Benjamin-Feir Instability of Rossby Waves on a Jet

By J. G. ESLER[1]*

[1]*University College London, United Kingdom*

SUMMARY

Large-scale waves on the extratropical tropopause have been widely observed to spontaneously organise into groups or wave packets. Here, a simple paradigm for this wave packet formation is presented. Firstly, a weakly nonlinear theory of Rossby wave propagation on a potential vorticity front, based on small non-dimensional wave amplitude ϵ, is developed. As is typical for systems allowing conservative one-dimensional wave propagation, the evolution of the wave envelope is governed by the nonlinear Schrödinger equation. The sense of the nonlinearity is consistent with Benjamin-Feir instability, where uniform wavetrains are unstable to sideband modulations, leading to the formation of wave packets. Next, numerical results from contour dynamics integrations show that the weakly nonlinear predictions for sideband growth rates are quantitatively accurate up to $\epsilon \sim 0.5$, and that unstable sideband growth is qualitatively similar, but slower than predicted, at higher values of ϵ. For $\epsilon \geqslant 0.6$ the formation of wave packets leads to wave breaking, this occuring at much lower initial wave amplitudes than for unmodulated uniform wavetrains previously studied. The numerical results reveal that the length and time scales of the Benjamin-Feir instability are broadly consistent with observed wave packet formation in the extratropics.

KEYWORDS: Baroclinic waves   Potential Vorticity Front   Wave Packets

1. INTRODUCTION

The instability of finite amplitude Stokes' wavetrains on water to sideband disturbances, or long wavelength modulations in the amplitude of the wavetrain, was first noted by Benjamin and Feir (1967). This discovery subsequently led to the development of a general theory for the Benjamin-Feir instability of travelling waves. A central result is that one-dimensional weakly nonlinear waves with an amplitude dependent dispersion relation $\omega = \omega_0(k) + \omega_2(k)a^2 + O(a^4)$ will be Benjamin-Feir unstable if and only if the Whitham (1974) criterion

$$\frac{d^2\omega_0}{dk^2}.\omega_2 < 0 \tag{1}$$

is satisfied. To determine the sign and magnitude of the nonlinear correction ω_2, and hence the stability characteristics of the uniform wavetrain, the detailed dynamics of the specific travelling waves must be investigated. Several approaches to this problem have been adopted including the Fourier decomposition method of Benjamin and Feir (1967) and the variational approaches described by Whitham (1974, § 13-14) and Bridges and Mielke (1995). Here we shall follow the weakly nonlinear wave envelope approach developed in the water wave context by Hasimoto and Ono (1972). Numerous travelling wave problems have been investigated using the above mentioned methods. In the geophysical context these include equatorial Kelvin waves (Boyd, 1983), oceanic frontal waves (Slomp and Swaters, 1999; Karsten and Swaters, 2000), internal gravity waves (Grimshaw, 1977; Sutherland, 2001) and Kelvin-Helmholtz waves near the onset of instability (Benjamin and Bridges, 1997).

For the case of a uniform Rossby wavetrain on an unbounded β-plane, Newell (1969) showed that the Benjamin-Feir instability is present, but that it is

invariably of secondary importance as the wavetrain is also subject to a resonant triad instability that operates on a faster timescale. Plumb (1977) extended Newell's results to the case of an infinite zonal channel, bounded by sidewalls at $y = 0,\ L$ in the meridional direction. In this case, the triad instability is absent for waves with the smallest possible meridional wavenumber ($l = \pi/L$) provided they have zonal wavenumber less than a critical value ($k < 0.681\pi/L$). For wavetrains satisfying these criteria, Plumb showed, the Benjamin-Feir mechanism becomes the dominant mode of instability.

Here it is argued that Benjamin-Feir unstable wavetrains, such as those discovered by Plumb, are relevant to the situation at the extratropical tropopause, where their evolution may resemble the observed formation of wave packets of synoptic-scale Rossby waves (often referred to as baroclinic waves). Observed wave packet formation has dynamics resembling that of a sideband instability, such as Benjamin-Feir, rather than a resonant triad instability. This paper will demonstrate that Benjamin-Feir instability always dominates in another simple, paradigmatic model describing Rossby wave propagation. In this model, which is arguably more relevant to the situation at the extratropical tropopause (e.g., Swanson et al., 1997), Rossby waves are confined to propagate on an interface separating regions of constant potential vorticity (PV hereafter). In our model, the stability properties of the Rossby wavetrain are not found to be subject to abrupt transitions, that depend for example on the ratio of the wavelength to a channel width, as in the case of Plumb (1977). Further, when our model is 'tuned' as far as possible to conditions at the extratropical tropopause, realistic-looking wave packet formation occurs efficiently due to the Benjamin-Feir mechanism on timescales of a few days. It is argued that the wave packet formation in our model therefore represents a useful dynamical paradigm for the observed wave packet formation.

The ubiquitous nature of Rossby wave packets at the extratropical tropopause (baroclinic wave packets), and the associated 'downstream development' of new eddies associated with their group propagation, have been identified in numerous observational studies (Chang, 1993; Lee and Held, 1993; Berbery and Vera, 1996; Chang and Yu, 1999; Chang, 2001). Coherent wave packets have been observed to propagate up to twice around the globe in the southern hemisphere, so although longitudinal variations in baroclinicity (due to e.g. large-scale topography and land-sea contrasts) act to modulate their amplitudes, they cannot account for the existence and persistence of the wave packets themselves. The wave packets are arguably easiest to identify using Hovmoeller (longitude-time) plots of unfiltered meridional wind at the height and latitude of the tropopause ($\sim$ 300 mb, 40-55°N or S). However, it is also of interest to examine their development on isentropic maps of PV (e.g. Hoskins et al., 1985), as this framework allows the waves to be observed directly as fluctuations to a (near) materially conserved tracer field. Figure 1 shows a sequence of plots from consecutive days of ECMWF operational analyses of Ertel's PV on the 340 K isentropic surface during southern hemisphere winter. The extratropical dynamical tropopause may be identified on this plot as the sharp jump in PV, and following widespread convention it is labelled with the -2 Potential Vorticity Unit (PVU) contour (PV values less than this are shaded). The Rossby waves can be identified with the wave-like disturbances to the -2 PVU contour. In Figure 1 it can be seen the waves are concentrated mostly in two wave packets at the locations of the black arrows, and that these wave packets exhibit group propagation at around 25-35° day^{-1} (consistent with e.g.

Figure 5 of Chang, 1999). By contrast the phase speeds of individual waves are $\sim 10°$ day^{-1}. Because the group speed of the waves exceeds their phase speed, new eddies are constantly formed at the downstream edge of each wave packet, while those at the upstream edge decay away. This phenomenon has been described by Chang (1993) as 'downstream development' and can be seen in Figure 1. Over the three days shown, within the larger wave packet propagating through the western hemisphere, the wave crests 'A' and 'B' decay, while a new crest forms at 'E'. In the smaller wave packet in the eastern hemisphere 'X' decays while a new crest forms at 'Z'. The robustness of the group propagation and the strong tendency for the waves to self-organise as wave packets, as discovered and detailed by the observational studies listed above, is notable even though the waves of largest amplitude are often observed to break. It is this wave breaking that leads to exchange of air masses between the stratosphere (high PV) and the troposphere (low PV) (Thorncroft et al., 1993).

