Breaking of balanced and unbalanced equatorial waves

F. Bouchut
Département de Mathématiques et Applications, Ecole Normale Supérieure, 45, rue d’Ulm, 75005, Paris, France

J. Le Sommer and V. Zeitlin
Laboratoire de Météorologie Dynamique, Ecole Normale Supérieure LMD, 24, rue Lhomond, 75005, Paris, France

(Received 8 October 2004; accepted 16 December 2004; published online 8 March 2005)

A clear-cut signature of a wave-breaking event is irreversible modification of the mean flow. In this paper, we provide examples of different breaking mechanisms and show that breaking scenario of equatorial waves in the β-plane shallow water model is determined by the degree of balance between the zonal component of the Coriolis force and the pressure gradient. Our analysis is based on a specially designed numerical method which guarantees two essential conditions to simulate nonlinear equatorial waves: (i) the scheme converges toward weak solutions including shocks and (ii) preserves the steadiness of balanced stationary solutions. This allows for accurate diagnostics of Lagrangian invariants of motion such as passive tracer density or potential vorticity. For unbalanced waves, the lack of balance leads to shock formation in finite time. In shock fronts, the variation of the dissipation rate induces a nonadvective potential vorticity flux and violates the local potential vorticity conservation valid for smooth solutions. This dissipative breaking mechanism is generic for unbalanced waves and is associated with enhanced mixing. For long, balanced (Rossby) waves, breaking consists in appearance of recirculation regions. It results in the formation of propagating patterns, the equatorial modons, which trap fluid particles. Such breaking occurs during the propagation of Rossby wave packets with positive geopotential anomaly and is strengthened by decreasing fluid depth. The modons are robust and collide quasielastically with Kelvin waves.


A rich variety of specific equatorial waves exist in the tropical atmosphere and ocean. Their dynamics may be studied in the framework of a rotating shallow water model, which is of frequent use in tropical meteorology and oceanography. Below we study essentially nonlinear dynamics of equatorial waves. Its most striking manifestation is wave-breaking leading, on the one hand, to the formation of shocks by geostrophically unbalanced waves, and, on the other hand, to the formation of modons by geostrophically balanced waves (the geostrophic balance is approximate equilibrium between the pressure and the Coriolis forces). The wave-breaking modifies the background flow, changes its transport properties, and induces mixing.

I. INTRODUCTION

In the tropical atmosphere and in the equatorial ocean, the change of sign of the Coriolis force at the equator is known to yield a specific wave activity in the so-called equatorial wave guide (Ref. 1, p. 440). Since the pioneering work of Matsuno, who first obtained a complete description of the spectrum of equatorial waves in the linear β-plane shallow water model, these waves have been extensively studied. In the ocean, equatorial wave dynamics is known to play an important role in El Niño phenomenon (see, e.g., Delcroix et al., 3 and Boulanger and Menkes). In the atmosphere, such waves have been tentatively associated with the mesoscale variability of convection in the tropics and with the propagation of the tropospheric Madden-Julian oscillation whereas their propagation and dissipation in the stratosphere is the key element of the stratospheric quasi-biennial oscillation. It is worth noting that, both in the atmosphere and in the ocean, the linear shallow water theory fits the observations to a good extent.

The rotating shallow water equations may be obtained from the full primitive equations by vertical averaging and by neglecting vertical correlations of the hydrodynamical fields (see the classical works of Jeffries and Obukhov). Hence, the results of this model should be interpreted in terms of vertically averaged fields.

Nonlinear corrections to the equatorial wave dynamics were first investigated independently by Boyd and Ripa in a series of papers in the 1980s. On the basis of decomposition in the eigenfunctions of the linear problem, Ripa proposed to file nonlinear coupling mechanisms among equatorial waves in two categories:

- **Resonant** triadic interaction is a weakly nonlinear mechanism which can be predicted directly from the dispersion relation.
- **Off-resonant** interactions are essentially nonlinear and require finite amplitudes to occur.

---

3Electronic mail: zeitlin@lmd.ens.fr
In what follows, we shall use this classification.

Resonant interactions among harmonic waves were classified by Ripa. Boyd studied specific resonance mechanisms including second harmonic resonance and long wave/short wave resonance. Since a single equatorial wave is not an exact solution of the nonlinear equations, even a system with a single harmonic wave is subject to such interactions, as noticed by Ripa. Off-resonant interactions also play an important role in the equatorial shallow water model. These essentially nonlinear interactions lead to the appearance of coherent patterns. First, Boyd showed that the system can support the propagation of solitary Rossby waves which involve a balance between the Coriolis force and the pressure gradient, namely, the geostrophic balance. Second, Boyd and Ripa found that Kelvin waves can break in finite time.

High resolution numerical simulations of these two phenomena were performed only recently. Boyd investigated the stability of solitary Rossby waves. Fedorov and Melville showed that the breaking of a Kelvin wave results in a moving discontinuity so that it leads to the formation of a quasisteady pattern called a Kelvin front. Le Sommer et al. found that such localized wave patterns (Kelvin fronts and solitary Rossby waves) appear spontaneously during the adjustment of localized perturbation close to the equator.

In the present paper, we show that off-resonant interactions not only give rise to coherent patterns but also lead to irreversible modifications of the mean flow, which will be taken as definition of wave-breaking in what follows. A mean flow characteristic of interest is the potential vorticity distribution and a useful indication of the mean flow properties is a passive tracer distribution. [As at midlatitude \( f \) plane, the potential vorticity is a Lagrangian invariant of motion whose distribution constrains the evolution of the slow, geostrophically balanced part of the flow, see Sec. II A.] To our knowledge, none of the previous studies have examined the breaking processes over the whole range of wave modes. The purpose of the present paper is to investigate numerically the effect of essentially nonlinear interactions on each equatorial wave mode in order to answer the following questions:

- What is the route to breaking for a given wave mode?
- How mean flow is modified by the wave breaking?
- What is the effect of wave breaking on the transport and mixing of tracers?

Such study requires a numerical tool which would handle properly both shock fronts and balanced quasi-steady Rossby waves. More precisely, we need to compute with a good accuracy both solutions with discontinuities and solutions which are geostrophically balanced. This justifies the special care we pay to designing a numerical method which is the central point of our study.

Being constructed by vertical averaging of the full three-dimensional primitive equations, the model cannot reproduce the vertical structure of the breaking events. Nevertheless, its results are compatible with the solutions of the full primitive equations in what concerns the potential vorticity redistribution, cf. Epifanio and Durran for a recent discussion. Our results, thus, give information on horizontal location of strong mixing events due to breaking, on the resulting redistribution of potential vorticity, and on the formation of secondary circulations.

The paper is organized as follows: Sec. II is devoted to the presentation of the equatorial shallow water model and of the spectrum of linear equatorial waves. The numerical method is described in detail in Sec. III. In Sec. IV, we illustrate shock-formation mechanisms in two limiting cases which provide useful insight on the resulting nonadvective potential vorticity flux. The ubiquity of the shock formation as a result of the breaking process for unbalanced waves is demonstrated in Sec. V. Finally, we study the breaking of balanced Rossby waves and the role of solitary Rossby waves in this process in Sec. VI. Some additional material is provided in the Appendixes including the quasielastic interaction between a solitary Rossby wave and a Kelvin front.

