alpha$ is fair and becomes rapidly better for $\nu > \alpha$. For $M = 0$, (4.3) does not reduce exactly to (3.5d), although Fig. 2 indicates satisfactory quantitative agreement for sufficiently large ν. For all values of α, gravity waves, which are well approximated by (4.3) do not experience an equatorial waveguide. These gravity waves gain large-scale relevance, if the Lamb parameter is small. On Earth this is the case for Lamb waves, atmospheric gravity waves, and barotropic gravity waves in the ocean, where the latter are of particular significance for the ocean's response to external tidal forces. On a slowly rotating planet like Venus, gravity waves (4.3) play a dominating role on large scales.

A second regime in wave number space is defined by the inequality $M^2 \gg \nu^2$. In this case $\mu \gg 1$ and the denominator on the RHS of the Matsuno equation (3.3) can be Taylor expanded. This expansion leads to an approximation of the Matsuno equation in terms of a prolate angular equation with the eigenvalue relation

$$\nu^3 - \epsilon_{PA}(L,M;\alpha)\nu - \alpha M \approx 0 .$$

The roots of this dispersion relation, which satisfy the defining inequality

$$\nu_r = -2\nu_0 \cos(\gamma + \tfrac{1}{3}\pi) < |M| \qquad (4.4)$$

where $\nu_0^2 = \epsilon_{PA}/3$ and

$$\gamma = \tfrac{1}{3}\arccos(\alpha M/2\nu_0^3)$$

represent Rossby waves. The domain of validity of (4.4) in Fig. 1(a) is the part below a horizontal line (not shown) $\nu = 1$. For $\alpha = 5$ and mode numbers $0 \le N_2 = L - |M| \le 5$ the Rossby frequencies $\sigma_r = \nu_r/\alpha$ are shown in Fig. 3 as a function of M in comparison to numerically computed eigenfrequencies of the full system. For $N_2 > 0$ the dispersion relation (4.4) has been evaluated with the expansion (4.1a) of ϵ_{PA} to $O(\alpha^2)$, while for $N_2 = 0$ (the Rossby branch of the mixed Rossby-gravity mode) the expansion (4.1b) to $O(\alpha^0)$ has been used. Except for small $|M|$ in the mixed Rossby-gravity mode, agreement is generally good. Inspection of Fig. 1(a) shows that for $\alpha = 5$ the frequency of the mixed Rossby-gravity mode at $M = -1$ is larger than 1 and thus outside the domain of validity of (4.4).

For $c \to \infty$, i.e., $\alpha \to 0$, the eigenfrequencies (4.4) reduce to the Margules limit (3.2c). In the opposite extreme $\alpha \gg 1$ they approach the familiar β-plane values. In contrast to the Margules limit and the β plane, the approximation (4.4) holds for Rossby waves at all values of the Lamb parameter and is not restricted to the vicinity of the equator. Since the regime-defining inequality involves wave numbers and frequencies only, (4.4) applies in the entire physical space. It includes in particular those Rossby waves which are frequently represented in terms of a "midlatitude β plane." As pointed out in the previous section such a concept is formally as well as physically inconsistent, while the approximation (4.4) is free of these difficulties. Baroclinic Rossby waves in the midlatitude Pacific play an important role in the long-term aftereffects of an El Nino event [12].

Linear β-plane dynamics approximate low-frequency waves in the equatorial waveguide for large values of the Lamb parameter. The dispersive properties of gravity and Rossby waves on the β plane are given by the roots

