

## MONOPOLAR AND DIPOLAR VORTEX SOLITONS IN TWO SPACE DIMENSIONS

John P. BOYD

*Department of Atmospheric, Oceanic and Space Sciences and Laboratory for Scientific Computation, University of Michigan, 2455 Hayward Avenue, Ann Arbor, MI 48109, USA*

Received 23 June 1989, Revised 1 August 1990

We compute solitary waves which solve  $\Delta u - y^2 u - (u + yu^2)/c = 0$  where  $c$  is the phase speed in the  $x$  direction and  $y = 0$  is the equator. This equation is a heuristic model for Rossby waves on the "equatorial beta-plane" in geophysical fluid dynamics. For positive  $c$  only, there are one-signed solutions ("monopole vortices") which are centered in the middle latitudes. When  $c \in [-\infty, -1/3]$  or  $c > 0$ , there are dipoles which have matching vortices of opposite sign in each hemisphere. A third class of solutions is composed of equator-spanning monopoles that are unsymmetric in  $y$ . In addition to these families of strict solitary waves, there are also quadrapole vortices which are "weakly non-local solitons" in the sense that they almost meet the usual criterion for solitary waves except for a weak—very weak—radiation to infinity. Orthogonal rational Chebyshev functions and Newton's iteration are used to compute numerical solutions, but four analytic approximations are also derived. Although the equation is only a crude model of geophysical waves, its rich diversity offers a good education in solitary waves in two space dimensions.

### 1. Introduction

In the Appendix, we derive

$$u_{xx} + u_{yy} - y^2 u - u/c - yu^2/c = 0$$

"AEW equation" (1.1)

where  $c$  is the phase speed. (The waves propagate parallel to the  $x$ -axis.) Since (1.1) is an equatorial wave equation that incorporates some ageostrophic effects, we have dubbed (1.1) the "Ageostrophic Equatorial Wave" equation or "AEW" for short.<sup>1</sup> (I thank the reviewer for suggesting this label.) In the rest of this article, we will describe effective numerical and analytical methods for computing its soliton solutions, i.e., those waves that satisfy the boundary condition

$$u(x, y) \rightarrow 0 \quad \text{as } |x^2 + y^2| \rightarrow \infty \quad (1.2)$$

<sup>1</sup> Note that in the review [1], this same equation is labelled the "EMR" equation because it is an equatorial version of models discussed by Malanotte-Rizzoli [2].

The AEW equation has some limitations as a model which are explained in the Appendix. The nonlinear shallow water wave equations would be a better model. However, these are a set of three equations in three unknowns (rather than a single equation like (1.1)) and have very complicated physics. Consequently, only Boyd [3, 4] and Marshall and Boyd [5] have computed nonlinear waves for the shallow water wave equations, and then only using perturbation theory. The reviews by Malanotte-Rizzoli [2], Flierl [6], Yano and Tsujimura [7], and McWilliams [8] and the many articles in Nihoul and Jamart [9] catalogue more than one hundred papers that have computed nonlinear planetary waves using various approximations similar to (1.1).

The AEW eq. (1.1) is no less limited than these other approximations, but it does introduce new ingredients: explicit variations in the Coriolis parameter and equator-spanning solitary waves. Thus,

this article is a logical extension to previous work on geophysical solitons.

In Section 2, we discuss the one-signed or “monopole” solutions in what a geophysicist would call an “ $f$ -plane” approximation: replacing the variable coefficients in (1.1) by constants. In this limit, both the phase speed  $c$  and the latitude of the center of the monopole,  $y_0$ , may be chosen arbitrarily and independently. In Section 3, we apply perturbation theory to compute the first order effect of variable  $y$ , and find that a bounded solution is possible only when there is an eigenrelation between  $y_0$  and  $c$ .

In Section 4, a different perturbation method is used to derive analytical approximations to dipolar vortices in the Korteweg-deVries-like limit  $c \rightarrow -1/3$ . Multipole vortices are also predicted by the perturbation theory, but as explained in Section 5, the KdV-type theory misses an important complication. This is revealed by a “far field” analysis, that is, examining the equation as  $|x| \rightarrow \infty$  so that the wave amplitude decays to very small values and the dynamics becomes linear. One finds that the multipoles leak energy to  $x = \pm\infty$  through radiation in the  $n = 1$  linear wave mode. However, the amplitude of this radiation is *exponentially* small in the amplitude of the multipole vortex, so it is still meaningful to speak of solitons in an approximate sense.

In the limit  $|c| \rightarrow \infty$  (Section 6), the dipole structure becomes independent of  $c$ . The limits  $c \rightarrow -\infty$  and  $c \rightarrow \infty$  are identical except for a change in sign of  $u(x, y)$ . Thus, both branches of the dipole ( $c \in [-\infty, -1/3]$  and  $c \in [0, \infty]$ ) are connected to form a single, continuous family of solutions.

The final analytic approximation for the dipole is derived in Section 7. In the limit  $c \rightarrow 0+$ , the dipole becomes two widely separated, weakly interacting monopoles of the sort described in Sections 3 and 4, one monopole in each hemisphere. The dipole is always antisymmetric with respect to  $y = 0$ .

Section 8 presents numerical results for dipoles which fill the gaps between the three analytical approximations or limits discussed earlier. Some

readers may prefer to read this dipole summary section first and then backtrack to the special cases described in Sections 4, 6 and 7. Section 8 has been written to be self-contained to facilitate this option.

Section 9 describes our numerical methods. The use of rational Chebyshev functions [10] makes it possible to numerically compute the solitary waves on a doubly infinite domain without the use of artificial boundaries. We also use the numerical algorithm to generate an *analytic* approximation to the monopole which is the ratio of two cubic polynomials; the accuracy is less than the thickness of the curve.

Section 10 discusses solitary waves that are *almost* symmetric about the equator,  $y = 0$ . (In Section 4, we prove that solutions that rigorously satisfy the symmetry condition  $u(x, y) = u(x, -y)$  for all  $x$  and  $y$  are impossible.) Perturbation theory shows that these unsymmetric monopoles have the  $x$ -structure of solitons of the modified Korteweg-deVries equation.

The final section is a summary and discussion of open problems.

A terminological note: we shall use “soliton” and “solitary wave” as synonyms. Some authors reserve “soliton” for solitary waves that collide elastically, such as the solitons of the Korteweg-deVries equation. However, it is now known that this property is very special, limited to the mostly one-dimensional equations that can be solved by the inverse scattering method. In this article, because we compute only isolated solitons, distinctions based on properties in solitary wave – solitary wave collisions are irrelevant in any event.

## 2. Radially symmetric monopoles on the “ $f$ -plane”

When the soliton is centered at some latitude  $y_0$  which is *large* in comparison to the radial  $e$ -folding scale of the wave, it follows that the  $y$ -dependent coefficients in the AEW equation will vary by only a few percent over the region where the vortex is large. It is therefore a legitimate approximation to replace  $y$  by its mean value,  $y_0$ . Geophysical fluid

dynamicists call this the “*f*-plane approximation” because the factors of *y* in (1.1) are really the non-dimensional Coriolis parameter *f*. The result is

$$u_{xx} + u_{yy} - y_0^2 u - u/c - y_0 u^2/c = 0$$

“*f*-plane AEW” (2.1)

**Theorem.** The *f*-plane AEW equation, (2.1), has a family of solutions which are one-signed, i.e., are vortex monopoles. These are radially symmetric about the reference latitude *y*<sub>0</sub> which denotes the center of the vortex. They may be written in terms of a single universal function as

$$u(x, y) = A W(Br) \quad \text{“}f\text{-plane monopole”}$$

(2.2)

where

$$r \equiv (x^2 + [y - y_0]^2)^{1/2} \tag{2.3}$$

$$A \equiv (1 + cy_0^2)/y_0 \tag{2.4}$$

$$B \equiv (1/c + y_0^2)^{1/2} \tag{2.5}$$

and where *W*(*r*) is the solution of

$$\Delta W - W - W^2 = 0$$

“Flierl–Petviashvili monopole” (2.6)

where  $\Delta$  is the usual Laplacian operator ( $= \partial^2/\partial r^2 + (1/r)\partial/\partial r$ ).

For independent, arbitrary values of the parameters *c* and *y*<sub>0</sub>, (2.2) through (2.5) solve the *f*-plane AEW equation, (2.1). However, in the next section, we show that the perturbation equations for the first order effects of the *y*-dependent coefficients of the exact AEW equation on the monopole are solvable only when *y*<sub>0</sub> = ±1/*c*<sup>1/2</sup>, which simplifies (2.4) and (2.5) to

$$A = \pm 2c^{1/2}, \quad B = (2/c)^{1/2} \tag{2.7}$$

where the sign of *A* is the sign of *y*<sub>0</sub>. Note that *W*(*r*) < 0 for all *r* so that the northern hemisphere monopole is negative while its mirror image for *y* < 0 is everywhere positive. Observe that (2.7) implies that all monopoles have *positive* phase speed, *c* > 0.

**Proof.** When (2.2) through (2.5) is substituted into (2.1), one may trivially cancel all the factors of *c* and *y*<sub>0</sub> to reduce (2.1) to (2.6).