II. BASIC EQUATIONS AND SPECTRUM OF EQUATORIAL WAVES

A. Model equations

Throughout this study, we work with the equatorial \( \beta \)-plane shallow water equations. This set of equations describes the evolution of a layer of incompressible fluid of variable depth \( h \) and horizontal velocity \((u, v)\) in a rotating frame:

\[
h_t + uh_x + vh_y + h(u_x + v_y) = 0, \quad u_t + uu_x + vu_y + gh_x = \beta y v,
\]

\[
v_t + uv_x + vu_y + gh_y = -\beta y u, \tag{1}
\]

where \( g \) is the acceleration due to gravity and \( f(y) = \beta y \) (Coriolis parameter) is twice the rotation rate. For a derivation of this model and its basic properties, see Gill’s textbook. This model can be derived for both the ocean and the atmosphere. In the oceanic case, the effect of wind forcing results in an equivalent topography term which is to be added to the right-hand side of Eq. (1) (Ref. 1, p. 435, and Ref. 24).

The model (1) is taken as a physically plausible nonlinear model although it has intrinsic limitations in what concerns the vertical structure of the flow (cf. Ripa). If linearized, this set of equations is equivalent to Matsumo’s model. See Ripa for a detailed discussion of the relevance of nonlinear terms.

If \( H \) is a typical scale for the fluid depth \( h \), we can introduce intrinsic velocity and length scales

\[
c = \sqrt{gH} \quad \text{and} \quad R_d = \frac{c}{\beta}. \tag{2}
\]

They are, respectively, the Kelvin wave speed and the equatorial Rossby radius. It is to be noticed that the set of Eqs. (1) is formally equivalent to gas dynamics equations in a rotating frame, with specific heat ratio \( \gamma = 2 \). Therefore \( c \) can be considered as an equivalent sound speed and \( h \) as an equivalent density.
Hereafter, for a given wave field, taking $\Delta H$ as maximum height deviation, we introduce the nondimensional amplitude

$$e(t) = \frac{\Delta H}{H} = \frac{\max_{x,y,t} h(x,y,t) - \min_{x,y,t} h(x,y,t)}{H}. \quad (3)$$

**B. Reminder on equatorial waves**

As noticed in the Introduction, the latitudinal variation in the rotation rate yields a particular wave spectrum for equatorial waves. Let us briefly remind the basic properties of equatorial waves. The reader is referred to the recent textbook by Majda$^{25}$ for a comprehensive introduction to equatorial waves.

Considering small perturbations over a rest state $h = H$, and looking for solutions $\approx \exp ik \left[ (x-y) \omega - \frac{x^2}{2} \right]$ that vanish at $y \rightarrow \pm \infty$, an infinite set of equatorially trapped waves is found (Ref. 1, p. 434). First, Kelvin waves (KW) are solutions such that $u = 0$. Then, other wave modes are described using an ad hoc projection onto parabolic cylinder functions (see Appendix B). A synthetic dispersion relation accounts for the whole set of equatorial waves

$$\left( \frac{\omega}{c} \right)^2 - \kappa^2 - \frac{k^2}{\omega} = \frac{2(n + 1)}{c} \beta,$$  \quad (4)

where $n$ refers to the corresponding latitudinal structure given by the expansion in parabolic cylinder functions. This leads to the standard classification of equatorial wave modes

- $n = -1$: eastward Kelvin waves (KW),
- $n = 0$: eastward Yanai waves (EYW), westward Yanai waves (WYW),
- $n \geq 1$: westward Rossby waves (RWN), eastward inertia-gravity waves (EIGWn), westward inertia-gravity waves (WIGWn).

The spectrum of equatorial waves is summarized in Fig. 1 and the corresponding spatial structures are presented in Fig. 2 for $n = 1$. Note that below we restrict our study to low meridional quantum numbers $n \leq 1$, although most of the results of this paper can be extended to higher $n$.

It is to be noted that the same wave types have different spatial structures depending on the direction of their phase propagation. In what follows, we will show that the nonlinear evolution of a wave strongly depends on its spatial structure. More precisely, an important criteria for the breaking scenario turns to be the degree of balance of the zonal pressure gradient by Coriolis force.

**C. Imbalance parameter**

The notion of geostrophic balance, i.e., the balance between the Coriolis and the pressure forces is one of the keystones of geophysical fluid dynamics. In fact in midlatitudes the synoptic-scale motions happen to be almost balanced, which greatly simplifies their study. For one-dimensional structures, geostrophically balanced configurations are steady. As we show below, at the equator this notion is of primary importance too.

In order to characterize a degree of imbalance of equatorial waves, we introduce, for a given wave, the geostrophic meridional velocity $v_g = g \partial_x h / \beta y$ and we define an imbalance parameter

$$\Lambda = \frac{|v(x,y) - v_g(x,y)|}{|v_g(x,y)|}. \quad (5)$$

It turns out that the value of $\Lambda$ is not very sensitive to the choice of position $(x,y)$, that is why we do not take spatial mean in Eq. (5). This parameter takes values between 0 and $+\infty$ and measures the degree of balance in the $x$ direction at a given location $(x,y)$. Being almost independent on location, it can be used to quantify the degree of zonal balance of each wave mode. Therefore we define a balanced wave to be such that $\Lambda \rightarrow 0$, so that $\beta y v \sim g \partial_x h$. Reversely, we call unbalanced the waves such that $\Lambda \gg 1$. Note that the imbalance parameter $\Lambda$ does not depend on the wave amplitude.

It is easy to verify that for a Kelvin wave $\Lambda_K = 1$. Such a wave is therefore the basic example of an essentially unbalanced wave. The most striking is the fact that, for Rossby waves RW1, RW2, and RW3, $\Lambda_R$ is zero so Rossby waves are essentially balanced waves. [We have checked this result with Mathematica up to $n = 3$ with several digits of accuracy.] We nonetheless expect this strict balance to break down at higher $n$.

One needs to go in deeper detail in order to classify EIGW, WIGW, EYW, and WYW. We therefore compute the numerical values of $\Lambda$ for these waves. We solve Eq. (4) numerically and compute $\Lambda$ for each wave. On the one hand, we obtain that for inertia-gravity waves (EIGW and WIGW1) $\Lambda > 1$ for any $k$ and any $y \in [-5R_a, 5R_a]$. These waves are therefore unbalanced. On the other hand, polarization relations between $h, u,$ and $v$ for Yanai waves (EYW and WYW) with a given $k$ imply that $\Lambda_Y$ does not depend on $(x,y)$. Figure 3 shows the variation of $\Lambda_Y$ as a function of $k$. EYW have $\Lambda_Y > 1$ and are therefore unbalanced. The situa-
tion is more ambiguous in what concerns WYW. WYW with small wavelengths are balanced whereas large-scale WYW are unbalanced.

Hence, there exist both balanced and unbalanced waves in the model. Moreover, as mentioned in the Introduction, unbalanced Kelvin waves can form discontinuities. Therefore, developing a numerical method in order to study breaking of equatorial waves requires special care for both types of motion. In the following section, we show how this could be done.