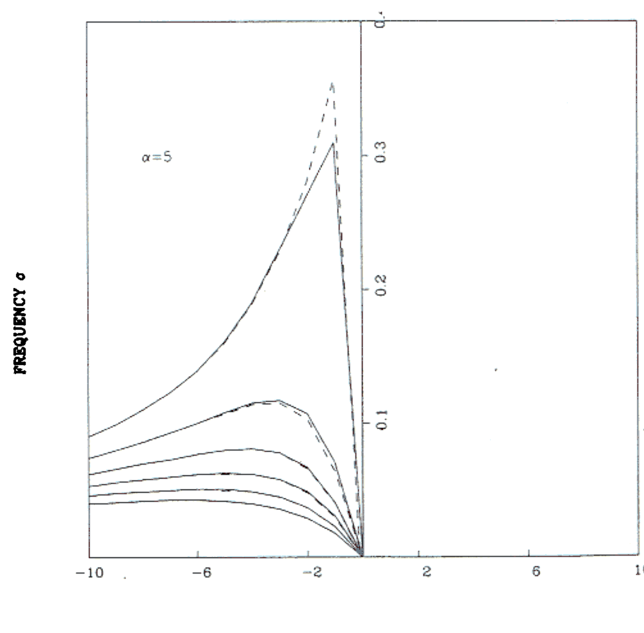

FIG. 3. Nondimensional Rossby frequencies $\sigma = \omega/2\Omega$ versus zonal wave number M for modes $0 \le N_2 \le 5$ and Lamb parameter $\alpha = 5$. Solid lines: (4.4), dashed lines: numerical computations.

of (3.6). For $N_2>0$, the low-frequency gravity waves are represented by the root

$$\nu_g \approx 2\nu_0 \cos\gamma \tag{4.5a}$$

while the Rossby modes are

$$\nu_r \approx -2\nu_0 \cos(\gamma+\tfrac{1}{3}\pi) \tag{4.5b}$$

with $\nu_0^2=[\alpha(2N_2+1)+M^2]/3$ and

$$\gamma=\tfrac{1}{3}\arccos(\alpha M/2\nu_0^3)\ .$$

A major insight of Matsuno's analysis of linear β-plane dynamics was the identification of the role of mixed Rossby-gravity and Kelvin waves. For $N_2=0$, only the root

$$\nu_{rg} \approx \tfrac{1}{2}M+\tfrac{1}{2}\sqrt{M^2+4\alpha} \tag{4.5c}$$

of (3.6) is taken, which combines the eastward gravity branch for $N_2=0$ with the highest westward Rossby branch into the mixed Rossby-gravity mode. Finally, the Kelvin wave is the trivial solution of the β-plane version of the Matsuno equation and its dispersion relation is consequently not included in (3.6). Rather, it has to be reevaluated from the equations of motion as

$$\nu_k \approx M\Theta(M)\ , \tag{4.5d}$$

where Θ here denotes the unit step function. For $\alpha=100$ and $N_2\le 3$ the dispersion relation (4.5) is shown in Fig. 4 in comparison to numerically computed eigenfrequencies of the full problem. The agreement is generally good. For increasing α the agreement improves, while for smaller α or higher frequencies discrepancies emerge. As a large-α approximation, (3.6) does not reduce to the Margules limits for vanishing Lamb parameter. While (3.6) does not exactly reduce to (3.8d) for strictly standing waves with $M=0$, Fig. 4 shows nevertheless that quantitative agreement is satisfactory. In Fig. 1 the domain of validity of the equatorial β plane is the lower right of the diagrams, below the dashed line and to the right of a vertical line (not shown) $\alpha=1$. The β plane thus excludes high-frequency gravity waves ($\nu\ge\alpha$) and barotropic Rossby waves ($\alpha\le 1$).

A prominent feature in Fig. 4 is the east-west asymmetry of the gravity modes (4.5a). For given frequency and mode number N_2 Fig. 4 clearly indicates

$$|M_W|>M_E\ .$$

In obvious contrast to (4.2) eastward propagating gravity waves have a higher phase speed than their westward counterparts at the same frequency and mode number. It will here be demonstrated that this difference is in fact an enhanced east-west asymmetry of the type (4.2). To this end notice that the dispersion relations (3.1c), (3.8d), and (4.3) utilize a mode number N_0, while (3.2c), (3.5d), (3.6), and (4.4) organize modes with respect to N_2. Both of these numbers represent physically different quantities: N_0 is the number of zeros of the mass perturbation r, while N_2 counts the zeros of the latitudinal momentum component j_2. To evaluate the role of these wave numbers consider the eigenfrequencies of gravity waves in either representation. Figure 5 shows numerically computed eigenfrequencies of gravity waves for $0\le N_0\le 10$ at Lamb parameters $\alpha=0,5,100$. Figure 6 depicts exactly the same eigenfrequencies, here, however, combined into modes according to constant N_2. It is pointed out that Fig. 5 does not show the roots of (3.6). These roots exhibit additional quantitative discrepancies with numerically calculated eigenvalues due to the fact that (3.6) involves only terms of $O(\alpha^0)$ of (4.1b). These quantitative discrepancies can partly be decreased by including higher-order terms of (4.1b) into (3.6). The qualitative feature of interest here is the fact that the β-plane ap-