The parameter-free function *W*(*r*) that solves (2.6) may be dubbed the “Flierl–Petviashvili monopole” because it was independently discovered by G. Flierl [11] (ocean solitons) and Petviashvili [12] (the Great Red Spot of Jupiter and plasma drift waves). By applying the numerical methods discussed in a later section, truncating the series, and combining terms we obtain the approximation

$$W(r) \approx (-2.3822 - 1.01327r^2 + 0.02417r^4) / (1 + [3/4]r^2 + [3/16]r^4 + [1/64]r^6)$$

The maximum absolute error is about 0.0105, which is 0.44% of the minimum of this negative-definite function. For a still more accurate approximation, which is needed to compute the corrections due to variable *y* in the next section, one may use the coefficients collected in Table 1, which are those of the sum

$$W(r) \approx \sum_{n=1}^{10} a_{2n} [\cos(2n \operatorname{arccot} [r/2]) - 1] \tag{2.9}$$

which is accurate to about 1 part in 10<sup>4</sup>.

### 3. Monopoles in the variable coefficient AEW: Perturbation theory and the eigenrelation

It is easy to substitute the *f*-plane monopole into the original, unapproximated AEW to correct for variable *y*. Defining

$$u = u^0 + u^1 + \dots \tag{3.1}$$

gives the first order perturbation equation

$$\Delta u^1 - (1/c + y_0^2)u^1 - (2y_0/c)u^0 u^1 = (y - y_0)\{2y_0 u^0 + [u^0]^2/c\} \tag{3.2}$$

By rescaling the spatial coordinates and *u*<sup>0</sup>, we can

Table 1  
Coefficients of the spectral series for the Flierl-Petviashvili monopole,  $W(r)$

$n$	1	2	3	4	5	6	7	8	9	10
$a_{2n}$	1.25538	-0.21751	-0.06430	-0.00551	0.00339	0.00271	0.00139	0.00059	0.00020	0.00003

reduce (3.2) to

$$V_{rr} + (3/r)V_r - V - 2W(r)V = W + \lambda W^2 \tag{3.3}$$

where

$$\lambda \equiv (1/c + y_0^2)/(2y_0^2) \tag{3.4}$$

$$u^1 = 2c(y - y_0)V([1/c + y_0^2]^{1/2} \times [x^2 + (y - y_0)^2]^{1/2}) \tag{3.5}$$

The perturbation eq. (3.3) looks harmless, but for arbitrary  $\lambda$ , it has no bounded solution. The problem is that because its coefficients are constant, the  $f$ -plane AEW equation is *translationally invariant*. It follows that if  $u^0(x, y)$  is a solution to (2.1), then  $u^0(x, y + \delta)$  must also be a solution for arbitrary constant  $\delta$ . For the special case of  $\delta \ll 1$ ,

$$u^0(x, y + \delta) \approx u^0(x, y) + \delta \partial u^0 / \partial y + O(\delta^2) \tag{3.6}$$

The only way (3.6) may solve the  $f$ -plane AEW equation for small  $\delta$  is if  $u_y^0$  is an *eigenfunction* of the linearized AEW equation with zero eigenvalue, that is to say, if  $u_y^0$  is a non-trivial homogeneous solution of (3.2).

It is well-known that a self-adjoint inhomogeneous differential equation with a non-trivial homogeneous solution is solvable if and only if the inhomogeneous term is orthogonal to the homogeneous solution. Unfortunately, (3.3) is not self-adjoint, but it is easy to apply the orthogonality condition associated with the original partial differential equation. The eigenfunction of (3.2) is proportional to

$$e_0 \propto (y - y_0) W_r(Br)/r \tag{3.7}$$

The condition that the two-dimensional integral of (3.7) with the R.H.S. of (3.2) is zero may be

simplified by extracting the non-negative factor  $\sin^2(\theta)$  to give

$$\int_0^\infty r dr \{ [W_r][r(W + \lambda W^2)] \} = 0 \tag{3.8}$$

Multiplying (2.6) by  $r^2 W_r$ , integrating over  $r$ , substituting the result in (3.8), and then integrating by parts prove that  $\lambda = 1$ , which is equivalent to

$$y_0 = 1/c^{1/2} \tag{3.9}$$

“Monopole eigencondition”

Figure 1 shows both the Flierl-Petviashvili monopole and the radially-dependent part of its first order correction. Table 2 gives the spectral coefficients of  $V(r)$ .

Swenson [13] derives a first order eigencondition similar to (3.8) when refining Flierl’s theory [11] of quasi-geostrophic solitary waves in a weak shear flow. In that model, however, the integral contains only a single term and there is no parameter to

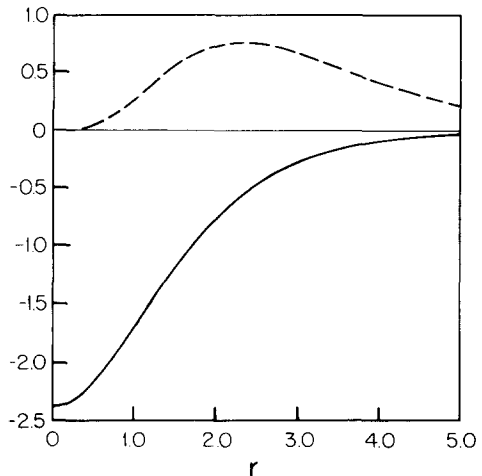


Fig. 1.  $W(r)$ , the universal shape function for the  $f$ -plane monopole, is the solid curve. The dashes show  $rV(r)$ , the radially dependent part of the first order correction to  $W(r)$  due to the variation of the AEW coefficients with  $y$ .

Table 2  
Coefficients of  $V(r)$ , the first order correction to the monopole

$n$	1	2	3	4	5	6	7	8	9	10
$a_{2n}$	-0.0545	-0.1810	0.0516	0.0262	0.0048	-0.0009	-0.0014	-0.0010	-0.0005	-0.0003

vary so that the integral is zero. In our application, the first order theory *reduces* the set of  $f$ -plane monopoles from what is predicted by the zeroth order analysis of the previous section, but in the Flierl-Swenson case, the solitons are entirely *eliminated*.

#### 4. Dipoles in the small amplitude, KdV-like limit

As  $c \rightarrow -1/3$ , the AEW dipoles have small amplitude and a very large zonal ( $x$ ) scale. One may systematically apply the singular perturbation technique known as the “method of multiple scales” as in Boyd [3, 4] to derive an analytic approximation to this dipoles. Since perturbation theory is well understood, we will take a different approach here.

The Hermite functions are a complete orthogonal basis set on  $y \in [-\infty, \infty]$ . It follows that without approximation, we may write

$$u = \sum_{n=0}^{\infty} A_n(x)\psi_n(y) \tag{4.1}$$

By using the eigenequation satisfied by the Hermite functions, which is

$$\psi_{n,yy} - y^2\psi_n = -(2n + 1)\psi_n \tag{4.2}$$

one may apply Galerkin’s method to (1.1) using the series (4.1) to obtain, also *without approximation*, the infinite set of equations ( $n = 0, 1, \dots$ )

$$\begin{aligned} &A_{n,xx} - [(2n + 1) + 1/c]A_n \\ &- (1/c) \sum_{k,m=0}^{\infty} A_k A_m \\ &\times \langle \psi_n, y\psi_k\psi_m \rangle / \langle \psi_n, \psi_n \rangle = 0 \end{aligned} \tag{4.3}$$

where  $\langle a, b \rangle$  denotes the inner product of  $a(y)$  with  $b(y)$ , i.e., the unweighted integral from  $-\infty$  to  $\infty$ .

If we expand the  $A_n(x)$  as Chebyshev series, we would obtain a Chebyshev/Hermite spectral method that offers the same exponential convergence as the double Chebyshev pseudospectral method described in Section 9. Unfortunately, (4.3) is harder to program, which is why the rational Chebyshev method is used in Sections 8 and 9. However, (4.3) is very useful for theoretical purposes as shown by the following.

**Theorem.** The AEW equation has no solutions which are symmetric about  $y = 0$ .

**Proof.** If the sum (4.1) is restricted to even  $n$ , that is, to basis functions such that  $\psi_n(y) = \psi_n(-y)$ , then all the inner products in the double sum are equal to zero because the factor of  $y$  makes the integrand antisymmetric about the origin. The resulting uncoupled, linear ordinary differential equations have no bounded solutions as may be verified by explicitly solving them in terms of hyperbolic functions.

It is always necessary to truncate the set (4.3) to a large but finite number  $N$  of equations. In general, this truncated set must be solved numerically. However, if we make the most drastic possible truncation, i.e., retaining just a single Hermite function in the basis, then we obtain

$$\begin{aligned} &A_{n,xx} - [(2n + 1) + 1/c]A_n - I_n A_n^2 / c = 0 \\ &\text{“1-mode truncation”} \end{aligned} \tag{4.4}$$

where  $I_n$  is the self-interaction coefficient

$$I_n = \langle \psi_n, y\psi_n\psi_n \rangle / \langle \psi_n, \psi_n \rangle \tag{4.5}$$

This single ODE in  $x$  may be solved *analytically*. However, this observation is useless unless there is a limit in which this drastic truncation gives a self-consistent approximation.

One such limit arises if the self-interaction of a given mode  $n$  is resonant while its interaction with the other modes is nonresonant. To understand this distinction between “resonant” and “nonresonant”, pretend that the nonlinear terms are known forcing functions, and consider the response of the linear terms, that is,

$$A_{n,xx} + [(2n+1) + 1/c]A_n = f_n(x) \quad (4.6)$$

It is easy to formally solve (4.6) via Green’s function to show that usually  $A_n \sim O(f_n(x))$ , a nonresonant response. However, if

$$c \approx -1/(2n+1) - \delta \quad \delta \ll 1$$

$$\text{“resonance condition”} \quad (4.7)$$

then  $A_n \sim O(f_n/\delta)$  so that this one coefficient is huge in comparison to the nonlinear interaction which is driving it.