FIG. 2. Geopotential anomaly and horizontal velocity associated with some equatorial waves. lower panel: westward traveling waves, upper panel: eastward traveling waves
III. DESCRIPTION OF THE NUMERICAL METHOD

A. Required properties

Based on the preceding remarks, we discuss in this section two properties which turn to be essential in numerical simulations of nonlinear equatorial waves. First, off-resonant interactions between equatorial waves can lead to shock formation. Through shocks, like in compressible gas dynamics, mass and momentum are conserved and energy is dissipated. Such conditions yield a unique weak solution. There are two alternative ways to obtain this unique solution. On the one hand, one can require four cross-shock conditions to be verified through a discontinuity moving at speed $c_w$. Namely, the Rankine-Hugoniot conditions

$$-c_n[h] + [hu_n] = 0,$$

$$-c_n[hu_n] + \left[ hu_n^2 + \frac{1}{2}gh^2 \right] = 0,$$

$$-c_n[hu_n] + [hu_nu_n] = 0,$$  \hspace{1cm} (6)

where the brackets denote the jumps of corresponding quantities across the shock, $u_n$ ($u_s$) is the cross-front (along-front) velocity and the entropy condition which states that the energy must be dissipated across a shock

$$-c_n[\eta] + [(\eta + gh^2)/2]u_n \leq 0,$$

with $\eta = \frac{h}{2}(gh + u_n^2 + u_s^2).$  \hspace{1cm} (7)

On the other hand, one can consider the limit solution at vanishing viscosity of a sequence sets of conservative equations of the form (A1) with dissipation included in the right-hand side (r.h.s.) As in gas dynamics, both approaches are equivalent since they lead to the same unique solution. At this point, we can express the first required property for the numerical method: (i) The scheme should converge toward weak solutions with increasing resolution. Second, in what concerns balanced solutions, we require a minimal property of steady state preservation, namely (ii) Steady, geostrophically balanced initial conditions should keep steady at a discrete level. In what follows, we propose a method which guarantees both conditions (i) and (ii). Property (i) is verified within the framework of multidimensional finite volume techniques (see Appendix A) with an appropriate choice of Riemann solver. Property (ii) is achieved by a special treatment of Coriolis source terms.

B. Source-term treatment

Hereafter, we describe the strategy we use in order to handle Coriolis source terms in Eq. (1) within the framework of finite volume techniques, so that property (ii) is verified at a discrete level. A brief reminder of the basic notations and concepts of finite volume methods for homogeneous conservation laws is given in Appendix A. The only subtlety consists in treatment of the r.h.s. of Eq. (1). We show that it could be reduced to the treatment of local, unidimensional topographic source terms for which efficient solvers have been recently developed (see Ref. 26).

The first step is to split the source term in two contributions in order to fit the dimensional splitting approach (see Appendix A). Writing Eq. (1) in conservation law formulation, we get Eq. (A1) with a r.h.s. of the form $S(U) \equiv (0, \beta yhu, -\beta yhu)$. The source term is then split into two contributions: $S=S_1+S_2$ with $S_1=(0, \beta yhu, 0)$ and $S_2=(0, 0, -\beta yhu)$. A this point, the scheme only needs to compute Riemann problem solutions (see Appendix A) for the following one-dimensional system:

$$\partial h + \partial_x(hu) = 0,$$

$$\partial(hu) + \partial_x(hu^2 + gh^2/2) = \beta yhu,$$

$$\partial(hu) + \partial_x(huv) = 0.$$

The third equation in Eq. (8) is decoupled from the first two equations.

The second step is to transform the r.h.s. of Eq. (8) using the apparent topography method introduced by Bouchut et al.\textsuperscript{26} We therefore introduce an apparent topography in the $x$-direction $Z(x)/g$ such that $\partial_x Z = -\beta y v$. Hence, approximating the Riemann problem solution for Eq. (8) simply requires to solve a Riemann problem for

$$\partial h + \partial_x(hu) = 0,$$

$$\partial(hu) + \partial_x(hu^2 + gh^2/2) + h\partial_x Z = 0,$$  \hspace{1cm} (9)

together with an advection equation for $u$. At a discrete level, $Z_i$ is chosen such that

$$Z_{i+1} - Z_i = -\Delta x\beta y_{i+1/2}(v_i + v_{i+1/2})/2.$$  \hspace{1cm} (10)

It should be noticed that, although the r.h.s. of Eq. (1) is not a gradient, Coriolis source terms can be treated by means of two potentials $Z^x(x)/g$ and $Z^y(y)/g$ at a discrete level. The reasons are that each direction is treated separately and that finite volume schemes deal with approximate solutions of Riemann problems instead of exact nonlinear solutions. Noticeably, verifying property (ii) now only requires us to use a scheme capable of computing the rest flows over topography in the shallow water model, a problem where considerable
progress in numerical methods was recently achieved (see, e.g., Ref. 26).

C. Well-balanced Riemann solver

We now explain how we compute solutions of the Riemann problem for system (9). Since we have described the dimensional splitting approach and the reduction of the source term in the form (9), we focus hereafter on the one-dimensional problem. Riemann problem solutions can be approximated within the framework of Appendix A by using fluxes which are discontinuous at the cell interface.25 In this way, we obtain the following update formula:

$$U_{i+1}^{t+1} = U_i^t - \frac{\Delta t}{\Delta x} \left( F_{i+1/2} - F_{i-1/2} \right), \quad (11)$$

where we have distinguished between the left and right numerical fluxes $F_l$ and $F_r$ so that:

$$F_{i+1/2}^{+} = F_r(U_i, U_{i+1}, Z_{i+1} - Z_i),$$

$$F_{i+1/2}^{-} = F_l(U_i, U_{i+1}, Z_{i+1} - Z_i). \quad (12)$$

The Riemann solver now computes fluxes at each cell interface. In this framework, property (i) is fulfilled if the fluxes $(F_r, F_l)$ are conservative and verify a discrete entropy relation. Property (ii) is verified if $(F_r, F_l)$ are well-balanced as defined by LeRoux and Greenberg.28 Several numerical schemes can combine such properties. The Riemann solver used in this paper is based on two ingredients: the hydrostatic reconstruction procedure and a relaxation Riemann solver.

The first ingredient of the procedure follows Audusse et al.26 who describe a general strategy, based on a local hydrostatic reconstruction, that allows them to derive a well-balanced scheme for system (9) from any Riemann solver of the homogeneous problem. In addition, the hydrostatic reconstruction method of Audusse et al.26 preserves the non-negativity of $h$. The second ingredient is a Riemann solver for the homogeneous system (A4) which verifies a discrete entropy relation. We choose to use a relaxation Riemann solver as proposed by Bouchut.29 An introduction to relaxation schemes used as approximate Riemann solvers is given by LeVeque and Pelanti.30 Our approximate Riemann solver is obtained from the exact solution of the Riemann problems for the relaxation system:

$$h_t + (hu)_x = 0,$$

$$(hu)_t + (hu^2 + \pi)_x = 0,$$

$$(h\pi u^2)_t + (h\pi u^3)_x + u_x = 0,$$

$$(hs)_t + (hsu)_x = 0, \quad (13)$$

where $\pi$ is initialized with $\pi_l = gh_l^2/2$, $\pi_r = gh_r^2/2$ for the left and right states, respectively. Under a subcharacteristic condition, this system can be used to approximate solutions of Eq. (A4).29 Now the eigenvalues of Eq. (13) are linearly degenerate (see Ref. 31, p. 283) so that it is straightforward to compute solutions of Riemann problems. [A system of hyperbolic PDE is said to be linearly degenerate if each eigenvalue of the Jacobian matrix $\partial_t F(U)$ is constant along integral curves of the corresponding eigenvector. In such a case the solution of the Riemann problem consists in discontinuities propagating at constant speed.] The numerical flux $F$ is given by the solution of the Riemann problem for system (13), so that $F= (hu, hu^2 + \pi)$ evaluated along the line $(x/t=0)$. It can be proven that this relaxation Riemann solver verifies a discrete entropy relation. Moreover, variable $s$ has been introduced to handle situations of drying which happen when $h \rightarrow 0$. Since the hydrostatic reconstruction preserves non-negativity of $h$, the resulting nonhomogeneous scheme also handles initial conditions with $h=0$ (which were not considered in this paper).