These considerations motivate the choice of model used in this paper, the intention being to demonstrate the possible importance of the Rossby wave Benjamin-Feir instability in a simple but relevant context. Previous analytical (Esler, 1997) and numerical (Lee and Held, 1993; Swanson and Pierrehumbert, 1994; Esler and Haynes, 1999) studies of wave packets in the quasi-geostrophic two-layer model have not revealed a simple paradigm explaining their formation. The approach taken follows Nakamura and Plumb (1994), Swanson et al. (1997) and Swanson (2000, 2001) by representing the tropopause as a single contour dividing two regions of constant quasi-geostrophic PV in a single layer shallow water flow. Rossby waves appear as waves on the PV jump defined by the contour, as opposed to wave-like disturbances propagating on a continuous planetary PV gradient. This type of simple model is relevant to real geophysical flows because Rossby waves typically act to reinforce and sharpen PV gradients on which they propagate. This PV gradient reinforcement occurs because away from the PV jump the Rossby waves act to homogenise PV gradients in nonlinear critical layers (see e.g. Rhines, 1994). This has led Haynes et al. (2001) to argue that it is the presence of the Rossby waves themselves that generates the sharp gradient in PV observed at the extratropical tropopause. For our choice of a single layer model, it is important to emphasise that we are concentrating on the dynamics of Rossby waves at the extratropical tropopause viewed in isolation, with the possibilities of both baroclinic growth and dissipation† suppressed.

Contour models have also been widely used to demonstrate dynamical paradigms in geophysical fluids, including Rossby wave breaking (Polvani and Plumb, 1992) and hydraulic effects (Haynes et al., 1993). Considerable attention has also been given to the long-wave limit for Rossby wave propagation (Pratt and Stern, 1986; Nycander et al., 1993), where dispersive and nonlinear effects become comparable. Weakly nonlinear analysis in the long-wave limit is to be distinguished from the nonlinear approach taken in section 2, where it is assumed that it is the *wave envelope* that varies on an asymptotically large spatial scale rather than the waves themselves. Connections between generic long-wave behaviour governed by the KdV equation and Benjamin-Feir instability governed

---

† A model with comparable dynamics, albeit in a profoundly different geometry, might be obtained by considering the behaviour of an isolated potential temperature front on the upper boundary of an Eady-type model. Baroclinicity may be suppressed in this model by assuming constant potential temperature on the lower boundary.

by the nonlinear Schrödinger envelope equation have recently been elucidated by Grimshaw et al. (2001).

In section 2 we derive the nonlinear Schrödinger envelope equation for Rossby waves in the single contour model. Various predictions for nonlinear frequency corrections and sideband growth rates are tested by numerical contour dynamics integrations in section 3. Finally in section 4 the implications for the dynamics of the atmosphere and oceans are discussed.

2. Weakly nonlinear behaviour of Rossby waves on a potential vorticity front

The basic model used in both this section and the next is that of an equivalent barotropic, quasi-geostrophic shallow water flow on an f-plane. The dynamics are determined by the position of a single contour at $y = \eta(x, t)$ separating two regions of constant quasi-geostrophic PV. Pratt (1988) refers to such a dividing contour as a PV front. The flow streamfunction ψ is related to the position of the contour through a Helmholtz equation of the form

$$f + \nabla^2\psi - \frac{\psi}{L_R^2} = q = \begin{cases} Q + \Delta/2 & \text{for} \quad y > \eta(x, t) \\ Q - \Delta/2 & \text{for} \quad y < \eta(x, t) \end{cases} \quad (2)$$

Here $L_R = \sqrt{gH}/f$ is the Rossby radius of deformation in terms of the layer depth H, gravity g and reference Coriolis parameter f. Note that f and the background PV value Q do not otherwise enter the dynamics below, and we are neglecting the planetary vorticity gradient. In the case where the contour is undisturbed ($\eta \equiv 0$) the basic velocity is given by $\mathbf{u} = (U(y), 0)$ where

$$U(y) = U_0 \exp\left(-|y|/L_R\right). \quad (3)$$

where U_0, the velocity at the contour, is determined below.

The kinematic equation that determines the evolution of the contour position $y = \eta(x, t)$ in time is

$$\eta_t + u\eta_x = v \quad \text{on} \quad y = \eta(x, t). \quad (4)$$

Following e.g. Dritschel (1986) the velocity $\mathbf{u} = (u, v) = (-\psi_y, \psi_x)$ can be obtained from the contour position $\eta(x, t)$ by evaluating a contour integral along the PV front, using the Green's function for the Helmholtz equation (2)

$$\mathbf{u}(\mathbf{x}) = -\Delta \int_C G(\mathbf{x}, \mathbf{x}') \, d\mathbf{x}'. \quad (5)$$

Here the integral is taken along the contour. The appropriate Green's function solution for (2) is given by (Swanson, 2001)

$$G(\mathbf{x}, \mathbf{x}') = -\frac{1}{2\pi} K_0 \left(\frac{|\mathbf{x} - \mathbf{x}'|}{L_R}\right) \quad (6)$$

where K_0 is a modified Bessel function in standard notation ($K_0(r) \to 0$ as $r \to \infty$).

Considering simple flows of this type, Dritschel (1988) showed as a corollary to Arnol'd's first theorem (Arnol'd, 1966), that a direct consequence of the conservation of streamwise impulse over the domain

$$J = \int\int yq \, dx \, dy \quad (7)$$

is that the integral of the square of the of the contour displacement is conserved in these flows, i.e.,

$$\frac{d}{dt} \int \eta^2 \, dx = 0. \tag{8}$$

This provides an important constraint on the evolution of the disturbance on the contour in both the weakly nonlinear theory and the fully nonlinear contour dynamics integrations to follow.