Since (4.7) may be satisfied for only a single mode at a time—if  $c \approx -1/3$ , then (4.7) applies only for  $n=1$ —it follows that at most a *single* mode may be resonant for a given value of  $c$ . This in turn implies that when  $c$  is close to one of the resonant values given by (4.7), it is consistent to truncate the Galerkin equations to just a single mode. The general solution of (4.4) is

$$A_n(x) = a \operatorname{sech}^2(bx) \quad n \text{ odd} \quad (4.8)$$

where

$$\begin{aligned} a &= -(6/I_n)cb^2, \\ b^2 &= [(2n+1) + 1/c]/4 \end{aligned} \quad (4.9)$$

but the one-mode truncation that leads to (4.8) and (4.9) is consistent only when the resonance condition is approximately satisfied. This in turn implies that the solitons must have *small amplitude* since  $a \sim O(\delta)$  when  $\delta \ll 1$ . The complete lowest order approximation is

$$u_n(x, y) = a \operatorname{sech}^2(bx) H_n(y) e^{-0.5y^2}, \quad \text{odd } n \quad (4.10)$$

where we have written  $\psi_n(y)$  as the product of the  $n$ -th Hermite polynomial with the Gaussian. The restriction to odd  $n$  is necessary because the self-

interaction coefficient is identically 0 for all the symmetric modes. The  $n$ -th Hermite polynomial has  $n$  zeros, one of which is at  $y=0$ , so the solitary wave is a multipole vortex in the sense that there are  $(n+1)$  regions of alternating sign along a line of constant  $x$ .

Because (4.4) is merely the  $x$ -integrated form of the Korteweg-deVries (KdV) equation (specialized to a steadily translating wave), we shall refer to (4.10) and (4.7) as the “small amplitude, KdV-like” regime.

The gravest mode,  $n=1$ , gives a dipole vortex:  $u(x, y)$  is positive in the northern hemisphere ( $y > 0$ ) and negative for negative  $y$ . For this special case, (4.10) becomes

$$\begin{aligned} u(x, y) &= 8.27\delta \operatorname{sech}^2([3/2]\delta^{1/2}x)y \\ &\quad \times \exp(-0.5y^2) \end{aligned} \quad (4.11)$$

where  $\delta$  is defined by

$$c = -1/3 - \delta, \quad \delta \geq 0 \quad (4.12)$$

For small amplitude and  $c \approx -1/3$ , the  $n=1$  soliton has elliptical contours of constant  $u$  which, as shown in Fig. 2, are highly elongated in the  $x$ -direction because the zonal scale of the wave is  $O(\delta^{-1/2})$ . The  $y$ -scale, in contrast, is always  $O(1)$  independent of  $\delta$ . (This mismatch in  $x$  and  $y$  scales is the basis for the perturbation technique used in [3] and Section 10, the method of “multiple scales”.) As the phase speed becomes more and more negative, the amplitude increases and the zonal scale decreases so that the contours of constant  $u$  become more and more circular.

Although the one-mode truncation can only be strictly justified for  $c \approx -1/3$ , the solution (4.8) to (4.9) predicts solitons which are positive for  $y > 0$  for all  $c \in [-\infty, -1/3]$  and solitary waves of the opposite sign for all  $c > 0$ . The numerical study described below shows that this is precisely what happens. The one mode solution also predicts a finite minimum zonal scale as  $c \rightarrow -\infty$ , and this, too, is correct. It is intriguing that such a simple-minded approximation may be *qualitatively* correct well beyond its formal region of validity.

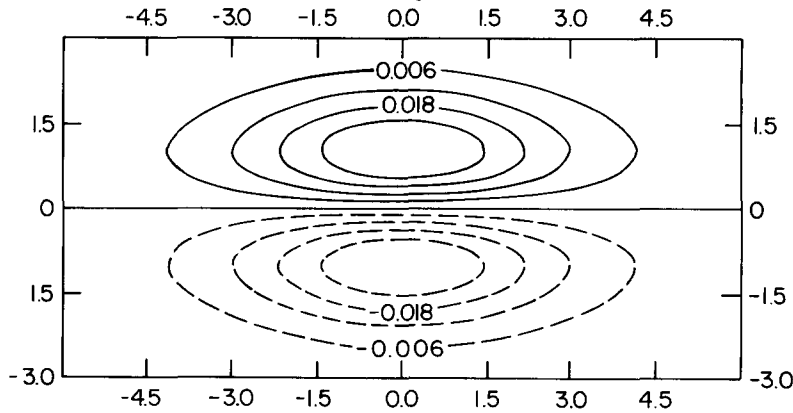


Fig. 2. Contours of constant  $u$  for  $b=1/3$  for the  $n=1$  (dipole) solitary wave for  $c \approx -1/3$ . Positive contours are solid; negative contours are dashed. Because of the four-fold symmetry evident in the graph, later figures show only a single quadrant of the  $x$ - $y$  plane for a given solution.

**5. Multipoles for small amplitude: weakly non-local solitary waves**

The “far field” is defined to be the region  $|x| \gg 1$  where the soliton centered at  $x = 0$  has decayed to such small amplitude that the *nonlinear* terms in the AEW equation may be *neglected*. In the “far field”, the Hermite–Galerkin equations (4.3) simplify to a set of uncoupled, linear ODE’s.

When  $c$  satisfies the constraint

$$c < -1/3 \text{ or } c > 0 \tag{5.1}$$

all the equations in (4.3) have hyperbolic solutions, and it is possible for the soliton to decay exponentially as  $|x| \rightarrow \infty$ . When the  $n=3$  mode is resonant, however,  $c \approx -1/7$ . Substituting this into the  $n=1$  equation gives

$$A_{1,xx} + 4A_1 \approx 0, \tag{5.2}$$

“far field”  $n=1$ ;  $c \approx -1/7$

i.e., the coefficient of the  $n=1$  Hermite function is *oscillatory* rather than exponentially decaying as  $|x| \rightarrow \infty$ .

The result is that the multipole vortices cannot be solitons in the strict sense of solutions which

decay exponentially as  $|x| \rightarrow \infty$ . Instead, multipoles are *generalized* solitons, members of a class known as “weakly non-local solitary waves”. The review by Boyd [1] discusses eight such examples.

A simple analysis explains the precise meaning of “weakly” non-local. When the amplitude  $a$  and pseudo-wavenumber  $b$  are sufficiently small, the resonant component  $A_n(x) \gg A_m(x)$  for all  $m \neq n$ . It follows that for the  $n=3$  soliton, it is good approximation for all  $x$  to write

$$A_{1,xx} + 4A_1 = A_3^2 \langle \psi_1, y\psi_3\psi_3 \rangle / \langle \psi_1, \psi_1 \rangle \tag{5.3}$$

Equation (5.3) is a linear, forced ODE.

The crucial point is that the scale of the forcing, which is  $O(1/b)$ , is very large in comparison to the  $O(1)$  length scale of the homogeneous solutions of (5.3), which vary on a length scale of  $O(1)$ . The result, as shown by Boyd [14], is that although  $A_1(x)$  is the same order of magnitude as the forcing, i.e.,  $O(a^2)$ , for small  $x$ , the amplitude of the oscillations in the “far field”  $|x| \gg 1$  is *exponentially small* in the perturbation parameter  $b$ . This conclusion is true for any R.H.S. in (5.3) provided that the forcing varies on a length scale of  $O(1/b)$ .

For a forcing which is proportional to  $b^4 \text{sech}^4(bx)$ , as is (5.3), one may show that in the

“far field” [1]

$$A_1(x) \sim [\text{constant}] \exp(-\pi/b) \times \cos(2[x + \text{phase}]), \quad |x| \gg 1 \tag{5.4}$$

This implies that when  $b = 1/10$ , the amplitude of the “far field” oscillations is only  $O(10^{-14})$  relative to the maximum of the solitary wave!

Boyd [1, 15, 16] describes numerical algorithms for computing the non-local soliton, complete with its oscillatory “wings”. These articles also explain the radiation condition and other technical details that justify (5.4). Williams and Wilson [17] have investigated equatorial multipole solitons using an time-dependent shallow water model initialized with the perturbative solution. They find that the  $n = 1$  dipole is a classical soliton. In contrast, the higher multipoles are long-lived for small amplitude, but decay through radiative leakage for large amplitudes—just as predicted here.

We omit a discussion of numerical solutions for “weakly non-local” AEW quadrupoles because such is given in the note by Boyd [18].

**6. Dipoles in the limit  $|c| \rightarrow \infty$ : the connection through infinity**

In the limit  $|c| \rightarrow \infty$ , the term  $(-u/c)$  in (1.1) becomes negligible in comparison to  $(-y^2u)$  everywhere except in the neighbourhood of the equator. Since the length scale remains  $O(1)$  everywhere, it follows that  $(-u/c)$  is negligible in comparison to the Laplacian at all latitudes.