Hence, the nonhomogeneous Riemann solver built in this way is by construction well-balanced, conservative, and verifies a discrete entropy relation so that properties (i) and (ii) are guaranteed for the whole two-dimensional scheme.

D. Practical implementation and summary of the basic properties of the method

The whole method we have presented so far is of the first order both in space and time. It is possible to extend this method to higher order by using a higher order spatial reconstruction method in space and a higher order update formula in time, see Ref. 31. We choose a second order minmod slope limiter technique in space and the second order Heun method in time. Moreover, the use of the finite volume techniques allows the implementation of periodic, radiation or wall boundary conditions and their extension to higher orders. In the simulations presented in what follows we use periodic boundary conditions in the $x$ direction and radiation conditions in the $y$ direction, except for the experiment presented in Fig. 15, where radiation boundary conditions are used in both directions.

In Sec. VI B, we use a Lagrangian tracer $a$ such that $\partial_t a + u \partial_x a + v \partial_y a = 0$. In order to compute its evolution with the same accuracy as the dynamical variables $(h, hu, hv)$, we introduce an associated conservative quantity $A = ha$, which verifies $\partial_t A + \partial_x A + \partial_y A = 0$. The latter equation is added to the complete set of Eq. (1). As a result, a linearly degenerate eigenvalue is added to the original system but it does not modify the evolution of the dynamical variables $h$, $hu$, and $hv$. In this paper we focus on transport due to breaking rather than mixing processes. Therefore, we did not add any explicit tracer diffusion operator in the scheme though it could be done in principle. Finally, an additional advantage of the

<table>
<thead>
<tr>
<th>Table I. Basic properties of the numerical procedure.</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
</tr>
<tr>
<td>b.</td>
</tr>
<tr>
<td>c.</td>
</tr>
<tr>
<td>d.</td>
</tr>
<tr>
<td>e.</td>
</tr>
</tbody>
</table>
whole procedure is a possibility to include real topography straightforwardly. It is sufficient (see Sec. VI B) to add the real topography to the apparent one.

A summary of the properties of the numerical procedure is given in Table I. A detailed description of our scheme may be found in Bouchut, Le Sommer, and Zeitlin. An interested reader will also find there the advanced comparisons of the numerical schemes of this kind with the standard ones.

The results we show hereafter are computed within a $200 \times 200$ horizontal grid. For localized wave packet simulations (see Sec. V B) zonal resolution was increased up to 600 points.

We apply three types of initial conditions and the corresponding parameters are given in the caption for each plot. For simulations of harmonic waves, we use horizontal structures as given by Gill. Nondimensional initial amplitude $\epsilon$ and wave-number $k$ are given in captions. The procedure applied to obtain initial fields for localized wave packets is adapted from Ripa and described in Appendix C. For each simulation we give the nondimensional amplitude $\epsilon$ and the peak wave-number $k_p$. Finally, the Rossby solitons simulations (Sec. VI B) were initialized with the fields obtained by Boyd. We give parameters according to the notation of this paper in Appendix F.

### IV. UNBALANCED WAVES: SHOCK FORMATION IN TWO LIMITING CASES

For unbalanced waves, breaking consists in shock formation. We illustrate in this section the mechanism and consequences of this process in two limiting configurations. We show how Kelvin waves break due to the lack of meridional velocity and how ultra-long inertia-gravity waves and Yanai waves break due to the lack of zonal pressure gradient. Before proceeding, let us briefly remind the effect of dissipation in the shock fronts.

#### A. Wave-mean flow interaction

As noticed by Buhler, a redistribution of the potential vorticity (PV)

$$q = \frac{\beta y + v_s - u_s}{h}$$  

(14)

is a convenient indicator of dissipative wave-mean flow interactions. Without dissipation, the PV is a Lagrangian invariant of motion ($d_q = 0$) which varies only due to advective fluxes. But, as mentioned previously, shocks induce dissipation which should yield a modification of the PV distribution.

A description of the wave-mean flow interaction in shallow water flows is possible within the “generalized Lagrangian-mean” (GLM) formalism. This work gave evidence of the equivalence between momentum flux deposition and PV redistribution. But, although the fully nonlinear GLM computation is possible, in principle, there are practical difficulties of extending the pseudomomentum definition to the $\beta$-plane (see Ref. 33).

Alternatively, following Lighthill, the dissipation rate across the shock is equal to the jump of the Bernoulli function relative to the shock

$$\left[ \vec{B} \right] = -\frac{\beta}{4} \frac{h^3}{h_r}$$  

(15)

where $\vec{B} = gh + [(u_n - c_n)^2 + u_s^2]/2$. The subscripts $f$, $r$ refer to the front or rear state of the shock, respectively, and the subscripts $n$, $s$ mean normal and parallel to the shock directions. By combining Rankine-Hugoniot conditions (6), one can obtain the jump in potential vorticity across a shock

$$[q] = -\frac{\partial [\vec{B}]}{h(u_n - c_n)}.$$  

(16)

This PV jump indicates that fluid particles change their PV as they pass through the shock front. The resulting PV redistribution is the consequence of a nonadvective PV flux along the shock front. A PV redistribution indicates a significant wave-mean flow interaction. A crucial feature in the wave-mean flow interaction due to shocks is its essential dependence on the along-front variation of the dissipation rate, cf. Eq. (16). Such dependence is crucial for the wave-mean flow interaction during Kelvin wave breaking.

#### B. Kelvin wave breaking

For a Kelvin wave, $v = 0$ so that $\Lambda = 1$. Therefore there is no Coriolis force to counterbalance nonzero zonal pressure gradient. This leads to shock formation.

This mechanism was pointed out by Boyd and Ripa. More precisely both authors argued that the zonal structure of the zonal velocity associated with a Kelvin wave evolves according to a simple wave equation which is known to support discontinuous solutions [see, e.g., Lighthill (Ref. 34, p. 151)] with discontinuities forming in finite time.

Ripa was convinced that this was an indication of the failure of his analytical approach. He concluded that some IGW should be emitted before a discontinuity forms. But the numerical simulations by Fedorov and Melville showed that both phenomena occur: The evolution of a finite amplitude Kelvin wave packet leads to shock formation in finite time. The shock front is then propagating at a different speed than the linear Kelvin wave. This in turn can excite resonantly EIGW. Moreover, Le Sommer et al. pointed out the distinction between positive and negative mass anomalies in this context. A Kelvin wave with positive mass anomaly breaks at its front whereas a negative mass-anomaly Kelvin wave breaks at its rear.

In Fig. 4, we plotted successive snapshots of the evolution of a harmonic Kelvin wave with $k = 1$ and $\epsilon = 0.5$. A shock front is visible at time $t = 6.0$. As noted above, the dissipation across the shock front induces a nonadvective potential vorticity flux along the shock. In our simulation, its effect is to bring PV anomaly from the southern hemisphere to the northern hemisphere which is illustrated in Fig. 5.

Kelvin wave breaking induces a significant interaction between the wave field and the mean flow. It should be stressed that the PV redistribution is an irreversible process: Since in Kelvin fronts dissipation is always maximum at the equator, the resulting nonadvective PV flux is systematically oriented northward. A Kelvin wave can deposit only easterly momentum. Another effect is to be expected in the presence
of Kelvin fronts. If a diffusive tracer was added, concentrated gradients in the shock should enhance mixing of tracer along the equator. However, this effect is beyond the scope of the present paper.