The asymptotic analysis of this section depends on the assumption of small contour slope, which can be most generally defined by imposing the global condition $|x_1 - x_2| \gg |\eta(x_1) - \eta(x_2)|$ for any two points with x-coordinates x_1, x_2. This condition is relaxed in the contour dynamics integrations of section 3. First, we note that equation (5) can be written in the form

$$(u, v) = -\Delta \int_{-\infty}^{\infty} G(\mathbf{x}, \mathbf{x}')(1, \eta'(x')) \, dx' \tag{9}$$

for any single-valued differentiable position function η. Next, we use the expansion

$$
\begin{aligned}
|\mathbf{x} - \mathbf{x}'| &= \left[(x - x')^2 + (\eta(x) - \eta(x'))^2 \right]^{1/2} \\
&= |x - x'| \left(1 + \tfrac{1}{2} \left(\frac{\eta(x) - \eta(x')}{x - x'} \right)^2 + ... \right),
\end{aligned}
\tag{10}
$$

which is clearly convergent if the small slope condition above is satisfied everywhere. Expanding K_0 in a Taylor series using this expression, the leading order terms are

$$K_0 \left(\frac{|\mathbf{x} - \mathbf{x}'|}{L_R} \right) = K_0 \left(\frac{|x - x'|}{L_R} \right) + K_0' \left(\frac{|x - x'|}{L_R} \right) \frac{(\eta(x) - \eta(x'))^2}{2L_R|x - x'|} + ... \tag{11}$$

and we obtain

$$(u, v) = \frac{\Delta}{2\pi} \int_{-\infty}^{\infty} \left[K_0 \left(\frac{|x - x'|}{L_R} \right) + K_0' \left(\frac{|x - x'|}{L_R} \right) \frac{(\eta(x) - \eta(x'))^2}{2L_R|x - x'|} + ... \right] (1, \eta'(x')) \, dx'. \tag{12}$$

We shall show below that if the PV contour satisfies the small slope condition then the terms included in (12) are sufficient to obtain the velocity field at the required order from the contour position. The solution for no displacement ($\eta \equiv 0$) is

$$U_0 = \frac{\Delta}{2\pi} \int_{-\infty}^{\infty} K_0 \left(\frac{|x'|}{L_R} \right) dx' = \frac{\Delta L_R}{2}. \tag{13}$$

We proceed by assuming that at first order, the contour displacement is given by a modulated harmonic wavetrain

$$\epsilon \eta_0 = \frac{\epsilon}{2k} \left(A(X, T, \tilde{T}) \, \exp\left(ikx - i\omega_0 t \right) + \text{ c.c.} \right), \tag{14}$$

Here A is an amplitude function that depends on 'slow' space and time‡ variables $X = \epsilon x$, $T = \epsilon t$ and $\tilde{T} = \epsilon^2 t$. The small parameter ϵ is a measure of the non-dimensional wave amplitude. It is easily seen that terms neglected in (12) are

‡ The use of only one 'slow' space variable may be justified by defining the domain length scale to be $\mathrm{O}(\epsilon^{-1})$. However, if a further space scale $\tilde{X} = \epsilon^2 x$ were to be introduced, the analysis would remain unaltered under the transformation $\tilde{T} \to \tilde{T} + \tilde{X}/\omega_{0k}$.

$O(\epsilon^4)$ and higher. Further, we assume that all fields may be expanded in ϵ, so that

$$
\begin{aligned}
\eta &= & \epsilon(\eta_0 + \epsilon\eta_1 + \epsilon^2\eta_2 + ...) \\
u &= U_0+ & \epsilon(u_0 + \epsilon u_1 + \epsilon^2 u_2 + ...) \\
v &= & \epsilon(v_0 + \epsilon v_1 + \epsilon^2 v_2 + ...)
\end{aligned}
\tag{15}
$$

As the u component of (12) is quadratic in η, it is clear that $u_0 \equiv 0$. Expanding the derivatives in (4) in the 'slow' space and time variables, and equating powers of ϵ gives,

$$
\eta_{0t} + U_0\eta_{0x} = v_0 \qquad \text{at leading order} \tag{16}
$$
$$
\eta_{0T} + U_0\eta_{0X} + \eta_{1t} + U_0\eta_{1x} = v_1 \qquad \text{at first order} \tag{17}
$$
$$
\eta_{0\tilde{T}} + \eta_{1T} + U_0\eta_{1X} + \eta_{2t} + U_0\eta_{2x} + u_1\eta_{0x} = v_2 \qquad \text{at second order.} \tag{18}
$$

Details of the evaluation of the integral expressions (12) are given in Appendix A, and we shall refer to results from there throughout the remainder of this section. The leading order velocity is shown there to be

$$
v_0 = \frac{\Delta L_R}{4(1 + k^2 L_R^2)^{1/2}} \left(iA(X, T, \tilde{T}) \exp\left(ikx - i\omega_0 t\right) + c.c. \right) \tag{19}
$$

Inserting η_0, v_0 into (16) we recover the linear dispersion relation for Rossby waves on the contour

$$
\omega_0 = U_0 k - \frac{\Delta k L_R}{2(k^2 L_R^2 + 1)^{1/2}} \tag{20}
$$

Note that this expression can also be obtained by using a linearisation condition relating the jump in the disturbance zonal flow $u' = -\psi_y'$ at the location of the contour to the contour displacement, $[\psi_y'] = -\Delta\eta$ (e.g. Swanson et al., 1997). We have chosen to avoid this approach in order to proceed to higher order in ϵ. An important consideration at this point concerns the possibility of resonant triad interaction. If non-trivial solutions existed to the triad equation

$$
\omega_0(k + \kappa) = \omega_0(k) + \omega_0(\kappa) \qquad k, \kappa > 0, \tag{21}
$$

then the primary instability to the uniform wavetrain (14) would occur at $O(\epsilon^2)$ and would involve the growth of the resonant waves with wavenumbers κ and $k + \kappa$. In the case of (20), however, the triad equation may be reduced to the polynomial equation

$$
\kappa^2 + k\kappa + k^2 + \frac{3}{L_R^2} = 0, \tag{22}
$$

which clearly has no real roots for any real values of k or L_R. This implies that the leading order instability of the wave-train is the Benjamin-Feir instability to be described below.

At first order (17) can be interpreted as a forced equation determining η_1. The forcing terms are all fundamental terms (i.e. proportional to $\exp\left(ikx - i\omega_0 t\right)$) implying that η_1 is also a fundamental term. If η_1 is a fundamental term, however, it can simply be incorporated into η_0 by a renormalisation of $A(X, T, \tilde{T})$. For this reason we may set $\eta_1 \equiv 0$ without loss of generality. From the expression (I.1) given in the appendix we can therefore write, (as (I.2)$\equiv 0$)

$$
v_1 = \frac{\Delta L_R}{4k(1 + k^2 L_R^2)^{3/2}} \left(A_X(X, T, \tilde{T}) \exp\left(ikx - i\omega_0 t\right) + c.c. \right) \tag{23}
$$

and in (17) this results in the group velocity condition

$$A_T + \omega_{0k} A_X = 0, \qquad \omega_{0k} = U_0 - \frac{\Delta L_R}{2(1 + k^2 L_R^2)^{3/2}} \tag{24}$$

Note that $\omega_{0k} = \partial \omega_0 / \partial k$ is also the group velocity derived from the linear dispersion relation (20). The group velocity condition on the wave amplitude A can be expressed as

$$A = A(\zeta, \tilde{T}) \qquad \text{with} \quad \zeta = X - \omega_{0k} T. \tag{25}$$