The magnitude of the nonlinear term is more subtle. If the amplitude remains  $O(1)$  as  $|C| \rightarrow \infty$ , then  $(-yu^2/c)$  becomes negligible. However, the remaining linear equation has no bounded solution. It follows that the magnitude of  $u(x, y)$  must increase with  $|c|$  so that the nonlinear term is comparable with the surviving linear terms in this limit. The conclusion is that the AEW equation reduces

to

$$u_{xx} + u_{yy} - y^2u - yu^2/c = 0 \tag{6.1}$$

“large  $|c|$  AEW equation”

It is trivial to show that the *general* solution of (6.1) is given by the single, universal function  $v(x, y)$  where

$$u(x, y) = cv(x, y) \tag{6.2}$$

and  $v(x, y)$  is the solution of the parameter-free equation

$$v_{xx} + v_{yy} - y^2v - yv^2 = 0 \tag{6.3}$$

The “far field” analysis shows that  $n = 1$  solitons exist only when  $c$  is on either of two intervals:  $c \in [-\infty, -1/3]$  and  $c \in [0, \infty]$ . The gap between these intervals would seem to imply two unrelated branches of solutions, but (6.2) gives a different perspective. It shows that the solution for  $c \rightarrow -\infty$  is merely the *negative* of the solution for  $c \rightarrow \infty$ . Thus, the two branches are a single, continuous branch *connected through infinity*.

By “infinity” we mean infinite phase speed and infinite amplitude. The function  $v(x, y)$  which is the connection through infinity has  $O(1)$  length scales, however. As  $c \rightarrow \pm\infty$ , the amplitude of the dipole increases without bound, but the structure tends to a limiting shape.

**7. Dipoles in the limit of small, positive  $c$ : widely separated, weakly interacting monopoles**

In Section 3, we showed that the AEW equation has monopole solutions whose  $e$ -folding scale is  $O(1/c^{1/2})$  with centers at  $y_0 = 1/c^{1/2}$ . It follows that the amplitude of these monopoles at the equator is  $O(\exp[-1.414/c])$ . It follows that if we place two monopoles of equal magnitude in opposite hemispheres, we obtain a dipole with an error which is exponentially small in  $c$ :

$$u(x, y) = 2c^{1/2} \{ W([2/c]^{1/2}) \times [x^2 + (y - 1/c^{1/2})^2]^{1/2} \}$$



$$\begin{aligned}
 & - W([2/c]^{1/2} \\
 & \times [x^2 + (y + 1/c^{1/2})^2]^{1/2}), \\
 & 0 < c \ll 1 \quad (7.1)
 \end{aligned}$$

where  $W(r)$  is the Flierl–Petviashvili monopole defined by (2.6) and illustrated in Fig. 1. Note that since the monopole is of opposite sign in opposite hemispheres, the superposition of two monopoles is antisymmetric about  $y=0$ —negative for  $y>0$  and positive south of the equator.

A more systematic analysis begins by multiplying the AEW equation by  $c$  to give

$$c[u_{xx} + u_{yy}] - cy^2u - u - yu^2 = 0 \quad (7.2)$$

The naive approach of neglecting all the terms multiplied by  $c$  when  $c \rightarrow 0+$  is clearly incorrect because this would leave an algebraic rather than a differential equation. Instead, one is forced to conclude (as verified by (7.1)) that the zonal and latitudinal scales are  $O(c^{1/2})$  in this limit.

The subtlety that lurks to snare the unwary is that  $O(c^{1/2})$  is the correct length scale only for the derivatives. The undifferentiated terms with coefficients proportional to  $y$  or  $y^2$  are controlled by the second, independent latitudinal scale  $y_0$ : the distance of the center of each of the two vortices from the equator.

The perturbative analysis of Section 3 shows that  $y_0 \approx 1/c^{1/2}$  so that this scale *increases* as  $c \rightarrow 0$  even though the derivative length scale is doing just the opposite. The resulting dipole approximation (7.1) is self-consistent. The question remains: Is  $y_0 \sim O(1/c^{1/2})$  the *only* possibility?

Inspecting (7.2), we see that we cannot neglect the nonlinear term because the resulting linear equation has no bounded solution. We attempted to solve (1.1) with the  $(-y^2u)$  term omitted, but without success. Numerical solution of the full, unapproximated AEW equation did show that the centers of the two vortices moved farther and farther from the equator as  $c$  decreased for as far as we could follow. Thus, it appears that (7.1) is the only dipole in the limit  $c \rightarrow 0$ . The numerical and analytical evidence is, as noted earlier, that

all dipolar solutions to the AEW equation form a single, continuous branch.

Figure 3 illustrates the small  $c$  limit for  $c = 1/16$ . As predicted, the dipole consists of two widely separated monopoles. (Because of the symmetry, only the northern hemisphere monopole is shown; the southern hemisphere is simply the mirror image of what is shown in Fig. 3.) Both contour plots, (a) and (b) shows the vortex is almost radially symmetric about its center at  $y_0 \approx 4 = 1/c^{1/2}$ . The Coriolis parameter ( $=y$  in our nondimensional coordinates) varies only by a few percent across the narrow diameter of the monopole. Fig. 3(c) shows the weak effects of the variations of  $y$ : the northern hemisphere vortex is a little shorter (by 6.4%) and wider than the monopole predicted by (7.1) and is shifted slightly towards the equator. However, the small  $c$  double monopole approximation (7.1) is clearly very accurate for this small value of  $c$ , 0.0625.

Because the monopole is so concentrated, this case is the most difficult numerical challenge. Inspection of the spectral coefficients showed that 8 basis functions in  $x$  were quite sufficient, but convergence in  $y$  is quite slow. Figure 3 shows that 40 basis functions in  $y$  were adequate. The computation with a total of 320 tensor product basis functions required about an hour on a Mac II, starting from (7.1) as the first guess. The numerical methods are discussed further in Section 9 below.

### 8. Dipoles: numerical results and summary

By applying the Newton/pseudospectral method described in the next section, we found that the three limits described in earlier sections form a single, continuous branch of dipolar solitary waves. All dipoles are antisymmetric with respect to  $y = 0$  and symmetric with respect to  $x = 0$ .

In the limit  $c \rightarrow -1/3$  (Section 4), the vortices are very elongated in the  $x$ -direction so that the contours of constant  $u(x, y)$  are quasi-ellipses with the semimajor axes aligned with the  $x$ -axis (Fig. 2). As  $c$  is decreased to larger and larger negative

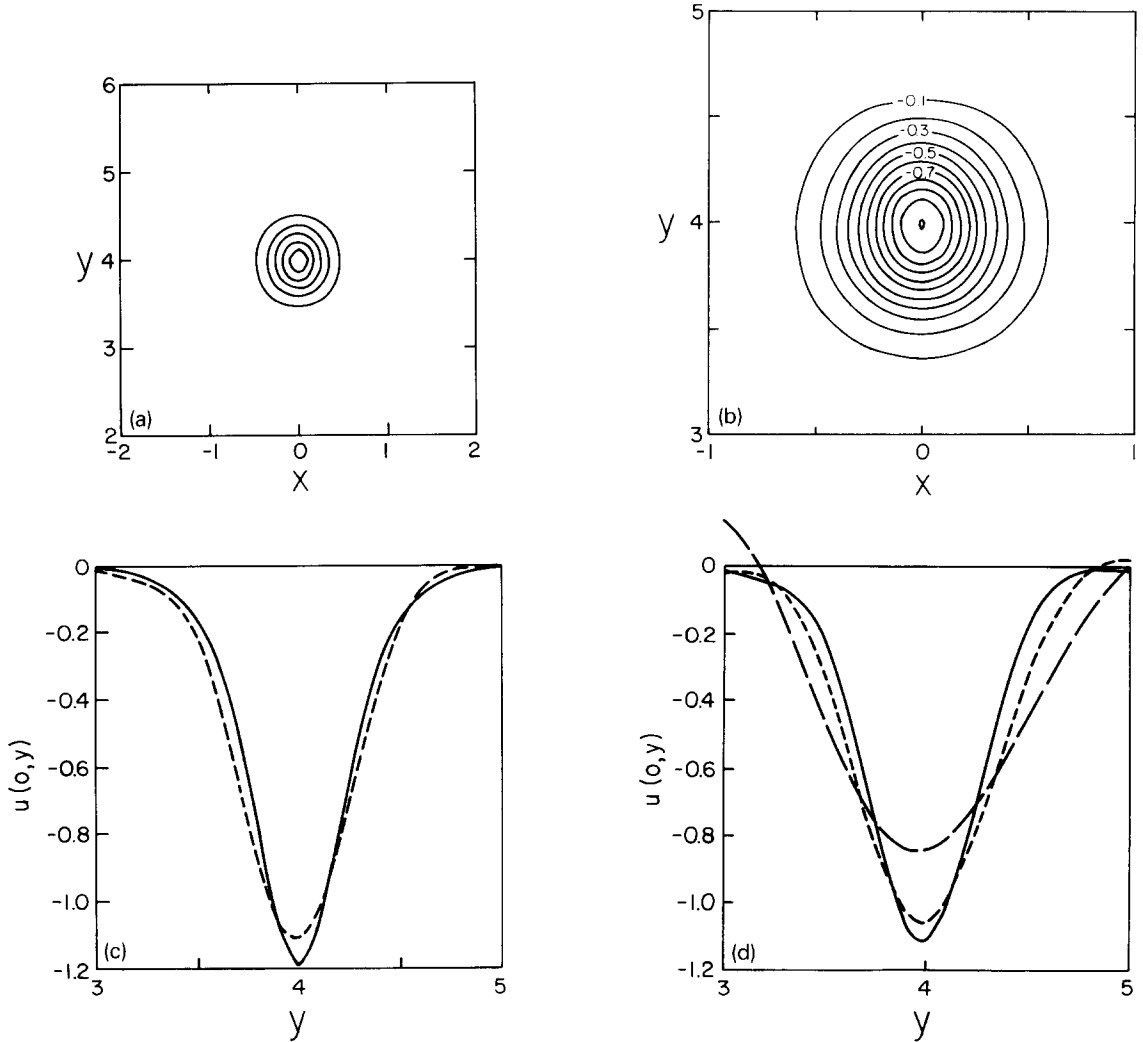


Fig. 3. (a) Contour plot of the dipole for  $c = 1/16$  on the domain  $[-3, 3] \times [0, 6]$ . Because the dipole is antisymmetric with respect to the equator, only the northern hemisphere ( $y > 0$ ) is shown; (b) Blow-up of (a), restricted to  $[-1, 1] \times [3, 5]$ ; (c) Cross-sections of the dipole along the line  $x = 0$  connecting the two vortex centers. Solid: the analytic, small  $c$  approximation. (The shape is that of a Flierl-Petviashvili monopole centered at  $y_0 = 4$ .) Dashed: the exact dipole as computed numerically; (d) Cross-section of the dipole for  $c = 1/16$  as computed using three different spectral truncations. Solid:  $8 \times 40$  basis with scale factors  $L_x = 0.7, L_y = 4$ . Short dashes:  $8 \times 30, L_x = 1, L_y = 4$ . Long dashes:  $8 \times 15$  truncation,  $L_x = 1, L_y = 4$ .

values, the contours become more and more circular. As  $c \rightarrow -\infty$  (Section 6),  $u(x, y) \rightarrow cv(x, y)$  where  $v(x, y)$  solves the limiting eq. (6.3). For all  $c \in [-\infty, -1/3]$ ,  $u(x, y) > 0$  for all  $y > 0$ .