C. Long IGW and YW breaking

Another interesting limiting configuration is the case of IGW or YW with $k \rightarrow 0$. For those waves, the zonal pressure gradient tends to zero whereas the meridional velocity is nonzero so that $\Lambda \rightarrow \infty$. This lack of balance also leads to shock formation. An illustration of such a behavior is given in Fig. 6 for a standing Yanai wave with $k=0$. In the linear regime, this wave is oscillating with frequency $\omega = \sqrt{c\beta}$ with a meridional structure corresponding to a parabolic cylinder function with $n=0$. In the finite amplitude case, we can see that shocks are emitted and propagate northward and southward. Due to the increase of Coriolis parameter $f(y)$ with latitude, the amplitude of these shocks is rapidly decreasing. This amplitude decrease is similar to that observed by Bouchut et al. in the context of geostrophic adjustment of frontal structures on the $f$ plane. Nevertheless, the overall situation is different here. First, the original large-scale wave disturbance is trapped around $y=0$ and we do not expect a steady adjusted state to be reached. As a consequence, small-scale wave emission and shock formation are not restricted to a short adjustment stage but, rather, are taking place periodically. Second, shocks are decreasing more rapidly because $f$ is not constant.

Figure 7 shows a similar behavior for IGW with $k=0$. In the linear limit, we expect an oscillating standing wave with a symmetric meridional structure given by a parabolic cylinder function with $n=1$, a behavior which is observed for small $\epsilon$. For finite amplitude, we see that the standing wave pattern is periodically emitting small scale waves which form shocks. It is interesting to note that small scale wave emission undermines the expansion in parabolic cylinder functions. Whereas this expansion is useful to describe the
linear regime, its use to describe finite amplitude waves would require a large number of modes for convergence. In particular, an expansion of the wave field obtained at time $t=8.0$ in Fig. 7, would involve components with $n \gg 2$.

FIG. 6. Time evolution and breaking of a YW with $k=0$ which illustrates shock formation due to a lack of geostrophic balance in zonal direction. Initial amplitude is $\max(v)=0.5c$. $h$ is plotted as a function of $y/R_d$ at successive times, since $k=0$. A reference line obtained in the linear regime where amplitude $\max(v)=0.01c$ is overlayed (dotted line).

FIG. 7. Same as 6 for an IGW with $k=0$. 

© 2005 American Institute of Physics. Downloaded 09 Mar 2005 to 194.254.66.226. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp
In what concerns wave-mean flow interaction, it should be stressed that, since all the fields are invariant in the z direction, so are shocks. Therefore, the jump amplitude does not vary and no PV deposit is expected which is checked in numerical simulations (not shown). But, since shocks are formed periodically around the same latitude, we expect a significant mixing to occur in those regions.

It should also be noted that since the initial amount of energy is fixed for a given wave and since energy is dissipated through shocks, wave emission and shock formation should be transient processes. Therefore, the following questions remain:

• What is the limiting amplitude for the initial condition such that no shocks are formed?
• For a given amplitude, how long does the shock formation last?
• What part of the available energy is dissipated through shocks during the transient relaxation?
• How much of momentum flux is transported from equatorial regions to higher latitudes during the transient relaxation?

Our numerical simulations indicate the following scenario: Shocks are formed and dissipate energy effectively only beyond a certain threshold, typically \( \epsilon = 0.3 \), which depends on the wave type. Shock formation is taking place periodically until dissipation decreases the main wave amplitude so that the threshold value is reached. If this scenario is correct, the long-time evolution is compatible with the linear description.

To conclude this section, shock formation and small-scale wave emission is a generic behavior for finite-amplitude unbalanced waves. Two limiting cases have been studied in detail. In the case of a Kelvin wave a significant dissipative wave-mean flow interaction is taking place whereas in the case of long gravity waves and Yanai waves there is no PV change due to shocks. Moreover, the question of the long-term behavior of wave-emission process remains. On route, enhanced mixing is expected at the shock formation location.

V. UN BALANCED WAVES: UB IQUIT Y OF THE SHOCK FORMATION

A. Monochromatic unbalanced wave breaking

So far, we have illustrated shock formation in two limiting cases. However, this behavior is generic for any unbalanced wave with sufficiently large amplitude.

In Fig. 8, we show the time evolution of an EYW with \( k = 1/R_d \) and \( \epsilon = 0.3 \). The initial condition is antisymmetric. [The vector \((h, u, v)\) can be decomposed into a symmetric part \((h_s, u_s, v_s)\) such that \( h_s(-y) = h_s(y), u_s(-y) = -u_s(y), v_s(-y) = v_s(y) \) and an antisymmetric part \((h_a, u_a, v_a)\) such that \( h_a(-y) = -h_a(y), u_a(-y) = u_a(y), v_a(-y) = -v_a(y) \).] Noticeably, equatorial waves with even meridional quantum number \( n \) are symmetric whereas odd \( n \) waves are antisymmetric.] Its evolution exhibits a strong difference between positive and negative mass anomaly regions. This results in the concentration of the gradients and the waves break at time \( t = 5.0 \).

Finally the wave pattern takes a chevron-type shape with shocks either to the north or to the south. A striking feature of the nonlinear evolution is the growth of a symmetric contribution to the fields \((h, u, v)\) as seen in Fig. 9. This component has the spatial structure of an EIG wave with \( k = 2 \).

As noticed by Ripa,\(^{12}\) at a given time, any field \((h, u, v)\) can be decomposed using the complete orthogonal basis of equatorial wave modes. Starting from an antisymmetric EYW with \( k = 1 \), the rapid growth of a symmetric EIGW with \( k = 2 \) is a result of interaction between EYW and EIGW. Such interaction between EYW and EIGW was not predicted by the weakly nonlinear approach by Ripa,\(^{13,14}\) who classified all possible resonances in this model. Neither is it the result of a second harmonic resonance of the type discussed by Boyd.\(^{15}\)

Both authors mentioned that harmonic EYW and EIGW are stable regarding the weakly nonlinear effects. The present mechanism is therefore due to an essentially nonlinear effect unnoticed previously. We believe that it is similar to the EIGW emission by Kelvin fronts observed by Fedorov and Melville.\(^{21}\) Namely, finite amplitudes induce a change in the perturbation phase speed which allows for new interactions to take place. Such interactions, which are not observed in the weakly nonlinear limit, will herein be referred to as essentially nonlinear interactions. The framework proposed by Ripa\(^{12}\) in order to study off-resonant interactions would possibly allow a description of the EIGW emission but would require further study which is out of the scope of the present paper.

Figure 10 presents the evolution and breaking of an EIGW with \( k = 1/R_d \) and \( \epsilon = 0.3 \). As before the harmonic wave propagates and concentrates its gradients so that it breaks at time \( t = 5.0 \). At time \( t = 6.0 \), a quasisteady V shape is observed. Note the strong similarity with the Fig. 9. The only visible difference is that the dissipation rate is symmetric in Fig. 10 whereas it is not in Fig. 9. This shows that the presence of the background EYW in the previous run had modified the free evolution of the EIGW which was resonantly excited. More precisely, the EYW compensates the gradient of the EIGW and inhibits shock formation. The result is the chevron shape of Fig. 8. As a conclusion, we can describe the scenario of EYW breaking as a two-stage mechanism: first, an EIGW is excited, second the EIGW breaks and forms shocks which are compensated by EYW gradients.