At second order (equation 18) it is clear from (1) and (1) that both fundamental and third harmonic terms (proportional to $\exp{(3ikx - 3i\omega_0 t)}$) are present. Following the arguments above, η_2 is assumed to have a third harmonic component only, as any fundamental component can be renormalised into A. We can therefore write

$$\eta_2 = \frac{A_H}{6k} \exp{(3ikx - 3i\omega_0 t)} + c.c. \tag{26}$$

where A_H is an as yet unknown amplitude. This allows us to write down an expression for v_2 from (I.1), (I.3) and (I.4)

$$
\begin{aligned}
v_2 \;=\; & \frac{3\Delta L_R^2}{8k(1 + k^2 L_R^2)^{5/2}} \left(i A_{XX}(X, T, \tilde{T}) \exp{(ikx - i\omega_0 t)} + c.c. \right) \\
& + \frac{\Delta L_R}{16 k^2 L_R^2} \left[\left((1 + k^2 L_R^2)^{1/2} - (1 + 4k^2 L_R^2)^{1/2} \right) \left(2i|A|^2 A \exp{(ikx - i\omega_0 t)} + c.c. \right) \right. \\
& + \left. \left((1 + k^2 L_R^2)^{1/2} - 2(1 + 4k^2 L_R^2)^{1/2} + (1 + 9k^2 L_R^2)^{1/2} \right) \left(i A^3 \exp{(3(ikx - i\omega_0 t))} + c.c. \right) \right] \\
& + A_H \frac{\Delta L_R}{4(1 + 9k^2 L_R^2)^{1/2}} \left(i \exp{(3(ikx - i\omega_0 t))} + c.c. \right)
\end{aligned}
\tag{27}
$$

Equating the fundamental terms in equation (18) results in the nonlinear Schrödinger (NLS) equation

$$\text{(NLS)} \qquad\qquad A_{\tilde{T}} - \tfrac{1}{2} i \omega_{0kk} A_{\zeta\zeta} + i\omega_2 |A|^2 A = 0, \tag{28}$$

$$\text{with} \qquad\qquad \omega_{0kk} = \tfrac{3}{2} \Delta L_R^2 \, k L_R \left(1 + k^2 L_R^2 \right)^{-5/2}, \tag{29}$$

$$\text{and} \qquad \omega_2 = -\tfrac{\Delta}{16 k L_R} \left[4(1 + k^2 L_R^2)^{1/2} - (1 + 4k^2 L_R^2)^{1/2} - 3 \right]. \tag{30}$$

Note that the well-known conservation property of the NLS equation

$$\frac{d}{dt} \int |A|^2 \, d\zeta = 0 \tag{31}$$

is the weakly nonlinear equivalent of the fully nonlinear Arnol'd constraint derived by Dritschel (1988) (equation 8 above).

To determine the details of the Benjamin-Feir instability, it is necessary to examine the stability properties of the uniform wavetrain solution to the amplitude equation (28), which has equation

$$A = \exp{(-i\omega_2 \tilde{T})}. \tag{32}$$

As discussed in the introduction the Benjamin-Feir instability is present if and only if the Whitham criterion (1) is satisfied ($\omega_{0kk}\,\omega_2 < 0$). For equivalent

barotropic Rossby waves on a PV front, equations (29-30) indicate that the Whitham criterion is satisfied for all $kL_R > 0$. The standard approach (e.g. Johnson, 1997) is to examine sideband modulations to the uniform wavetrain of the form

$$A(\zeta, \tilde{T}) = \exp\left(-i\omega_2 \tilde{T}\right) \left[1 + A_1 e^{i(q\zeta - \sigma\tilde{T})} + A_2 e^{-i(q\zeta - \sigma^*\tilde{T})}\right] \qquad (33)$$

where the star denotes the complex conjugate. Since the Whitham criterion (1) is satisfied the dispersion relation for the sidebands reveals a positive growth rate with

$$\text{Im } \sigma = -\omega_2 \left(-q^2 \left(\frac{\omega_{0kk}}{\omega_2}\right) - \frac{q^4}{4} \left(\frac{\omega_{0kk}}{\omega_2}\right)^2\right)^{1/2}. \qquad (34)$$

The fastest growing sideband mode has growth rate $\epsilon^2 \text{Im } \sigma_F$ and wavenumber ϵq_F with

$$\text{Im } \sigma_F = -\omega_2 \quad \text{and} \quad q_F = \sqrt{2}\left(-\frac{\omega_2}{\omega_{0kk}}\right)^{1/2}. \qquad (35)$$

In the short wave limit $kL_R \to \infty$, we have $\omega_2 \to -\Delta/8$, maximising Im σ_F and indicating that sideband growth will be fastest for short barotropic waves in an infinite domain. However, in this limit the wavenumber of the fastest growing sideband $q_F \to \sqrt{8/3}L_R^{-1}$, indicating a further separation of scales between the scale of the sideband wavenumber and the fundamental wave (with wavenumber k). This separation of scales means that in a finite domain sideband instability will always be suppressed for large enough L_R, as will be discussed further in section 3 below.

By equating third harmonic terms in (18) it is possible to determine the magnitude and sign of the third harmonic η_2. The constant A_H in (26) can be evaluated after some algebra, as

$$A_H = A^3(\zeta, \tilde{T}) \frac{1}{8k^2 L_R^2} \frac{(1 + 9k^2 L_R^2)^{1/2} - 3(1 + 4k^2 L_R^2)^{1/2} + 3(1 + k^2 L_R^2)^{1/2} - 1}{(1 + k^2 L_R^2)^{-1/2} - (1 + 9k^2 L_R^2)^{-1/2}}$$

$$(36)$$

Figure 2a plots the relative amplitude of the third harmonic compared to the fundamental wave ($|A_H| / 3|A|$) as a function of kL_R. In Figure 2b the fundamental wave (dashed curve) and fundamental plus third harmonic (solid curve) are both plotted for $kL_R = 1$ and $\epsilon|A| = 1$, to illustrate how small the predicted shape correction is. Note that the sign of the shape correction is such that all weakly nonlinear waves are 'omega'-shaped, in the terms of Nakamura and Plumb (1994), as opposed to 'sawtooth'-shaped. At finite amplitude (corresponding to $\epsilon > 1$), Nakamura and Plumb found that short waves ($kL_R > 5.25$) were sawtooth-shaped, but longer waves with $kL_R < 5.25$ were omega-shaped, and that this significantly affects the amplitude at which wave breaking takes place. The abrupt bifurcation at $kL_R \approx 5.25$ between the two types of steady solution has no weakly nonlinear equivalent. These issues will be discussed further below.