In the limit  $c \rightarrow \infty$ ,  $u(x, y) \rightarrow cv(x, y)$  where  $v(x, y)$  is the same large  $|c|$  universal shape function as for  $c \rightarrow -\infty$ . The difference in the sign of  $c$ , however, implies a difference in the sign of  $u$ : for

all  $c \in [0, \infty]$ , the solitary wave is *negative* for positive  $y$ .

The function  $v(x, y)$  has two maxima in absolute value at  $y \approx \pm 1.4$ . The outer contours are elliptical with semimajor axes aligned along the  $x$ -axis, but the inner contours around the maximum and the minimum are quasi-circular.

As  $c$  decreases (Section 7), the amplitude of  $u$

monotonically decreases to 0—first linearly with  $c$  and then as  $c^{1/2}$  when  $c \ll 1$ . The vortex centers move farther and farther from the equator so that the northern hemisphere minimum and southern hemisphere maximum occur at  $|y| \approx 1/c^{1/2}$ . The vortices become narrower and narrower as  $c$  tends to zero so that the solution for small  $c$ , as predicted by Section 7, consists of two widely separated monopoles of small diameter. The contours of constant  $u$  take the form of narrow, closely spaced concentric circles centered on the maximum and minimum of  $u(x, y)$ ; the area between the monopoles is mostly empty as they rapidly decay below whatever is the smallest nonzero contour.

Figure 4 is a composite that shows one quadrant of each of four representative cases:  $c = \pm 1$  and  $c \rightarrow \pm\infty$ . Because of the four-fold symmetry, each dipole can be recreated by symmetry from the quadrant illustrated. We are forced to omit the limiting cases  $c \approx -1/3$  and  $c \approx 0$  because the contours of constant  $u$  extend to infinity in the  $x$ -

direction in the limit  $c \rightarrow -1/3$  while the centers of the constituent monopoles diverge to infinity as  $c \rightarrow 0+$ . Nonetheless, Fig. 4 does illustrate the overall, finite amplitude behavior.

**9. Numerical methods: the pseudospectral method with rational Chebyshev functions on the infinite interval**

To solve the AEW equation numerically, we set up a Newton iteration by writing

$$u^{(i+1)} \equiv u^{(i)} + d(x, y), \ll 1 \tag{9.1}$$

where the superscript denotes the iteration number. Then discarding terms that are  $O(d^2)$ , we repeatedly solve the linear equation

$$\begin{aligned} d_{xx} + d_{yy} + [-1/c - y^2 - 2yu^{(i)}/c]d \\ = -\{u_{xx}^{(i)} + u_{yy}^{(i)} + [-1/c - y^2]u^{(i)} \\ - y[u^{(i)}]^2/c\}, \end{aligned} \tag{9.2}$$

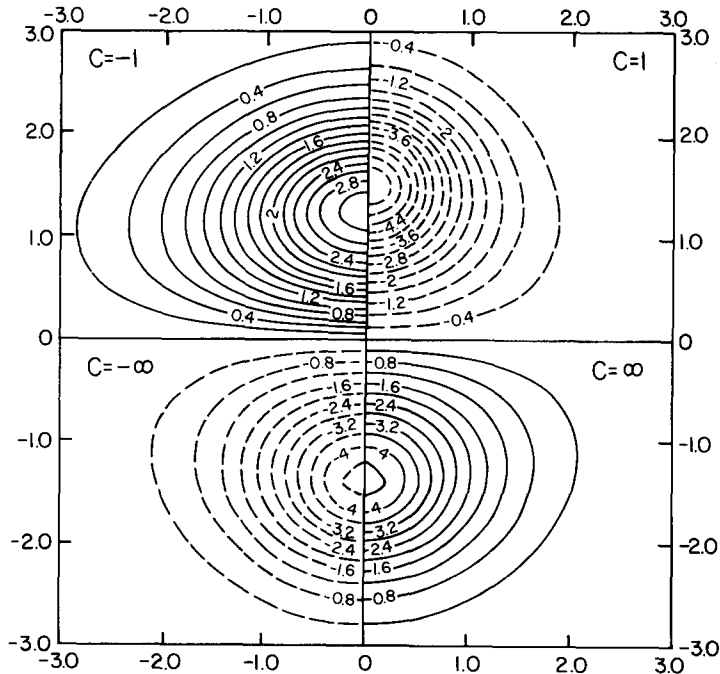


Fig. 4. A composite showing the contours of  $u(x, y)$  in one quadrant for each of four different cases:  $c = 1$  (upper right),  $c \rightarrow \infty$  (lower right),  $c \rightarrow -\infty$  (lower left), and  $c = -1$  (upper left). Each case may be reconstructed in the three quadrants which are not displayed by recalling that  $u(x, y)$  is symmetric about  $x = 0$  and antisymmetric with respect to  $y = 0$ .

update  $u(x, y)$  via (9.1), and repeat until  $d(x, y)$  is sufficiently small.

To solve the linear PDE (9.2), we employed the pseudospectral method as in Boyd [19]. The main difference from that earlier article is that we here employ a tensor product basis appropriate for  $x \in [-\infty, \infty]$ ,  $y \in [-\infty, \infty]$ :

$$u(x, y) = \sum_{m=1}^M \sum_{n=1}^N a_{mn} \phi_{mn}(x, y) \quad (9.3)$$

where

$$\begin{aligned} \phi_{mn}(x, y) \equiv & [\text{TB}_{2m}(x/L_x) - 1] \\ & \times [\text{TB}_{2n+1}(y/L_y) - \text{TB}_1(y/L_y)] \end{aligned} \quad (9.4)$$

where the  $\text{TB}_n(x)$  are the rational Chebyshev functions of Boyd [10]. The parameters  $L_x$  and  $L_y$  are mapping constants that may be adjusted to improve the rate-of-convergence. All the cases shown in Fig. 4 were calculated using  $M = 8$ ,  $N = 15$ ,  $L_x = 2$ , and  $L_y = 3$ . We omit a full discussion because [10] gives all the algorithmic details and theory, but a few comments are in order.

First, because the solitary wave is symmetric in  $x$  and antisymmetric in  $y$ , the basis set can be restricted to only those functions which have the same symmetry. Hence, all the  $\text{TB}_m(x)$  in the  $x$ -basis are of even degree while all the  $y$ -basis functions are of odd degree.

Second, we know that  $u(x, y) \rightarrow 0$  as  $|x^2 + y^2| \rightarrow \infty$ . Consequently, we lose no generality by choosing basis functions which also vanish at infinity. Subtracting 1 or  $\text{TB}_1$  from each basis function guarantees that each  $\phi_{mn}$  individually tends to zero in this limit and reduces the size of the basis set with no loss of accuracy.

Collectively, these two tricks allow an  $8 \times 15$  basis set to give the same accuracy as a general  $18 \times 32$  basis—a reduction from 576 basis functions to only 120 via symmetry. The dense  $120 \times 120$  matrix that is the discretization of (9.1) is solved via Gaussian elimination; a single iteration requires about 4.5 minutes on a 6 Mhz IBM-AT.

The number of degrees of freedom is larger in  $y$  than in  $x$  because the dipole has only a single

maximum in  $x$  (at  $x = 0$ ), but a maximum/minimum pair in  $y$ .

One vice of Newton's method is that it requires a first guess. The analytic, KdV-like approximation (Section 4) gives a good first guess for  $c \approx -1/3$ . One may then increase  $c$  in small steps, using the results at previous values of  $c$  to extrapolate a first guess for  $u(x, y)$  at the new phase speed. This parameter-marching is usually called the "method of continuation" [20]. As noted above, the anti-symmetric dipole forms a single, continuous branch of solutions without limit points, so refinements such as Keller's "pseudo-arclength continuation" [19] are unnecessary.