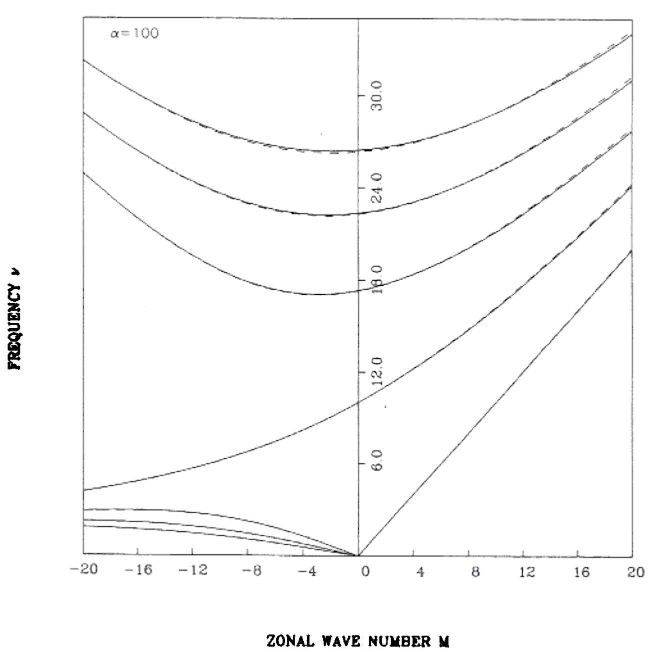

FIG. 4. β-plane dispersion relation. Nondimensional frequencies $\nu=a\omega/c$ versus zonal wave number M for modes $N_2\le 3$ and Lamb parameter $\alpha=100$. Solid lines: (3.6), dashed lines: numerical computations.

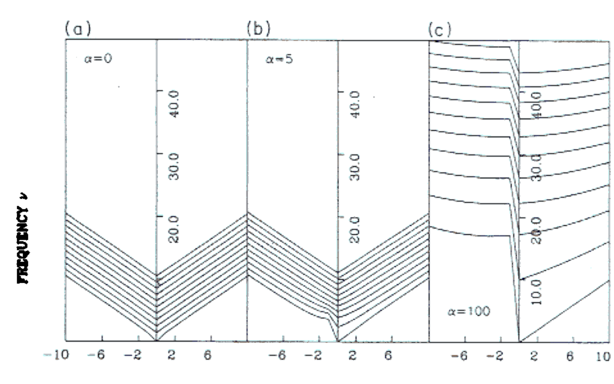

FIG. 5. Enhanced east-west asymmetry of low-frequency gravity waves: westward branches of low gravity modes are shifted upwards with increasing α. Numerically computed, nondimensional gravity frequencies for $0\le N_0\le 10$. (a) For Lamb parameter $\alpha=0$. (b) For Lamb parameter $\alpha=5$. (c) For Lamb parameter $\alpha=100$.

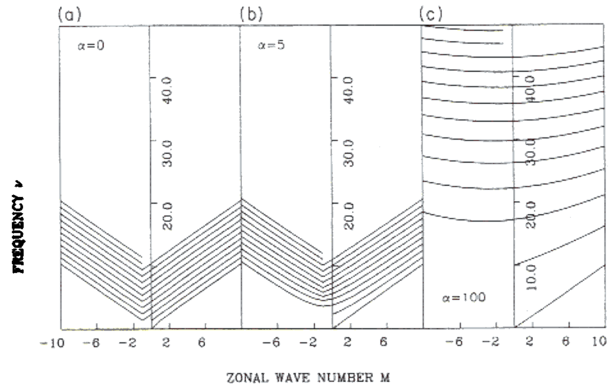

FIG. 6. As Fig. 5, but modes labeled by $N_2 = N_0 - \mathrm{sgn}(M)$.