3.   NUMERICAL INTEGRATIONS OF THE SINGLE CONTOUR MODEL

In order to verify the weakly nonlinear theory described above, and to investigate its range of validity we have integrated (4) together with (5) numerically

using a single-layer contour dynamics package, see e.g. Dritschel (1989) and Dunn et al. (2001) for details. In summary, the contour is represented by a series of nodes with a typical density of 40-100 nodes per wavelength in the cases to be presented. Velocities at each node are calculated by summing the contributions from every other node along the contour, via an expression derived from (9) using integration by parts (Dritschel, 1986). Periodicity in the x-direction is imposed, with channel length L_x, by repeating the velocity summation over the contour with x-displacements changed by nL_x for $n = 1, -1, 2, -2, ...$ etc. until the remaining node contributions become negligible. Timestepping is done by the fourth-order Runge-Kutta scheme with stepsize $\delta t = 0.005\ \Delta^{-1}$. Normally the contour dynamics algorithm includes a renoding procedure after each timestep, in order to increase the density of nodes in regions of high curvature. However, we found that the renoding procedure had a weak dissipative effect on the waves when an integration continued for a long time (in particular the $\epsilon = 0.3$ case described below). As we are primarily interested in the linear sideband growth stages of the wave, when the contour curvature remains approximately constant, we have turned off the renoding algorithm for integrations with $\epsilon \leqslant 0.7$, and retained it for larger values of ϵ.

Following Swanson et al. (1997), the single contour divides two regions of constant PV in a domain that is periodic in the x-direction and infinite in the y-direction. The contour represents a geophysical PV 'barrier' such as the extratropical tropopause (e.g. Haynes et al., 2001). The size of the domain in the x-direction is taken to be $L_x = 12\pi k^{-1}$, so that there are six waves across the periodic domain. In most of the experiments described below the Rossby radius L_R is set so that $kL_R = 1$.

Swanson (2001) describes some of the issues connected to choosing appropriate scalings for Δ and L_R that best correspond to the situation at the extratropical tropopause, and discusses some of the caveats that must be considered when comparing such an idealised model to the atmosphere. For the case of baroclinic Rossby waves in the extratropics it is important to remember that the single layer contour model cannot describe their vertical structure. Crudely, their vertical structure can be described by a pair of phase-locked 'counter-propagating' Rossby waves at the tropopause and at the surface, that interact constructively to cause wave growth (e.g. Holton, 1992). The single layer contour model corresponds dynamically to the case where the surface wave is absent (or weak), as would occur if the both the surface potential temperature and the latitudinal gradient in lower tropospheric PV were homogenised. The success of the single contour model in describing the basic dispersive properties of baroclinic Rossby waves (see below) might be explained by the relative weakness of the surface wave in fully resolved turbulent baroclinic flows such as the extratropical atmosphere.

Swanson argues that sensible choices are $L_R = 700$ km and $\Delta = 2 \times 10^{-4} \mathrm{s}^{-1}$. This choice of L_R corresponds to a wavelength of 4400 km for waves with $kL_R = 1$, and $L_x = 26400$ km in agreement with observations of extratropical baroclinic waves (e.g. Pierrehumbert and Swanson, 1995). The basic velocity along the contour $U_0 = \Delta L_R/2 = 70 \mathrm{ms}^{-1}$, wave phase speed $\omega_0/k = 20.5\ \mathrm{ms}^{-1}$ and wave group velocity $\omega_{0k} = 45.3\ \mathrm{ms}^{-1}$ for waves with wavenumber $kL_R = 1$. Compared with Rossby waves at the tropopause these values are roughly in the correct ratio but are somewhat high (suggesting that a lower value of $\Delta = 1 \times 10^{-4} \mathrm{s}^{-1}$ might be more appropriate). Chang (1999) estimates for the Southern Hemisphere summer $U_0 \approx 30 \mathrm{ms}^{-1}$, $\omega_0/k = 8$-$12\ \mathrm{ms}^{-1}$ and $\omega_{0k} = 22\ \mathrm{ms}^{-1}$.

For mesoscale, oceanic baroclinic Rossby waves propagating on the Antarctic Circumpolar Current, suitable scales might be $U_0 = 0.05 - 0.1$ ms^{-1}, $L_R = 36$km (Sinha and Richards, 1999).

We test the following two predictions of the weakly nonlinear theory.

(1) *The frequency correction for finite amplitude wavetrains:*
For the uniform wavetrain solution of the NLS equation given above, the frequency correction is given by

$$\omega = \omega_0 + \epsilon^2 \omega_2 \qquad (37)$$

where ω_0 and ω_2 are as in equations (20) and (30) respectively.

(2) *The growth rate of sideband disturbances to the uniform wavetrain:*
In an infinite domain, from equation (35) the fastest growing sideband mode has growth rate $\epsilon^2 \omega_2$. However, in a periodic domain with dimension L_x the allowable sideband wavenumbers satisfy $\epsilon q = 2\pi n L_x$, where $n = 1, 2, 3, ...$, and the growth rate is given by equation (34).

Unless otherwise stated in the following experiments the contour dynamics model is initialised with a sinusoidal wavetrain (equation 14 with $A \equiv 1$), plus a small random perturbation. In these cases, because no attempt is made to correct the wave shape an oscillation similar to that described by Nakamura and Plumb (1994) occurs, where the wave oscillates between having a sinusoidal and an 'omega'-shape. Several experiments where the model was instead initialised with $\eta_0 + \epsilon^2 \eta_2$ were found to reduce the amplitude of this oscillation by a factor of 10-50 for $\epsilon = 0.3 - 0.5$. Using this more 'accurate' initialisation had no effect on the calculated frequency of the fundamental or the sideband growth rates.

The growth rates and frequencies shown in Figures 3 and 4 are obtained by careful analysis of the evolution of the wave spectrum, which is obtained at each time-step by Fourier decomposition of the contour position. The frequency of the fundamental wave may then be calculated from the evolution of the wave-6 component. Linear sideband growth is diagnosed as taking place during the early stages of the evolution of the wavetrain, when sidebands wavenumbers grow exponentially in amplitude at a constant measurable rate as predicted by equation (33). For example, during wave-1 sideband growth, wave-5 and wave-7 grow exponentially with the same growth rate (typically measured to agree to within 1%). Wave-2 sideband growth involves the exponential growth of wave-4 and wave-8, and wave-3 sideband growth the exponential growth of wave-3 and wave-9. Well-defined periods of sideband growth were detectable in all of the experiments presented.

Figure 3 shows plots of predicted $(\omega_0 + \epsilon^2 \omega_2)$ and calculated frequency against non-dimensional wave amplitude ϵ for uniform wavetrains of wavenumber $kL_R = 1$ (left) and $kL_R = 2$ (right). This shows good agreement between the numerical results (triangles) and the weakly nonlinear predictions (curves) up to $\epsilon = 0.7$.