The rescalings discussed in earlier sections are very helpful, however. A blind march to  $c \approx -\infty$  would obviously require a large number of steps. To compute  $v(x, y)$ , the shape function for large  $|c|$ , we actually solved

$$v_{xx} + v_{yy} - y^2 v - yv^2 = \tau v \quad (9.5)$$

where  $\tau$  is an artificial computational parameter. When  $\tau = 1$ , (9.5) is identical with the full AEW equation (1.1) for  $c = 1$ . When  $\tau = 0$ , (9.5) simplifies to the large  $|c|$  equation, (6.3). As we march from  $\tau = 1$  to 0 in small steps, the *shape* of  $v(x, y; \tau)$  slowly deforms from that of  $u(x, y; c = 1)$  to that of  $v(x, y)$ . Thus, the limiting shape  $v(x, y)$  is computed *directly* in a finite number of steps. The alternative, which is to *asymptotically* approach the limit by computing  $u(x, y; c)$  for larger and larger  $c$ , is much more costly.

We employed two checks. The first was to apply finite differences with a very small grid spacing to evaluate the residual, that is, we evaluated

$$\begin{aligned} R(x, y; h) \equiv & [u(x+h, y) + u(x-h, y) \\ & + u(x, y+h) + u(x, y-h) \\ & - 4u(x, y)] / (h^2) \\ & - [1/c + y^2]u(x, y) \\ & - yu^2(x, y)/c \end{aligned} \quad (9.6)$$

where  $h = 0.0001$  and  $u(x, y)$  is evaluated by summing the spectral series. If  $u(x, y)$  were the exact solution, then the residual would be everywhere

zero. For  $c = -1$ , the worst case of the four in Fig. 4, the maximum value of  $R(x, y; h)$  was smaller than the maximum of  $u$  by a factor of about 150. This method provides a good independent check on the correctness of the pseudospectral program, but it has the disadvantage of extreme pessimism. The reason is that small errors in the coefficients of the high degree basis functions are *magnified* by the double *differentiation*, roughly by a factor of  $n^2$  where  $n$  is the degree of the function. Consequently,

$$\begin{aligned} \max|u(x, y) - u_{\text{numerical}}(x, y)| \\ \sim O(N^2 \max|R(x, y; h)|) \end{aligned} \tag{9.7}$$

where  $N$  is the degree of the highest basis function.

Consequently, we made a second check: comparing solutions with different resolution. Figure 5 displays the difference between two solutions for

the worst of the four solutions shown in Fig. 4,  $c = -1$ . Pseudospectral errors normally oscillate uniformly over the computational domain. This is clearly true in Fig. 5 except that the error rapidly decays for large  $|y|$  because of the equally rapid decay of the exact and numerical solutions.

The fact that the error rapidly oscillates in  $x$  suggests that main source of error is  $x$ -resolution. Figure 4 shows that for  $c = -1$ , the soliton has a larger zonal scale than for the other cases. For this case, it would have been better to either use larger  $L_x$  or distribute the degrees of freedom more evenly between  $x$  and  $y$ . The maximum difference, however, is only  $\pm 0.0013$ , which suggests that the 120-basis function solution is accurate to at least three digits.

The strategy for computing the Flierl–Petviashvili monopole,  $W(r)$ , and the perturbative correction to it is similar, but with some differences. In

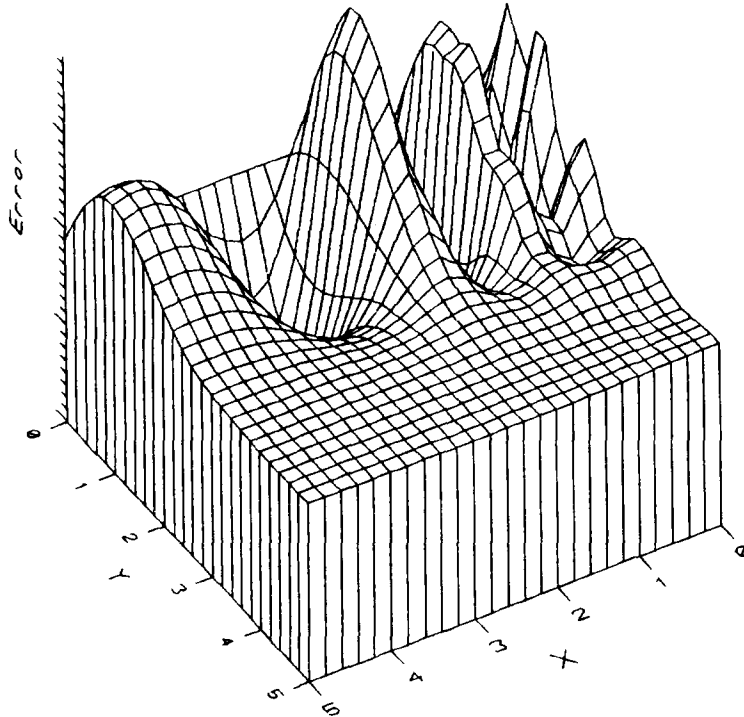


Fig. 5. A surface perspective plot of the upper right quadrant of the difference between the solution with an  $8 \times 15$  basis (120 basis functions) and the solution with 10 degrees of freedom in  $x$  and 18 in  $y$  (180 basis functions). The nearest corner is  $x = 5, y = 5$  and the far corner is the origin  $(0, 0)$ . The tick marks on the vertical axis run from  $-0.0013$  to  $0.0012$ . The phase speed  $c = -1$ . The maximum of the difference is smaller than the maximum of  $u(x, y)$  for this case by about 2400, suggesting that the  $8 \times 15$  solution is accurate to at least three decimal places.

cylindrical coordinates where  $r \in [0, \infty]$ , the obvious basis set is the orthogonal rational functions on the semi-infinite interval of Boyd [21]. However, if one expands a general function  $f(r, \theta)$  in polar coordinates:

$$f(r, \theta) = \sum_{n=0}^{\infty} f_n(r) \cos(n\theta) + g_n(r) \sin(n\theta) \tag{9.8}$$

one may show [20] that the  $f_n(r)$  and  $g_n(r)$  are symmetric about  $r=0$  for even  $n$  while the coefficients for odd  $n$  are antisymmetric about  $r=0$ . Consequently, we used the symmetric factors of our two-dimensional basis functions to compute  $W(r)$  and the first order correction (divided by  $r$ ),  $V(r)$ .

Devising a first guess for the monopole is harder than for the dipole because (2.6) is intrinsically nonlinear and no analytic solutions are available to initialize the continuation method. Two guessing strategies were successful. The first was to solve (2.6) using the pseudospectral method with a *single* basis function,  $\phi_1(r) \equiv TB_2(r/2) - 1$ . This gives a linear equation whose solution is  $a_1 = 1$ . The resulting error is only about 16%, which is sufficiently small that Newton's iteration will converge if  $a_1\phi_1(r)$  is the first guess.

The second strategy is to pick a function that resembles the expected solution. We arbitrarily chose  $W_0 \equiv \exp(-r)$  because a "far field" analysis shows that  $W(r)$  must decay as  $\exp(-r)$  as  $r \rightarrow \infty$ . The key idea is to then *perturb* the problem with an *artificial forcing* which is the residual of this first guess, multiplied by a marching parameter  $\tau$ . Thus, (2.6) becomes

$$W_{rr} + (1/r)W_r - W - W^2 = (1 - \tau)\{W_{0,rr} + (1/r)W_{0,r} - W_0 - W_0^2\}. \tag{9.9}$$

For  $\tau=0$ ,  $W_0$  is the *exact* solution of (9.9) even if  $W_0$  is a terrible approximation to the Flierl-Petviashvili monopole. As we march from  $\tau=0$  to  $\tau=1$  using small steps in  $\tau$ ,  $W(r; \tau)$  slowly deforms

from  $W_0(r)$  to the desired monopole. This method is very powerful because it makes it unnecessary to find a special case which can be solved analytically. The only price for a poor guess for  $W_0(r)$  is that a poor guess will require more steps in  $\tau$ .

Of course, "more steps" may occasionally mean restarting the march with a different guess if  $W(r; \tau)$  runs off to infinity or folds back at a limit point for  $\tau_{\text{limit}} < 1$ , but this did not happen here.

### 10. Unsymmetrical, equator-spanning monopoles for small amplitude: the MKdV-like limit

By applying the method of multiple scales [3], one may show that the AEW equation has another class of solutions which are unsymmetric with respect to the equator. As in Section 4, we make two assumptions: (i) the zonal scale is *large*; and (ii) the wave amplitude is *small*.

However, analysis of Section 4 must be modified because the lowest order solution is the *latitudinally symmetric* function

$$u^0 = A(x) e^{-0.5x^2}. \tag{10.1}$$

The single mode truncation fails because the nonlinear term generated by (10.1) is  $(-y[u^0]^2)$ , which is *antisymmetric* with respect to  $y=0$ . In contrast, the  $y$ -dependent part of (10.1) is  $\psi_0(y)$ , the  $n=0$  Hermite function, which is *symmetric* with respect to the equator. If we truncate to just the  $n=0$  mode, then the nonlinear coefficient is *zero*.

The correct remedy is to apply the method of multiple scales in the same manner as for the even- $n$  equatorial modes in Boyd [3]. For the equatorial problem, one obtains the Modified Korteweg-deVries (MKdV) equation. For the present problem, one finds that  $A(x)$  in (10.1) has the same form as the soliton of the MKdV equation.

The assumption of large zonal scale implies that

$$\partial/\partial x \sim O(b), \quad b \ll 1 \tag{10.2}$$

where  $b$  is a parameter that will be dubbed the "pseudowavenumber". The assumption of small

amplitude is

$$A(x) \sim O(b), \quad b \ll 1 \tag{10.3}$$

If one had no models for guidance, one would assume arbitrary powers of  $b$  in (10.2) and (10.3), work through the analysis, and then adjust the exponents of  $b$  to produce a nontrivial, self-consistent approximation. We will spare the reader this groping-in-the-dark and simply assume what works, taking also

$$c = c_0 + c_2 + \dots \tag{10.4}$$

$$u = u^0 + u^1 + u^2 \tag{10.5}$$

where  $c_2 \sim O(b^2)c_0$ ,  $u^1 \sim O(b)u^0$ , and  $u^2 \sim O(b^2)u^0$ .