Finally, for a WIGW with \( k = 2/R_d \) and \( \epsilon = 0.3 \), breaking also happens rapidly as shown in Fig. 11. The shock is now centered about the equator. Note that, for a given wavelength, shock formation requires a bigger amplitude for WIGW than for EIGW which is probably due to the asymmetry of the dispersion curves between westward and eastward propagating inertia-gravity waves.

B. Unbalanced wave packets

The results we obtained for monochromatic unbalanced waves can be extended to the case of unbalanced wave packets. In Appendix C, we present a procedure which we used to generate wave packets for a given wave mode. Subsequent simulations show that shock formation resulting from the breaking process is ubiquitous for unbalanced wave packets.
FIG. 8. Time evolution and breaking of an EYW with $k = 1/R_d$ and $\epsilon = 0.3$.

FIG. 9. Symmetric component from an initial EYW with $k = 1/R_d$ and $\epsilon = 0.3$. 
FIG. 10. Time evolution and breaking of an EIGW with $k = 1/R_d$ and $\epsilon = 0.3$.

FIG. 11. Time evolution and breaking of a WIGW with $k = 2/R_d$ and $\epsilon = 0.3$. 
Figure 12 shows a snapshot of the evolution of an EIGW packet with $\epsilon=0.3$. We can see that shocks have been formed and propagate across the whole wave pattern. A similar behavior is observed for WIGW in Fig. 13 although for a bigger wave amplitude. Interestingly, in this case $k_0$ has been chosen such that the group velocity of the wave packet is zero. Hence, shocks will modify cumulatively the potential vorticity field at the same location. This mechanism yields a strong PV deposit of almost 0.2 (not shown) at time $t=10.0$. So, shock formation is a typical behavior for unbalanced wave packets which can induce a significant wave-mean flow interaction.

It is worth noting that essentially nonlinear interactions are also effective in the case of wave packets. It is illustrated in Fig. 14 which shows the meridional velocity at the equator as a function of time for a EYW packet with $k_0=1/R_d$. Although the overall dispersion of the wave packet is only slightly modified by finite-amplitude effects, the emission of short wavelength EIGW with $k'_0=2/R_d$ is visible. Therefore, the nonlinear resonance mechanism is not restricted to the case of a monochromatic wave but is generic for EYW.

To conclude, shock formation is ubiquitous for sufficiently large amplitudes for any unbalanced wave. Such breaking results in an irreversible wave-mean flow interaction and a strong effect is expected regarding the mixing of tracer. On route to breaking, positive and negative geopotential anomaly regions evolve in a highly asymmetric manner. Positive anomalies are propagating faster than predicted by the linear theory while negative anomalies are propagating slower. This difference in phase speed gives rise to essentially nonlinear resonant interactions, an example of such interactions being the EIGW emission by EYW.

**VI. LONG, BALANCED ROSSBY WAVE BREAKING**

In previous sections, we gave evidence that nonlinearity leads to shock formation in the case of unbalanced waves. Hereafter we will show that, for weakly dispersive balanced Rossby waves the scenario of breaking is radically different. In this case, dispersion can counterbalance nonlinearity so that localized patterns can emerge spontaneously. After briefly reminding the known propagating solutions of the model, we show what role these solutions play in the breaking process and how breaking modifies the flow field.

**A. Reminder on equatorial solitary waves**

For long Rossby waves, the equilibration of nonlinearity by dispersion in Eq. (1) yields steadily propagating, localized Rossby wave patterns. These waves were first investigated by Boyd\(^{17}\) who derived by means of asymptotic expansions a Korteweg–de Vries equation for the zonal structure of $u$ for $n=1$ Rossby waves. This classical equation is known to support solitary wave solutions (see, e.g., Ref. 36). Equatorial

---

Figure 12. Snapshot of an EIGW packet with $k_0=1.0/R_d$ at time $t=4.0$.

Figure 13. Snapshot of a WIGW packet at time $t=3.2$ with $k_0=0.293$ such that $c_g(k_0)=0$.

Figure 14. (Color). Hovmöller diagram of $v(x,0,t)$ for an EYW packet with $k_0=1.0/R_d$. Two different runs are superimposed for comparison. Black (green) contours correspond to $\epsilon=0.3$ ($\epsilon=0.01$). The small window shows contours of $h(x,0,t)$ (black contours) and $v(x,0,t)$ (black contours). The excitation of EIGW with $k'_0=2/R_d$ is observed.
solitons of the form \( u \propto \text{sech}^2[B(x-st)] \) were thus found, see Appendix F. Previous numerical studies showed a good agreement with this analytic solution for small amplitude waves.\(^{24,37}\)

Furthermore, Boyd\(^\text{18}\) obtained first-order corrections to solitons. The first-order perturbative solution he found has the property that upon a threshold \( B > 0.532 \), there exists a recirculation region in the fluid. He called such waves equatorial modons. Boyd\(^\text{20}\) investigated numerically the stability of this first order solution. Up to \( B < 0.532 \), he obtained a good correspondence between his analytic solution and the numerics but for \( B > 0.532 \) another branch of propagative solutions was found. These solutions mainly differ in the recirculation region. In order to avoid confusion, in what follows, we call “soliton” a steadily translating Rossby wave which does not trap fluid and “modon” a propagative pattern which traps fluid with no implicit reference to the perturbative definition of Boyd. The following numerical simulations will highlight the underlying dynamical relation between solitons and modons.

**B. Solitary waves breaking on a sloping topography**

The relation between solitons and modons is demonstrated by showing a smooth transition from solitons to modons. We considered the evolution of a soliton in a fluid layer which thickness is decreasing smoothly. As indicated in Sec. II, this can be viewed as the evolution of a soliton in a region submitted to a westerly wind forcing. The slope considered in this experiment is negative. This case is probably not the most geophysically relevant situation since the easterly trade winds in the tropics tend to sustain a positive slope of equatorial ocean thermocline. We should rather consider such a test as a thought experiment in order to illustrate the continuous transition from solitons to modons.

There are several studies on the effect of the variation of the thermocline depth on equatorial waves spectrum.\(^{38,39}\) The effect of the thermocline depth variation was taken into account using a WKB approach so that wave reflection was neglected. But none of these studies dealt with mass transport.

Consider a basin which depth is varying linearly from \( h = H \) for \( x > 0 \) to \( h = 0.4H \) for \( x \leq -20 \). The initial condition is a zeroth-order Boyd soliton centered at \( x = 15 \) with \( B = 0.25 \). The tracer field is initialized so that \( \phi(x,y,0) = a_x + i a_y = x + iy \). The evolution of the soliton is shown in Fig. 15. Tracer contours of \( a_x \) and \( a_y \) are also plotted. The wave is initially propagating toward the West, as expected. At this stage, the tracer field is advected reversibly: no trapping is visible at time \( t = 75.0 \). Then the wave propagates in the sloping topography region. Both its phase speed and meridional extent are diminishing, as expected from a WKB approach.\(^{39}\) During this propagation through the sloping topography the wave is partially reflected forming a small amplitude Kelvin wave. This is not visible in the figure since this emission is below the lowest shown value of \( h \). The amplitude of the reflected Kelvin wave is increasing with steepening topography (not shown). The tracer contours are getting more elongated as the wave propagates. Finally the wave enters the left region.

\[ \text{FIG. 15. Propagation of a solitary Rossby wave on a sloping topography. Lagrangian tracer contours are superimposed.} \]
of the basin and keeps a quasisteady shape and phase speed as shown at time $t=150$. The wave is now enclosing a recirculation region. It has been transformed into a modon. Note that we can trace how mass is trapped in modons. Mass trapping does not occur next to maxima of $h$ around which velocity vectors are circling due to geostrophic balance, but rather near the equator where $h$ is minimum.