Figure 4 shows predicted and calculated sideband growth rates for the $kL_R = 1$ experiments. The dotted curve illustrates the predicted maximum sideband growth rate Im σ_F in the case where the channel is infinite (see equation 35). The solid curve shows the predicted growth rate of the wave-1 sideband (with $\epsilon q = 2\pi/L_x$) and the dashed curve the wave-2 sideband (with $\epsilon q = 4\pi/L_x$). Note that the wave-1 sideband first becomes unstable only at $\epsilon \approx 0.264$, as the wavelength of the fastest growing sideband, which scales with ϵ^{-1}, is longer than the channel at lower wave amplitudes. To confirm convergence between theory and model

we therefore made one experiment with double the channel length $L_x = 24\pi/k$ (12 waves in the channel). In this case the weakly nonlinear predicted growth rate (long-dashed line in Figure 4) and calculated growth rate (cross) differed by less than 1%. The weakly nonlinear predictions for wave-1 sideband growth rates are reasonable for $\epsilon = 0.3, 0.4$, but at higher values of ϵ they tend to overestimate the growth rates. In addition the amplitude of transition from the wave-1 sideband disturbance to the wave-2 sideband disturbance is predicted to take place at $\epsilon \approx 0.592$, but in fact occurs between $\epsilon = 0.6$ and $\epsilon = 0.7$. Dimensionally, using the atmospheric scalings mentioned above, the e-folding timescale for sideband growth at $\epsilon = 0.3$ is 30 days and at $\epsilon = 0.7$ is 7.5 days.

Figure 5 shows snapshots of the growing sideband disturbances for $\epsilon = 0.3$ (left) and $\epsilon = 0.6$ (right). A wave-1 sideband grows in the $\epsilon = 0.3$ case and the wavetrain evolves into a single wave packet. Although this wave packet resembles an envelope soliton at time $t = 800\pi\Delta^{-1}$, this is far from an invariant solution and the envelope shape oscillates strongly until the end of this integration at $t = 2000\pi\Delta^{-1}$. For the $\epsilon = 0.6$ case, the wave-1 unstable sideband also comes to dominate and the wavetrain again breaks up into a single wave packet by $t = 200\pi\Delta^{-1}$. After this time, the largest wave in the wavetrain breaks, shedding a filament as has been previously described by e.g. Polvani and Plumb (1992). (This integration was repeated with renoding in order to validate the evolution when high curvature develops in the breaking region.)

At higher amplitudes the weakly nonlinear theory cannot be used to predict sideband growth rates. However a further series of experiments with $\epsilon = 0.7 - 2.0$ each showed clear evidence of linear sideband growth of small perturbations to the uniform wavetrain, of the form described by equation (33). Wave-2 sideband growth is dominant for $1.2 \geqslant \epsilon \geqslant 0.7$ and a wave-3 sideband growth is found to be dominant for $\epsilon \geqslant 1.5$. Sideband growth rates were also found to peak at around $\epsilon = 1.5$, when the dimensional e-folding timescale is around 3.8 days, a realistic timescale for wave packet formation in the southern hemisphere. Figure 6 shows the evolution of the wavetrain for $\epsilon = 1.2$ and $\epsilon = 2.0$. In the $\epsilon = 1.2$ case the wavetrain evolves into two wave packets, and in one of these packets the largest wave eventually breaks, shedding a filament. For $\epsilon = 2.0$, wave-3 sideband growth dominates, and this sideband modulates the wave-6 wavetrain (reference to 'wave packets' in this case might be misleading as there is little separation of scales between the wave envelope and the waves). In this case the eventual wave breaking has a different character, with small vortices being shed from the largest waves.

Finally, we compared some experiments at higher amplitudes with those of Nakamura and Plumb (1994) who looked at the short time evolution of uniform wavetrains in a near identical model in order to investigate wave breaking. They found that for $kL_R > 5.2$ the required amplitude for wave breaking of the uniform wavetrain was in the range $\epsilon \approx 2 - 2.5$ (see their Figure 10, and note that we have converted their quantities into the notation used in this paper) whereas in the range $2 < kL_R < 5.2$ the minimum breaking amplitude was higher at $\epsilon \approx 4.5\text{-}5.5$. Nakamura and Plumb describe how the jump in breaking amplitude at $kL_R \approx 5.2$ corresponds to a transition in the equilibrium shape of the waves between 'omega'-shaped and 'sawtooth'-shaped. Note that we have shown that weakly nonlinear waves are all 'omega'-shaped to leading order, so this shape transition is strongly nonlinear in that it takes place only at $O(1)$ wave amplitudes. Our results (not shown) confirm that at $kL_R = 1$ wave breaking takes place for $\epsilon = 5.0$. The

character of the wave breaking in this unperturbed wavetrain is quite different to that following sideband growth. Instead of shedding filaments or small vortices the contour separates into a train of opposite signed vortices. A further experiment with $kL_R = 10$ confirmed the results of Nakamura and Plumb for short waves. In this case the wave breaking mechanism for the uniform wavetrain with $\epsilon = 2.5$ involves 'sawtooth'-shaped waves that shed filaments.

Our experiments have shown that Benjamin-Feir instability leads to Rossby wave breaking for contours with similar initial non-dimensional wave amplitudes to those observed for tropopause Rossby waves. The transition that takes place in our experiments, between low amplitude waves that eventually break to shed filaments ($\epsilon = 0.6$ case in Figure 5) and higher amplitude waves that break forming cut-off vortices ($\epsilon = 2$ case in Figure 6), may be compared to different types of observed wave breaking at the tropopause, as exemplified by the paradigms LC1 and LC2 described in detail by Thorncroft et al., (1993). In a previous contour dynamics-based modelling study of wave breaking at the tropopause, Peters and Waugh (1996) found that cut-off vortices typically formed only in regions of diffluent background flow. Here, by allowing the Benjamin-Feir instability to develop, we have shown that this type of wave breaking can develop from an initially uniform background flow. Cut-off vortices are known to form due to wave breaking near the tropopause in non-diffluent regions, so it is reasonable to conclude that the presence of wave packets due to Benjamin-Feir instability may be necessary for their development.

4. CONCLUSIONS

In this paper we have used a single contour model of a potential vorticity front to show that Benjamin-Feir instability is a viable mechanism for wave packet formation on the extratropical tropopause. The instability was found to develop on timescales of a few days, leading to the formation of two wave packets encircling the globe as seen in observations (c.f. Figure 1), when a uniform wave-train with realistic wave amplitude ($\epsilon \sim 1$) was used as an initial condition. As in the well-known water wave example (Hasimoto and Ono, 1972) as well as for a class of Rossby waves in a β-channel discussed by Plumb (1977), the uniform wavetrain was found to be primarily unstable to sideband disturbances that grow exponentially and cause the development of nonlinear wave packets. In the weakly nonlinear limit the wave amplitude evolves according to the nonlinear Schrödinger (NLS) equation. Wave growth due to baroclinic instability and decay due to dissipation are therefore not necessary ingredients for wave packet formation, although both effects are present in the two-layer model study of Esler (1997). In the latter study the wave amplitude is governed by the Ginzburg-Landau equation, the complex nonlinear coefficient of which has strong dependence on the damping coefficients. Direct comparison with the contour model is difficult as the model assumptions render the equation invalid in the limit of weak damping.