The lowest order problem is

$$u_{yy}^0 - [1/c_0 + y^2]u^0 = 0. \tag{10.6}$$

The  $n = 0$  solution is (10.1) provided that

$$c_0 = -1 \tag{10.7}$$

At next order,

$$u_{yy}^1 + [1 - y^2]u^1 = -y[u^0]^2. \tag{10.8}$$

If one uses the identities [22],

$$\begin{aligned} [d^2/dy^2 + (1 - y^2)]\psi_n(y) \\ = -(2n)\psi_n(y) \end{aligned} \tag{10.9}$$

$$y e^{-0.5y^2} = \sum_{n=0}^{\infty} a_{2n+1} \psi_{2n+1}(y) \tag{10.10}$$

where

$$\begin{aligned} a_{2n+1} = \pi^{1/4} (2/3)^{3/2} \{([2n+1]!)^{1/2} \\ \times (-1/3)^n 2^{-(n+1/2)} / n!\} \end{aligned} \tag{10.11}$$

one finds

$$\begin{aligned} u^1 = \{A^2(x)/2\} \\ \times \sum_{n=0}^{\infty} \{a_{2n+1}/(2n+1)\} \psi_{2n+1}(y) \end{aligned} \tag{10.12}$$

At next order,

$$\begin{aligned} u_{yy}^2 + [1 - y^2]u^2 \\ = -2yu^0u^1 - u_{xx}^0 - c_2u^0 \end{aligned} \tag{10.13}$$

Because  $\psi_0(y)$  is a solution of the homogeneous form of (10.13), the second order solution is unbounded unless the R.H.S. of (10.13) is orthogonal to  $\psi_0(y)$ . This condition is equivalent to the statement that the Hermite series of the forcing in (10.13) has a zero coefficient for  $\psi_0(y)$ . Technically, a similar condition applies to (10.8), but is irrelevant because the R.H.S. of that equation contains only Hermite functions of odd degree.

Multiplying the R.H.S. of (10.13) by  $\psi_0(y)$  and integrating over  $y$  from  $-\infty$  to  $\infty$  gives

$$A_{xx} + c_2A + 0.157135A^3 = 0 \tag{10.14}$$

This ODE can be solved in terms of elliptic functions. One finds that this nonlinear eigenvalue problem has a solution

$$A(x) = \pm 3.56762b \operatorname{sech}(bx) \tag{10.15}$$

if and only if

$$c_2 = -b^2 \tag{10.16}$$

Unfortunately, this  $n = 0$  soliton has no counterpart in equatorial oceanography. One can show [22, 23] that modes which have finite phase speeds in the limit of infinitely large zonal scale exist only for  $n \geq 1$ . Consequently, no attempt was made to trace the full solution branch. Nevertheless, the mere existence of such unsymmetrical solitary waves is another dramatic illustration of the diversity of behavior that is possible even in a simple differential equation such as the AEW equation.

### 11. Summary and open problems

Although this study is limited to the steadily translating nonlinear waves of a very simple equation, we find a rich diversity of behavior. Monopoles, that is, vortices that are everywhere one-signed, exist for arbitrary phase speed  $c$  and arbitrary soliton center  $y_0$  in the “ $f$ -plane” approximation. This approximation simplifies the AEW equation by replacing the factors of  $y$  in the coefficients by the constant  $y_0$ . It is interesting that the *shape* of the monopole is described by a single,

simple shape function; the rest of the two-parameter  $f$ -plane family is obtained merely by rescaling the amplitude and width of this single parameter-free function,  $W(r)$ . It is even more interesting that allowing the coefficients to be variable gives a solvability condition at first order in perturbation theory which reduces the two-parameter family to one by creating an “eigenrelation” between the phase speed  $c$  and the soliton center latitude  $y_0$ . The perturbation theory is accurate when  $0 \leq c \ll 1$ .

The AEW equation has a family of dipoles which are antisymmetric about the “equator”,  $y = 0$ , and one-signed everywhere within a hemisphere. In the limit  $c \rightarrow -1/3$ , the dipoles are of very small amplitude and very large zonal scale. The method of multiple scales or an equivalent argument given in Section 4 shows that dipoles may be approximated by the product of  $\text{sech}^2$  in  $x$  and  $y \exp(-0.5y^2)$ .

The second analytic regime is for small positive  $c$ . Again, the dipole has very small amplitude. The dipole is approximated by two widely separated, rapidly decaying monopoles of opposite sign in opposite hemispheres.

In the opposite extreme of very *large* amplitude, one obtains a third analytic regime: for  $|c| \gg 1$ ,  $u(x, y) \equiv v(x, y)/c$  where  $v(x, y)$  is a single, parameter-free function that must be computed numerically. Although solitons do not exist on the range  $c \in [-1/3, 0]$ —the interval of the phase speeds of *linear* waves that are sinusoidal in  $x$ —the antisymmetric dipoles form a continuous branch of solutions. The two branches for negative and positive  $c$  are connected through infinity because  $u(x, y)$  in both the limit  $c \rightarrow \infty$  and  $c \rightarrow -\infty$  may be described by the same shape function,  $v(x, y)$ .

The quasi-geostrophic equation, which is a close cousin of the AEW equation, does not allow monopoles. However, it does have dipolar solutions known as “modons” [6]. Instead of being antisymmetric about the equator,  $y = 0$ , these are antisymmetric about some reference latitude  $y = y_0$ , and solve a constant coefficient nonlinear equation analogous to the AEW in the “ $f$ -plane”

approximation. It is trivial to show that such solutions do not exist on the AEW  $f$ -plane (2.1): the quadratically nonlinear term will be symmetric about  $y = y_0$  even if  $u(x, y)$  is antisymmetric, but the other terms preserve the antisymmetry. All solutions must have a symmetric component.

Monopoles but no modons; just the reverse for quasi-geostrophy. It is clear that solitons are sensitive in subtle ways to the characteristics of the equations that model them. The AEW/quasi-geostrophic comparison is a two-dimensional counterpart of the duality between the Korteweg-deVries equation and the Regularized Long Wave equation: both are equally consistent with the perturbation theory used to derive them from more complex systems, but the KdV is integrable and the RLW is not, and has solitons of both positive and negative sign.

Another subtlety is that the method of multiple scales and the Hermite-Galerkin resonance argument of Section 4 both predict the existence of higher mode solitons: waves whose latitudinal structure is described by high order Hermite functions. As explained in Section 5, these modes leak energy to large  $|x|$  through radiation whose latitudinal structure is described by the first Hermite function,  $\psi_1(y)$ , and therefore technically do not exist! The rate of energy leakage, however, is exponentially small in the perturbation parameter: when the error in the lowest order is  $O(10\%)$ , the amplitude of the radiation for large  $|x|$  is  $O(10^{-14})$  relative to that of the localized, isolated part of the wave!

For such quasi-solitons, a rigorous existence proof is not merely irrelevant, as such proofs often are in engineering, but exceedingly misleading. The physical effect (in this case, the radiative leakage) which make the soliton mathematically nonexistent is an effect which is exponentially small when the quasi-soliton is small. Experimentally, small amounts of damping and other small errors would make such quasi-solitons indistinguishable from true solitons for sufficiently small amplitude. Boyd [1] is a comprehensive



review of such “weakly non-local” solitons, which are now known to occur in plasma physics, particle physics, and oceanography, and probably in many other fields as well.

The numerical calculations show that the orthogonal rational functions of Boyd [10] are extremely effective in solving two-dimensional soliton problems, both in polar coordinates, which we used to compute the axisymmetric monopoles, and in Cartesian coordinates, which we employed in our dipole studies. Refinements such as preconditioned iterations (for solving large matrix problems) and pseudoarclength continuation (for following a solution curve around what is known variously as a “limit point”, “fold”, or turning point”) were unnecessary. Gaussian elimination, Newton’s iteration, and simple continuation were sufficient along with a few hours of time on an IBM-AT.

The major open problem for the AEW equation is: are there other solitary waves besides the ones computed here? In general, this is a very difficult question.

We are on surer ground in predicting that the numerical and analytical ideas developed here will be useful in attacking other, harder problems. Soliton theory, like so much of applied mathematics, is an art as much as science. A problem like this, which is simple but has complex and diverse solutions, is useful in becoming sensitive to the possibilities.

**Acknowledgments**

This work was supported by NSF Grants OCE8509923, OCE8800123 and DMS8716766. I thank the editor and reviewer for constructive suggestions. The reviewer pointed out the analytical proof that  $\lambda = 1$  in (3.8) and suggested the name “Ageostrophic Equatorial Wave” equation. I also thank Academic Press for permission to reprint Fig. 4 from [1].

**Appendix. A heuristic derivation of the AEW equation**

The nonlinear equatorial shallow water wave equations have dipole solitons which are antisymmetric about the equator,  $y = 0$ . Boyd [2, 3] has analyzed these dipoles, but only using the method of multiple scales in the limit  $c \rightarrow -1/3$ . After initial attempts at a numerical study were unsatisfactory, I decided to do this preliminary study of a simpler model. The shallow water wave equations have the exact conservation law

$$Dq/dt = 0 \tag{A.1}$$

where  $D/dt$  denotes the usual convective derivative and where  $q$ , the potential vorticity, is defined by

$$q \equiv (v_x - u_y + y)/(1 + \phi) \tag{A.2}$$

where  $u$  and  $v$  are the usual (nondimensional) Cartesian velocities and  $\phi$  is the nondimensional pressure (or equivalently, surface height). As it stands, (A.1) is not a closed model because  $u$ ,  $v$ , and  $\phi$  are independent quantities. One of the technical complications is that the shallow water wave equations are a set of three coupled equations rather than a single scalar equation like the AEW, (1.1).