Thus, when a solitary Rossby wave propagates on a sloping topography, it breaks forming a modon which encloses a recirculation region. Such a recirculation region is moving with the whole wave pattern so that the underlying particle distribution is irreversibly modified.

C. Rossby wave packet breaking: asymmetry between positive and negative geopotential anomalies

Now that we have shown the dynamical transition of solitons into modons, we can investigate the role of such localized propagating solutions during the breaking of a long Rossby wave packet. As previously, breaking is said to occur when a background flow property is irreversibly modified during wave propagation. Considering the last experiment, we expect the breaking of long Rossby waves to consist in an irreversible mass redistribution due to the formation of modons.

We initialize the fields with a $n=1$ Rossby wave packet according to Appendix F with Gaussian zonal structure $F(x)$ with zonal scale $L=10R_d$ and $\epsilon=0.43$. Note that $\epsilon \sim 0.3$ corresponds to the transition amplitude between equatorial solitons and equatorial modons. The subsequent evolution is shown in Fig. 16. The wave packet is propagating westward. During this propagation, it is split up in a series of isolated structures of decreasing amplitude. Each of them is either an equatorial soliton or an equatorial modon depending on its amplitude. Modons emerge from the initial wave packet and the mass field is irreversibly modified. Noticeably, we can observe an increase in the maximum amplitude of the whole pattern from $\epsilon=0.43$ at time $t=0.0$ to $\epsilon=0.62$ at time $t=180.0$. We should emphasize that such behavior is restricted to positive mass anomaly initial wave packets.

In order to investigate the possible breaking of negative geopotential anomaly Rossby wave packets, we performed the experiment shown in Fig. 17. Initial conditions are strictly identical to those of Fig. 16 except for the sign of the perturbation which is now chosen negative. Such initial condition is not an (asymptotic) propagating solution of system (1), but simply a weakly dispersive, long Rossby wave packet. Due to nonlinear effects, the wave packet is propagating slower than in the linear regime (corresponding contours are overlayed in red, see caption). This results in the emission of slow small scale Rossby waves. Although the spatial coherence of the wave packet has been lost, such behavior does not fit the definition of wave breaking given in the Introduction.

Breaking of long Rossby wave packets crucially depends on the sign of the geopotential perturbation. However, since no shocks appear during Rossby wave breaking, no essentially dissipative wave-mean flow interaction (as in shocks) is taking place. Nonetheless, Rossby wave breaking does induce a wave-mean flow interaction as (at least for positive

FIG. 16. Splitting of a Rossby wave packet in a sequence of equatorial solitary waves.
FIG. 17. Evolution of a negative geopotential anomaly Rossby wave packet. In red, we overlay \( b \) contours obtained with the same initial condition but with \( \varepsilon=0.01 \) in order to compare with the linear regime.

geopotential anomalies) it leads to irreversible formation of mass-trapping regions which are identified by the passive tracer evolution. The potential vorticity, being a Lagrangian invariant, will behave in a similar way. Moreover the concentration of tracer contour visible in Fig. 15 indicates that strong mixing should occur in the region of mass trapping.

VII. SUMMARY AND DISCUSSION

In this paper, we have studied the breaking of equatorial waves within the framework of rotating shallow water model. Irreversible changes of the mean-flow properties during the wave propagation were chosen as indicators of wave breaking.

We have demonstrated crucial dependence of the breaking scenario of a given wave on the degree of balance between the zonal pressure gradient and Coriolis force. This was achieved by means of a high resolution finite volume numerical method which is specially designed in order to capture both the rapid and possibly discontinuous solutions and the slow motions which involve a balance between Coriolis force and the pressure gradient.

On the one hand, breaking of unbalanced waves was shown to consist in shock formation. Energy dissipation in these shocks induces nonadvective potential vorticity fluxes which modify irreversibly the potential vorticity distribution. Such a breaking mechanism is valid for the whole range of unbalanced waves including KW, EYW, EIGW, and WIGW (see Sec. II B) and it is associated with enhanced mixing. A new essentially nonlinear interaction between EIGW and EYW was found.

On the other hand, the breaking of long, balanced Rossby waves is a process which rearranges irreversibly the fluid particles, and hence the PV distribution. This breaking mechanism leads to the formation of propagating patterns, modons. The modons arise spontaneously during the evolution of a zonally elongated, positive geopotential anomaly Rossby wave packet or during the propagation of Rossby solitons in a region of decreasing fluid depth.

We should stress an essential difference between positive and negative geopotential anomalies. As we have shown, the scenario of breaking strongly depends on the sign of the geopotential anomaly. In the case of Rossby waves, this highlights the so-called cyclone-anticyclone asymmetry of the shallow water model (see, e.g., Ref. 40). The asymmetry between positive and negative geopotential anomalies is due to the effect of pseudocompressibility in Eq. (1). Positive anomalies tend to propagate faster than predicted by linear theory whereas negative anomalies are propagating slower. This difference in phase propagation gives rise to a new interaction through nonlinear resonance mechanism. This is illustrated by EIGW/Eyw interaction.

We should emphasize the dynamical variety of the equatorial shallow water model which exhibits a rich panel of nonlinear phenomena ranging from shocks to solitary waves. Dealing with such rich patterns, natural questions of their stability and possible interactions arise. An indication of the robustness of these patterns which are obtained as a result of the breaking processes is given in Appendix D, where we demonstrate that the interaction between a Rossby soliton and a Kelvin front is quasielastic.

Finally, let us note that the description of off-resonant self interactions in the equatorial shallow water model would not be complete without mentioning the case of short scale Yanai and Rossby waves. For those waves, neither shock nor pattern formation was observed in our numerical simulations (not shown). Needless to say that their nonlinear dynamics is far from being trivial and Appendix E provides an illustration of the effect of strong nonlinearity on those balanced waves.

ACKNOWLEDGMENTS

The authors are grateful to Serguei B. Medvedev for his helpful comments on Secs. IV and V. The work of V. Zeitlin was supported by the French National Program “PATOM.”

APPENDIX A: MULTIDIMENSIONAL FINITE VOLUME METHOD

Here, we remind the classical framework of multidimensional finite volume methods for homogeneous systems of conservation laws. A recent comprehensive review of these methods is proposed by LeVeque.31 Consider, for example, the set of Eq. (1) with \( \beta=0 \) written in its conservative form

\[
\partial_t U + \partial_x F(U) + \partial_y G(U) = 0 \quad (A1)
\]

with \( U=(h,hu,hv) \) and its fluxes \( F, G \) are defined by

\[
F(U) = \begin{pmatrix}
hu \\
hu^2 + gh^2/2 \\
hu v
\end{pmatrix}, \quad G(U) = \begin{pmatrix}
hv \\
hu v \\
hv^2 + gh^2/2
\end{pmatrix}. \quad (A2)
\]

Note that \( G(U)^T=F(U) \) with the notation \( X^T=(x_1,x_2,x_3) \) for any vector \( X=(x_1,x_2,x_3) \). This is due to the isotropy of the original conservation law.