The weakly nonlinear results presented here also indicate that equivalent barotropic Rossby wavetrains are unstable for all zonal wavenumbers $kL_R > 0$ and that the sideband growth rate is fastest for short waves ($kL_R \to \infty$). For short waves, however, the most unstable sideband scales with L_R, so a uniform wavetrain would need to contain many wavelengths in order that the instability be expressed. For long waves ($kL_R \to 0$) the grow rate of the instability tends to zero.

In a finite periodic domain such as the extratropical atmosphere, the Benjamin-Feir instability is therefore strongest for waves with a scale comparable to the Rossby radius, $kL_R \sim 1$, i.e. observed synoptic-scale waves. Another notable result is that sideband growth leads to the breaking of Rossby waves with much lower initial amplitudes compared with initially uniform wavetrains, and this wave breaking is spatially localised. If the initial wave amplitude is sufficiently large ($\epsilon \geqslant 1.2$), then this wave breaking leads to the formation of small cut-off vortices. Cut-off vortex formation is a ubiquitous process at the tropopause (e.g. Thorncroft et al., 1993) that has not previously been noted to develop in contour dynamics integrations except where initial conditions are significantly asymmetric.

In the oceans PV barriers or fronts supporting Rossby waves are known to exist in the Antarctic circumpolar current and the Atlantic gulf stream (see Rhines 1994, for a discussion of these and other geophysical flows) and it is possible that careful analysis of observations in these regions may reveal Rossby wave packet formation. Similarly, the banded jets on the giant planets are likely associated with a series of PV fronts, each supporting Rossby waves.

APPENDIX A

Evaluation of the Velocity Field from the Contour Displacement

In this appendix we show how the velocity fields $(u(x), v(x))$ along the contour are obtained up to the third order in ϵ, given the contour position function η, using the integral expression (12). Concentrating first on the meridional velocity v we have

$$
\begin{aligned}
v(x) \quad = \quad & \epsilon \frac{\Delta}{2\pi} \int_{-\infty}^{\infty} K_0 \left(|x - x'|/L_R \right) \eta_0'(x') \, dx' && \text{(I.1)} \\
+ \quad & \epsilon^2 \frac{\Delta}{2\pi} \int_{-\infty}^{\infty} K_0 \left(|x - x'|/L_R \right) \eta_1'(x') \, dx' && \text{(I.2)} \\
+ \quad & \epsilon^3 \frac{\Delta}{2\pi} \int_{-\infty}^{\infty} K_0 \left(|x - x'|/L_R \right) \eta_2'(x') \, dx' && \text{(I.3)} \\
+ \quad & \epsilon^3 \frac{\Delta}{2\pi} \int_{-\infty}^{\infty} \frac{K_0'(|x-x'|/L_R)}{2|x'-x|} \eta_0'(x') \left(\eta_0(x') - \eta_0(x) \right)^2 \, dx' && \text{(I.4)} \\
+ \quad & O(\epsilon^4)
\end{aligned}
$$

In order to evaluate the first part of this integral (I.1) we note that

$$
\begin{aligned}
\eta_0(x') \quad = \quad & \frac{1}{2k} A(X', T, \tilde{T}) e^{ikx' - i\omega_0 t} + c.c. \\
= \quad & \frac{1}{2k} \Big(A(X, T, \tilde{T}) + \epsilon(x' - x) A_X(X, T, \tilde{T}) \\
& \qquad\qquad + \frac{\epsilon^2}{2}(x' - x)^2 A_{XX}(X, T, \tilde{T}) + O(\epsilon^3) \Big) e^{ikx' - i\omega_0 t} + c.c.
\end{aligned}
$$

giving

$$
\begin{aligned}
\eta_0'(x') \quad = \quad & \frac{1}{2k} \Big(ikA(X, T, \tilde{T}) + \epsilon ik(x' - x) A_X(X, T, \tilde{T}) + \frac{\epsilon^2}{2} ik(x' - x)^2 A_{XX}(X, T, \tilde{T}) \\
+ \quad & \epsilon A_X(X, T, \tilde{T}) + \epsilon^2(x' - x) A_{XX}(X', T, \tilde{T}) + O(\epsilon^3) \Big) e^{ikx' - i\omega_0 t} + c.c.
\end{aligned}
$$

Under a change of variables $\tilde{x} = x' - x$ and using the integral identities listed in Appendix B, (I.1) can be straightforwardly evaluated giving

$$(\text{I.1}) \quad = \quad \epsilon \frac{\Delta L_R}{4(1 + k^2 L_R^2)^{1/2}} \left(iA(X, T, \tilde{T})e^{ikx - i\omega_0 t} + c.c. \right)$$

$$+ \quad \epsilon^2 \frac{\Delta L_R}{4k(1 + k^2 L_R^2)^{3/2}} \left(A_X(X, T, \tilde{T})e^{ikx - i\omega_0 t} + c.c. \right)$$

$$+ \quad \epsilon^3 \frac{3\Delta L_R^2}{8k(1 + k^2 L_R^2)^{5/2}} \left(iA_{XX}(X, T, \tilde{T})e^{ikx - i\omega_0 t} + c.c. \right)$$

The leading order term yields the expression for v_0 (19) given in section 2, and the linear dispersion relation is recovered. From the renormalisation arguments discussed in section 2, it turns out that $\eta_1 \equiv 0$ and hence (I.2) $\equiv 0$. The next order meridional velocity v_1 is therefore given by (23) and the group velocity condition (24) is obtained. At next order, however, the nonlinear integral given by expression (I.4) must be evaluated. To leading order we find that

$$\eta_0'(x') \left(\eta_0(x') - \eta_0(x) \right)^2 \quad = \quad \frac{1}{8k^2} \left(i|A|^2 A e^{ikx - i\omega_0 t} \left(3e^{ik\tilde{x}} - e^{-ik\tilde{x}} - 2e^{2ik\tilde{x}} \right) + c.c. \right)$$

$$+ \quad \left(iA^3 e^{3(ikx - i\omega_0 t)} \left(e^{ik\tilde{x}} - 2e^{2ik\tilde{x}} + e^{3ik\tilde{x}} \right) + c.c. \right)$$

where to leading order we can take $A = A(X, T, \tilde{T})$ in each case. Using the appropriate Bessel function identities listed in Appendix B this expression allows (I.4) to be evaluated, giving,

$$(\text{I.4}) \quad = \quad \epsilon^3 \frac{\Delta L_R}{4k^2 L_R^2} \left[2i|A|^2 A \left((1 + k^2 L_R^2)^{1/2} - (1 + 4k^2 L_R^2)^{1/2} \right) e^{ikx - i\omega_0 t} \right.$$

$$+ \quad iA^3 e^{3(ikx - i\omega_0 t)} \left. \left((1 + k^2 L_R^2)^{1/2} - 2(1 + 4k^2 L_R^2)^{1/2} + (1 + 9k^2 L_R^2)^{1/2} \right) + c.c. \right]$$