proximation necessarily provides the eigenfrequencies in the N_2 representation due to the very nature of the Matsuno equation. It is furthermore emphasized that on the spherical surface only eigenfrequencies at integer values of the zonal wave number M are physically meaningful. Lines in Figs. 5 and 6 have merely an auxiliary function indicating which eigenfrequencies are considered as members of one mode. For increasing α, the N_0 representation, Fig. 5, develops a jump between $M=0$ and -1, while for frequencies $\nu \geq \alpha$ the characteristics of Fig. 2, namely, the east-west asymmetry corresponding to (4.2), are retained. The N_2 representation, on the other hand, yields the familiar β-plane diagram, Fig. 5(c), for large α. With decreasing α, this representation develops a minimum at $M=-1$. In contrast to the implications of (3.6) this minimum does not approach the axis $M=0$ nor does the steepness of eastern and western branches decrease with increasing N_2.

The comparison of Figs. 5 and 6 demonstrates the modification of the dispersion of low-frequency gravity waves with varying Lamb parameter: for larger α the westward branches not only steepen, but are translated to higher frequencies. Due to this translation eigenfrequencies with constant

$$N_2 = N_0 - \mathrm{sgn}(M)$$

can be combined into one mode. This relation between N_0 and N_2 ensures the same symmetry with respect to the equator for the eigensolutions of all members of one mode in both representations. The different east-west asymmetry of the β plane thus emerges in fact as an enhanced east-west asymmetry of the type (4.2). The reason for the upward translation of westward low-frequency gravity branches is the emergence of Rossby waves in this part of wave number space: the uniqueness of a wave prohibits the intersection of the dispersion curves of Rossby and gravity waves. The amplified east-west asymmetry of gravity waves in the β-plane regime is thus the necessary condition for the coexistence of Rossby waves and low-frequency gravity waves in the Yoshida waveguide.

Although the labeling of gravity modes by the mode number N_2 leads to good quantitative agreement in the β-plane regime, it is not quite unproblematic. In contrast to the discrete wave number M, the standard β-plane concept refers to a continuous wave number $k=M/a$ so that in particular the group velocity $u=(\partial_k \omega)_{N_2}$ as the speed of energy propagation is always well defined. In this framework, standing gravity waves on the β plane ($k=0$) exhibit a finite group velocity, while westward propagating gravity waves with wave number k_0, defined by $u(k_0)=0$, are not associated with a flow of energy. In the N_0 representation, on the other hand, a group velocity of the above form is not even defined in this domain of wave number space. Figures 5 and 6 thus primarily indicate that neither the mode number N_0 nor N_2 is appropriate over the entire wave number space. The only parameters consistent in the entire wave number space are the zonal wave number M and the degree L of the prolate angular function.

Under terrestrial conditions the dispersive properties of the β-plane regime are those of equatorial, baroclinic low-frequency waves in the ocean. These waves are of major significance in climate dynamics. The now widely known El Nino phenomenon, for instance, is essentially triggered by baroclinic Kelvin waves, generated by changing winds in the western equatorial Pacific [11].

V. CONCLUSION

With respect to dynamics as well as diagnostics, Laplace's tidal equations are the linear shallow water equations on the rotating spherical surface. These equations are a spatially 2D, covariant, and separable approximation to the covariant and nonseparable 3D problem. Covariant differentiation renders tensor analysis on the non-Euclidean spherical surface straightforward. The prolate spheroidal wave operator plays a fundamental role in the system's wave equation. Only in special cases can its latitudinal eigenfunctions be represented as simple combinations of trigonometric functions and prolate angular functions. The form of the wave equations indicates nevertheless that these functions are a key element of the general solution.