At high latitudes, however, one may apply the “quasi-geostrophic” approximation to show that all three unknowns in (A.2) may be written in terms of a single variable  $\psi$ , the streamfunction:

$$v \approx \psi_x, \quad u \approx -\psi_y, \quad \phi \approx y\psi \tag{A.3}$$

In strict quasi-geostrophy, one also makes some additional approximations that include linearizing the denominator of (A.2), but we shall retain (A.2) as is. Malanotte-Rizzoli [2] similarly incorporates some non-geostrophic effects in one of her model equations.

If we shift into a coordinate system which is travelling with wave at phase speed  $c$ —we shall use  $x$  to denote the zonal direction in this reference

frame everywhere except in the three previous equations—and apply (A.3), conservation of potential vorticity becomes

$$J(q, \psi + cy) = 0 \quad (\text{A.4})$$

where we assume that the wave is stationary and independent of time in the moving reference frame and where the Jacobian operator is, as usual, defined by

$$J(A, B) = A_y B_x - A_x B_y \quad (\text{A.5})$$

It is easy to show [23] that the vanishing of the Jacobian implies that the contours of  $A$  parallel those of  $B$ , which in turn means that (A.4) is equivalent to

$$q = F(\psi + cy) \quad (\text{A.6})$$

for some arbitrary  $F(x)$ . If we assume that all the contours of  $(\psi + cy)$  extend to infinity, then we may evaluate (A.6) by observing that  $\psi \rightarrow 0$  for large  $|x^2 + y^2|$ —this is the very definition of an “isolated” soliton. We find  $F(x) = x/c$ . Substituting this into (A.6), multiplying both sides by  $(1 + y\psi)$ , and cancelling common factors gives

$$\Delta\psi + (-1/c - y^2)\psi = y\psi^2/c \quad (\text{A.7})$$

which is simply the AEW equation, (1.1). We use  $u$  as the unknown instead of  $\psi$  in the rest of the article to avoid confusion with  $\psi_n(y)$ , which is the standard notation for the Hermite functions, but the proper geophysical interpretation of the unknown in the AEW equation is that it is the *streamfunction*.

Several comments are in order. First, the quadratic nonlinearity in (A.7) arises because we retained the non-geostrophic  $(1 + \phi) \approx (1 + y\psi)$  in the denominator of the potential vorticity. The analogous quasi-geostrophic equation is *linear* and has no soliton solutions. To obtain quasi-geostrophic solitary waves, it is necessary to relax one or more of the implicit assumptions of quasi-geostrophy by (i) including nongeostrophic effects, as here; or (ii) adding topographic variations; or (iii)

assuming that some contours of  $(\psi + cy)$  are *closed*, and thus cannot be determined by evaluating (A.6) at infinity, which permits “modons” [6].

Second, it is usual to approximate terms that vary with  $y$  as constants except perhaps where differentiated. We differ from Malanotte-Rizzoli’s own model by allowing  $y$  to vary so as to obtain an equation which is (hopefully!) valid even at the equator  $y = 0$ , and allows solitons that span both hemispheres.

For  $c \approx -1/3$ , one may compare the AEW solitons of Section 4 with those of Boyd [2, 3]. The linear part of (A.7) agrees exactly with the linear equation satisfied by the north-south current of the equatorial shallow water wave equations in the limit of infinite zonal scale. The result is that the linear dispersion relation is correct in the KdV limit, and so also is the latitudinal shape of  $v$ . One may show, however [22] that the structures of  $u$  and  $\phi$  are distorted by about 30% by the ansatz (A.3). Further, the nonlinearity in (A.6) is too weak; the AEW solitons are about 2.4 times larger than their equatorial counterparts.

For larger  $c$ , the differences between AEW and shallow water waves are likely to be greater. There are good reasons why the approximation of writing  $u, v$ , and  $\phi$  in terms of a streamfunction  $\psi$  is normally employed only at high latitudes!

Of course, this would suggest that the monopole analysis of Sections 2 and 3 would be trustworthy. Unfortunately, the predicted minimum of  $y\psi$  is  $-4.76$ , independent of  $y_0$ , which implies that the denominator of the potential vorticity changes sign. Physically, this would imply that the layer depth vanishes, and a different model is needed. It is striking, however, that the same monopole arose from the quite different analyses of Flierl [11] and Petviashvili [12].

The conclusion is that analyzing solitons of the full fluid equations is quite complex because apparently small modifications to approximating equations can completely alter the properties of solitons, or even alter their very existence. Yano and Tsujimure [7] give the most recent review; part of the problem is that there are at least five

fundamental parameters which may all vary over broad ranges.

In this article, we shall be content to use (1.1) as a good heuristic model. One final defense is appropriate, however. Flierl's quasi-geostrophic model replaces the Coriolis parameter by a constant but has  $y$ -varying terms because of the mean current. Swenson [13] showed that this leads to a contradiction at perturbative first order. In our model, the Coriolis parameter is allowed to vary, and monopoles exist. The AEW equation cannot be rigorously justified, but the identification of better geophysical models that are both simple and fully consistent is still a frontier.

## References

- [1] J.P. Boyd, "New directions in nonlinear periodic waves: polynoidal waves, imbricated solitons, weakly nonlocal solitary waves, and numerical boundary value algorithms", in: T.-Y. Wu and J.W. Hutchinson, eds., *Advances in Applied Mechanics* 27, Academic Press, New York (1989) 1-82.
- [2] P. Malanotte-Rizzoli, "Planetary solitary waves in geophysical flows", *Adv. Geophys.* 24, 147-224 (1982).
- [3] J.P. Boyd, "Equatorial solitary waves. Part I: Rossby solitons", *J. Phys. Oceanogr.* 10, 1699-1717 (1980).
- [4] J.P. Boyd, "Equatorial solitary waves. Part 3: Westward-traveling modons", *J. Phys. Oceanogr.* 15, 46-54 (1985).
- [5] H.G. Marshall and J.P. Boyd, "Solitons in a continuously stratified equatorial ocean", *J. Phys. Oceanogr.* 17, 1016-1031 (1987).
- [6] G.R. Flierl, "Isolated eddy models in geophysics", *Ann. Revs. Fluid Mech.* 19 493-530 (1987).
- [7] J.-I. Yano and Y.N. Tsujimura, "The domain of validity of the KdV-type solitary Rossby waves in the shallow water  $\beta$ -plane model", *Dyn. Atmos. Oceans* 11, 101-129 (1987).
- [8] J.C. McWilliams, "Geostrophic vortices", in: *Nonlinear Topics in Ocean Physics: Proceedings of the 1988 Fermi School*, ed. by A.R. Osborne and L. Bergamasco, North-Holland, in press (1990).
- [9] J.C.J. Nihoul and B.M. Jamart, eds., *Mesoscale/Synoptic Coherent Structures in Geophysical Turbulence*, Elsevier, Amsterdam (1989).
- [10] J.P. Boyd, "Spectral methods using rational basis functions on an infinite interval", *J. Comp. Phys.* 69, 112-142 (1987).
- [11] G.R. Flierl, "Baroclinic solitary waves with radial symmetry", *Dyn. Atmos. Oceans* 3, 15-38 (1979).
- [12] V.I. Petviashvili, "Red Spot of Jupiter and the drift soliton in a plasma", *JETP Lett.* 32, 619-622 (1981).
- [13] M. Swenson, "A note on baroclinic solitary waves with radial symmetry", *Dyn. Atmos. Oceans* 10, 243-252 (1986).
- [14] J.P. Boyd, "A comparison of numerical and analytical methods for the reduced wave equation with multiple spatial scales", *Appl. Numer. Math.*, in press (1991).
- [15] J.P. Boyd, "A numerical calculation of a weakly non-local solitary wave: the  $\phi^4$  breather", *Nonlinearity* 3, 177-195 (1990).
- [16] J.P. Boyd, "Weakly non-local solitons for capillary-gravity waves: fifth-degree Korteweg-deVries equation", *Physica D*, in press (1991).
- [17] G.P. Williams and R.J. Wilson, "The stability and genesis of Rossby vortices", *J. Atmos. Sci.* 49 207-240 (1988).
- [18] J.P. Boyd, "Weakly nonlocal solitary waves", in: J.C.J. Nihoul and B.M. Jamart, eds., *Mesoscale/Synoptic Coherent Structures in Geophysical Turbulence*, Elsevier, Amsterdam (1989) 103-112.
- [19] J.P. Boyd, *Chebyshev and Fourier Spectral Methods*, Springer, Berlin (1989) 792 pp.
- [20] J.P. Boyd, "Solitons from sine waves: analytical and numerical methods for non-integrable solitary and cnoidal waves", *Physica* 21D, 227-246 (1986).
- [21] J.P. Boyd, "Orthogonal rational functions on a semi-infinite interval", *J. Comp. Phys.* 70, 63-88 (1987).
- [22] J.P. Boyd, *Dynamics of the Equatorial Ocean*, Springer, Berlin, to appear, 450 pp.
- [23] J. Pedlosky, *Geophysical Fluid Dynamics*, 2d. ed., Springer, New York (1987) 500 pp.