Moreover, let us introduce a regular computation grid and the cell-centered variable \( U_{i,j} \). A finite volume scheme

Downloaded 09 Mar 2005 to 194.254.66.226. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp
computes the evolution of $U_{ij}$ using the values of $U$ in the neighboring cells. More precisely, these contributions are split in four interface fluxes as presented in Fig. 18 (upper panel). This yields the following first-order update formula:

$$U_{ij}^{n+1} = U_{ij}^n - \frac{\Delta t}{\Delta x} (F_{i+1/2,j} - F_{i-1/2,j}) - \frac{\Delta t}{\Delta y} (G_{i,j+1/2} - G_{i,j-1/2}).$$

(A3)

Under a Courant-Friedricks-Lewy condition on the time step, the flux between two adjacent cells (e.g., $F_{i-1/2,j}$) is obtained from the resolution of the Riemann problem between the right and left states (e.g., $U_{i-1,j}$ and $U_{i,j}$) see Fig. 18 (lower panel).

The routine which is used to compute interface fluxes, e.g., $F_{i+1/2,j} = F(U_{i,j}, U_{i+1,j})$, is called the Riemann solver. Given the aforementioned symmetry between $F$ and $G$, it is possible to use the same Riemann solver in both directions $x$ and $y$. Therefore, we only need to know how to deal with the Riemann problem between the states $U_l$ and $U_r$ for the following set of equations:

$$h_{l} + (hu)_{x} = 0,$$

$$(hu)_{t} + (hu^2 + gh^2/2)_{x} = 0,$$

$$(hv)_{t} + (huv)_{x} = 0.$$  (A4)

This procedure is known as dimensional splitting.\textsuperscript{31} Note that the third equation in Eq. (A4) is an advection equation which is now decoupled so that we have arrived at a simple one-dimensional problem.

**APPENDIX B: REMINDER ON PARABOLIC CYLINDER FUNCTIONS**

For a given $n$, the parabolic cylinder function $D_n$ is defined as the unique solution satisfying the equation

$$D''_n + (2n + 1 - y^2)D_n = 0,$$  (B1)

that is decaying at $\pm \infty$. This leads to the following formula:

$$D_n(y) = \frac{H_n(y)e^{-y^2/2}}{\sqrt{2^n n! \pi}}.$$  (B2)

$H_n$ are Hermite polynomials which are defined by

$$H_n = 2nH_{n-1} - H_{n-1},$$  (B3)

where $n \geq 1$ and $H_0 = 1$.

**APPENDIX C: LOCALIZED WAVE PACKET SYNTHESIS**

Having chosen a wave type (WIGW, EIGW, RW, EYW, or WYW), a meridional quantum number $n \geq 0$, a peak wave-number $k_0$, and a nondimensional amplitude $\epsilon$, we define the initial conditions by

$$u^0(x,y) = \frac{1}{2} \left[ q_n(x)D_{n+1}(\tilde{y}) - r_n(x)D_{n-1}(\tilde{y}) \right],$$

$$v^0(x,y) = v_nD_n(\tilde{y}),$$

$$h^0(x,y) = \frac{c}{2g} \left[ q_n(x)D_{n+1}(\tilde{y}) + r_n(x)D_{n-1}(\tilde{y}) \right].$$  (C1)

with $\tilde{y} = y/R_d$. $D_n$ refers to the $n$th parabolic cylinder function which definition is reminded in Appendix B. The structure function ($q_n, v_n, r_n$) is obtained by Fourier synthesis as

$$q_n(x) = \frac{i}{\pi} \frac{1}{\sqrt{2R_d}} \int_0^{\infty} F(k)A^{(+)}(k)e^{ikx}dk,$$

$$v_n(x) = \frac{1}{\pi R_d} \int_0^{\infty} F(k)A^{(0)}(k)e^{ikx}dk,$$

$$r_n(x) = -\frac{i}{\pi} \frac{1}{\sqrt{2R_d}} \int_0^{\infty} F(k)A^{(-)}(k)e^{ikx}dk,$$  (C2)

where $A^{(+)}$, $A^{(0)}$, $A^{(-)}$ are given by Ripa\textsuperscript{12} in his formula (2.27) and depend on the wave type of interest. We choose the spectrum $F(k)$ to be of the form

$$F(k) = \alpha e^{-\sqrt{(k - k_0)^2 + \delta^2}},$$  (C3)

where $\alpha$ is determined such that the resulting $h(x,y,0)$ has a nondimensional amplitude $\epsilon$.

**APPENDIX D: QUASIELASTIC COLLISION BETWEEN A ROSSBY SOLITARY WAVE AND A KELVIN WAVE**

In order to highlight a surprising robustness of the patterns that emerge from breaking equatorial waves, let us consider the interaction between a Rossby solitary wave and a Kelvin wave. The initial condition consists in the sum of two distinct patterns. A Kelvin wave with Gaussian zonal structure is centered at $x = 10.0$. Its zonal extent is $L_z = 3R_d$ and its...

amplitude $\epsilon = -0.3$. A solitary Rossby wave is initialized at $x = 45.0$. We choose Boyd’s zero order solution with amplitude $\epsilon \sim 0.18$. Snapshots of the evolution of this initial condition are shown in Fig. 19. At time $t = 15.0$, a shock front has formed at the rear of the Kelvin wave packet. This mechanism is described in Sec. IV B. Dissipation in this shock induces a slow decrease of the Kelvin wave amplitude as it is propagating eastward. On the other hand, the solitary Rossby is moving westward with no change of shape. At time $t = 30.0$, the two wave patterns collide but their interaction is quasielastic, as seen at time $t = 45.0$. The quasielastic character of the interaction can be verified by additional checks (not shown). This almost-noninteraction between Kelvin and Rossby waves was observed for a wide range of parameters. The only effect we have registered was a slight delay in the solitary Rossby wave propagation for high amplitude $\epsilon > 0.5$.

APPENDIX E: SMALL SCALE BALANCED WYW PACKETS

In this Appendix we briefly report the effect of strong nonlinearity on almost balanced small-scale WYW packets. Figure 20 presents dispersion of a WYW packet centered about $k_0 = 2/R_d$ with amplitude $\epsilon = 0.25$. Since Yanai waves are antisymmetric with respect to the equator, we plotted $v$ contours along the equator. By comparison with the linear regime (contours overlayed), we see that the wave experiences a slight frequency shift in the nonlinear regime. Such a frequency shift was predicted by Le Sommer et al. on the basis of scale analysis.

APPENDIX F: ROSSBY SOLITONS

Boyd looked for propagating Rossby wave packets of the form

$$u = \left( -9 + 6y^2 \right) / 4e^{-y^2/2} F(x)$$

FIG. 19. Interaction between Kelvin wave and Rossby solitary wave. See text for details.

FIG. 20. Hovmöller diagram of $v(x,0,t)$ for a WYW packet with $k_0 = 2.0/R_d$. As previously, a run was performed with $\epsilon = 0.01$ and plotted in green.
\[
v = 2\gamma e^{-y^2/2}F'(x),
\]

\[
h = 1 + (3 + 6\gamma^2/4)e^{-y^2/2}F(x),
\]

where \(\gamma = y/R_2\) and \(F(x)\) prescribes the zonal structure of the wave-packet. He obtained Rossby solitons of the form \(F(x) \approx \text{sech}^2(Bx)\).

In our experiments, we used the same zonal structure as an initial condition in Sec. VI B. In Sec. VI C, we impose \(F(x) \approx \exp(-x^2/2L_x^2)\) in order to build a Gaussian Rossby wave packet.

40 L. Polvani et al., Chaos **4**, 177 (1994).