The rotating spherical surface provides an anisotropic and inhomogeneous arena for the propagation of Rossby and long gravity waves. The fundamental latitude-longitude anisotropy is due to its topology, while the Coriolis effect introduces an additional east-west asymmetry, physically similar to the asymmetry induced by the shear of a stationary, zonal flow. The joint action of rotation and a finite intrinsic phase speed of shallow waters lead to a latitudinal inhomogeneity, which confines low-frequency waves to an equatorial waveguide. Prolate angular functions represent this inhomogeneity on the finite spherical surface, where the width of the Yoshida waveguide is controlled by the Lamb parameter. As this parameter distinguishes barotropic and baroclinic phase speeds, it also permits the consistent incorporation of vertical variability into rotating, spherical shallow waters as internal variability of an otherwise 2D fluid. For linearizations around a stationary, zonal flow, the Lamb parameter captures the necessary shear effects of such flows on the spherical surface.

In the wave number space of shallow waters on the rotating spherical surface three regimes with distinct dispersion properties have been identified. For gravity waves at and above the inertial frequency the rotating sphere is globally homogeneous and an equatorial waveguide does not exist. In addition to the latitude-longitude anisotropy, these gravity waves exhibit Coriolis-induced east-west asymmetry with stepper westward branches of the dispersion diagram. Barotropic gravity waves in the ocean, generated by external tidal forces, propagate in this regime. A second, low-frequency regime of Rossby waves can be distinguished for all values of the Lamb parameter. Besides Margules's nondivergent Rossby waves and Rossby waves on the equatorial β plane, this regime includes midlatitude Rossby waves, which are frequently represented in terms of a "midlatitude β plane." The approximations obtained here avoid the inherent inconsistencies of this concept. Finally, for large values of the Lamb parameter, the β plane admits gravity and Rossby waves with frequencies well below the inertial frequency, coexisting in the Yoshida waveguide. As a consequence of the emergence of Rossby waves in the westward part of wave number space, the east-west asymmetry of low-frequency gravity waves is enhanced: westward gravity branches are not only steeper than their eastward counterparts, but are also translated to higher frequencies. Under terrestrial conditions the β-plane regime represents baroclinic low-frequency waves in the equatorial ocean. These waves are of major significance for the terrestrial climate problem, as indicated by the crucial role of baroclinic Kelvin waves in the equatorial Pacific for the El Nino phenomenon. Exact dispersion relations for strictly standing waves and for inertial waves, as well as the Margules limits, provide benchmarks for analytical approximations and numerical computations of the eigenfrequencies of linear shallow water on the rotating spherical surface.

ACKNOWLEDGMENTS

The authors wish to thank the Physical Oceanography Section of ONR for the basic support of COAPS, NASA for the support of D.M., and P. Swarztrauber for kindly providing the Swarztrauber-Kasahara code. We appreciate helpful discussions with Dr. S. D. Meyers.

APPENDIX

For curvilinear coordinates λ (longitude) and φ (latitude) on the surface of a sphere of radius a the covariant metric is given as

$$g_{mn}=\begin{vmatrix} a^2\cos^2\varphi & 0 \\ 0 & a^2 \end{vmatrix} \tag{A1}$$

and the covariant components of the Levi Cività pseudotensor assume the values

$$\epsilon_{mn}=(-1)^n\sqrt{\det g_{ab}}$$

if $m\neq n$ and vanish otherwise. The nonvanishing values of the Christoffel symbols are

$$\Gamma^1_{12}=\Gamma^1_{21}=-\tan\varphi\ ,$$

$$\Gamma^2_{11}=\sin\varphi\cos\varphi\ .$$

The Ricci tensor is defined as a contraction of the Riemannian

$$G_{mn}=g^{ab}G_{ambn}$$

and has for the particular metric (A1) the form

$$G_{mn}=a^{-2}g_{mn}\ .$$

The Riemannian corresponding to (A1) can be written as

$$G_{ambn}=g_{ab}G_{mn}-g_{an}G_{mb}\ .$$

The covariant derivative of a covariant vector is

$$V_m;_n=\partial_n V_m-\Gamma^a_{mn}V_a$$

and of a contravariant vector

$$V^m;_n=\partial_n V^m+\Gamma^m_{na}V^a\ .$$

For the covariant derivative of a mixed tensor one has

$$T^m{}_{n;a}=\partial_a T^m{}_n-\Gamma^b_{an}T^m{}_b+\Gamma^m_{ab}T^b{}_n\ .$$

The Laplacian of a scalar is

$$\Delta S=(\partial^n S);_n=(\partial_n+\Gamma^m_{mn})\partial^n S$$
$$=a^{-2}(\cos^{-2}\partial^2_\lambda+\partial^2_\varphi-\tan\varphi\partial_\varphi)S$$

and its eigenfunctions are the spherical harmonics. The twofold covariant differentiation of a vector

$$V_a;_{mn}=V_a;_{nm}+G_{bamn}V^b$$

does not commute, since the Riemannian does not vanish. The Laplacian of a covariant vector is

$$g^{ab}V_n;_{ab}=(\partial_a+\Gamma^b_{ba})\partial^a V_n-2\Gamma^a_{bn}\partial^b V_a$$
$$+g^{ab}(\Gamma^c_{nb}\Gamma^m_{ac}+\Gamma^c_{ab}\Gamma^m_{nc}-\partial_b\Gamma^m_{na})V_m$$

with longitudinal component

$$g^{ab}V_{1;ab}=a^{-2}(\cos^{-2}\varphi\partial^2_\lambda+\partial^2_\varphi+\tan\varphi\partial_\varphi+1)V_1$$
$$-2a^{-2}\tan\varphi\partial_\lambda V_2$$

and latitudinal component

$$g^{ab}V_2;_{ab}=a^{-2}(\cos^{-2}\varphi\partial^2_\lambda+\partial^2_\varphi-\tan\varphi\partial_\varphi-\tan^2\varphi)V_2$$
$$+2(a\cos\varphi)^{-2}\tan\varphi\partial_\lambda V_1\ .$$

The gradient of a spherical harmonic

$$E_n=\partial_n Y^M_L$$

is an eigenvector of the vector Laplacian with eigenvalue $E=[L(L+1)-1]/a^2$. With the gradient of a spherical harmonic the rotated gradient

$$E_n=\epsilon_{mn}\partial^m Y^M_L$$

is also an eigenvector. For the non-Euclidean geometry of the spherical surface the gradient of the divergence of a vector becomes

$$V^a;_{an}=g^{ab}(V_n;_{ab}+\epsilon_{na}\epsilon^{rs}V_s;_{rb})-G_{an}V^a\ . \tag{A2}$$

[1] P. Laplace, Mem. Acad. R. Sci. Paris 75 (1778) [*Oeuvres*, (Gauthier-Villars, Paris, 1893)], Vol. 9.
[2] M. Margules, Sitzungsber. Kaiser Akad. Wiss. Wien Math. Naturwiss. **102**, 11 (1893).
[3] S. Hough, Philos. Trans. R. Soc. London Ser. A **191**, 139 (1898).
[4] T. Matsuno, J. Meteorol. Soc. Jpn. **44**, 23 (1966).
[5] K. Yoshida, J. Oceanogr. Soc. Jpn. **15**, 159 (1959).
[6] M. Longuet-Higgins, Philos. Trans. R. Soc. London Ser. A **262**, 511 (1968).
[7] D. Müller, B. Kelly, and J. O'Brien, Phys. Rev. Lett. **73**, 1557 (1994).
[8] J. Pedlosky, *Geophysical Fluid Dynamics* (Springer, New York, 1983).
[9] A. Gill, *Atmosphere-Ocean Dynamics* (Academic, New York, 1982).
[10] D. Adamec and J. O'Brien, J. Phys. Oceanogr. **8**, 1050 (1978).
[11] A. Busalacchi and J. O'Brien, J. Geophys. Res. **86C**, 10 901 (1981).
[12] G. Jacobs *et al.*, Nature **370**, 360 (1994).
[13] L. Landau and E. Lifshitz, *Classical Field Theory* (Pergamon, Oxford, 1972).
[14] J. Simmonds, *A Brief on Tensor Analysis* (Springer, New York, 1993).
[15] D. Müller, Phys. Rev. A **45**, 5545 (1992).
[16] C. Flammer, *Spheroidal Wave Functions* (Stanford University Press, Stanford, 1957).
[17] P. Swarztrauber and A. Kasahara, SIAM J. Sci. Stat. Comput. **6**, 464 (1985).
[18] *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. Stegun (Dover, New York, 1970).