From (12) the integral expression for the $O(\epsilon^2)$ correction to the zonal velocity u_1 takes the form

$$u_1 = \frac{\Delta}{2\pi} \int_{-\infty}^{\infty} \frac{K_0'(|x - x'|/L_R)}{2L_R|x' - x|} \left(\eta_0(x') - \eta_0(x) \right)^2 dx'.$$

This can be evaluated as above giving

$$u_1 \quad = \quad -\frac{\Delta L_R}{16k^2 L_R^2} \left[4|A|^2 \left((1 + k^2 L_R^2)^{1/2} - 1 \right) \right.$$

$$+ \left(Ae^{2(ikx - i\omega_0 t)} + c.c. \right) \left. \left(2(1 + k^2 L_R^2)^{1/2} - (1 + 4k^2 L_R^2)^{1/2} - 1 \right) \right].$$

which yields the following expression needed for equation (18)

$$u_1 \eta_{0x} = -\frac{\Delta L_R}{32k^2 L_R^2} \left[i|A|^2 A \left(2(1 + k^2 L_R^2)^{1/2} + (1 + 4k^2 L_R^2)^{1/2} - 3 \right) e^{ikx - i\omega_0 t} \right.$$

$$+ iA^3 e^{3(ikx - i\omega_0 t)} \left. \left(2(1 + k^2 L_R^2)^{1/2} - (1 + 4k^2 L_R^2)^{1/2} - 1 \right) + c.c. \right].$$

APPENDIX B

Integral identities used in Appendix A

The following identities are valid for all real m, including $m = 0$, and are used to evaluate the integral expressions in Appendix A (e.g. http://functions.wolfram.com/).

$$\int_{-\infty}^{\infty} K_0(|x|)\, e^{imx} dx = \pi(1 + m^2)^{-1/2}$$

$$\int_{-\infty}^{\infty} K_0(|x|)\, x\, e^{imx} dx = i\pi m(1 + m^2)^{-3/2}$$

$$\int_{-\infty}^{\infty} K_0(|x|)\, x^2 e^{imx} dx = \pi(1 - 2m^2)(1 + m^2)^{-5/2}$$

$$\int_{-\infty}^{\infty} \frac{K_1(|x|)}{|x|} e^{imx} dx = -\pi(1 + m^2)^{1/2}$$

Here $K_n(x)$ is the modified Bessel function that is uniquely defined as the solution of the equation

$$r^2 \frac{d^2\psi}{dr^2} + r\frac{d\psi}{dr} - (n^2 + r^2)\psi = 0, \qquad (B.1)$$

satisfying $K_n(x) \to 0$ as $x \to \infty$. Note that K_0 and K_1 are related through the identity

$$K_0'(x) = -K_1(x).$$

REFERENCES

Arnol'd, V. I.	1969	On an a priori estimate in the theory of hydrodynamical stability, *Izv. Vyssy. Uchebn. Zaved. Matematika*, **54**, 3-5, 1966, (English Transl. *Am. Math. Soc. Trans. Series 2*, **79**, 267-269.)
Benjamin, T. B. and Feir, J. E.	1967	The disintegration of wave trains on deep water. *J. Fluid Mech.*, **27,** 417–430
Benjamin, T. B. and Bridges, T. J.	1997	Reappraisal of the Kelvin-Helmholtz problem. Part 2. Interaction of the Kelvin-Helmholtz, superharmonic and Benjamin-Feir instabilities. *J. Fluid Mech.*, **333**, 327–373
Berbery, E. H. and Vera, C. S.	1996	Characteristics of the Southern Hemisphere winter storm track with filtered and unfiltered data. *J. Atmos. Sci.*, **53**, 468–481
Boyd, J. P.	1983	Equatorial solitary waves. Part 2: Envelope solitons. *J. Phys. Oceanogr.*, **13**, 428–449
Bridges, T. J. and Mielke, A.	1995	A proof of the Benjamin-Feir instability. *Arch. Rat. Mech. Anal.*, **133**, 145–198
Chang, E. M.	1993	Downstream development of baroclinic waves as inferred from regression analysis. *J. Atmos. Sci.*, **50,** 2038–2053
Chang, E. M.	1999	Characteristics of Wave Packets in the Upper Troposphere. Part II: Seasonal and Hemispheric Variations. *J. Atmos. Sci.*, **56**, 1729–1747
Chang, E. M.	2001	The structure of baroclinic wave packets. *J. Atmos. Sci.*, **58,** 1694–1713
Chang, E. M., and Yu, D. B.	1999	Characteristics of wave packets in the upper troposphere. Part I: Northern Hemisphere winter. *J. Atmos. Sci.*, **56**, 1708–1728.
Dritschel, D. G.	1986	The nonlinear evolution of rotating configurations of uniform vorticity. *J. Fluid Mech.*, **172,** 157–182
Dritschel, D. G.	1988	Nonlinear stability bounds for inviscid, two-dimensional, parallel or circular flows with monotonic vorticity, and the analogous three-dimensional quasi-geostrophic flows. *J. Fluid Mech.*, **191,** 575–581
Dritschel, D. G.	1989	Contour dynamics and contour surgery: Numerical algorithms for extended high-resolution modeling of vortex dynamics in two-dimensional inviscid, incompressible flows. *Comput. Phys. Rep.*, **10**, 77–146
Dunn, D. C., Macdonald, N. R. and Johnson, E. R.	2001	The motion of a singular vortex near an escarpment. *J. Fluid Mech.*, **448,** 335–365
Esler, J. G.	1997	Wave packets in simple equilibrated baroclinic systems. *J. Atmos. Sci.*, **54**, 2820–2849
Esler, J. G., and Haynes, P. H.	1999	Mechanisms for wave packet formation and maintenance in a quasigeostrophic two-layer model. *J. Atmos. Sci.*, **56,** 2457–2489
Grimshaw, R. H. J.	1977	The modulation of an internal gravity wave packet and the resonance with the mean motion, *Stud. Appl. Maths*, **56**, 241–266
Grimshaw, R. H. J., Pelinovsky, D., Pelinovsky, E., and Talipova, T.	2001	Wave group dynamics in weakly nonlinear long-wave models. *Physica D*, **159**, 35–57
Hasimoto, H. and Ono, H.	1972	Nonlinear modulation of gravity waves. *J. Phys. Soc. Japan*, **33**, 805–811
Haynes, P. H., Johnson E. R. and Hurst R. G.	1993	A simple model of Rossby-wave hydraulic behaviour. *J. Fluid Mech.*, **253**, 359–384
Haynes, P. H., Scinocca, J. F. and Greenslade, M. D.	2001	Formation and maintenance of the extratropical tropopause by baroclinic eddies. *Geophys. Res. Lett.*, **28,** 4179–4182
Holton, J. R.	1992	*An introduction to dynamic meteorology*, Third Edition, Academic Press. 511 pp
Hoskins, B. J., McIntyre, M. E. and Robertson, A. W.	1985	On the use and significance of isentropic potential vorticity maps. *Q. J. R. Meteorol. Soc.*, **111,** 877–946
Johnson, R. S.	1997	*A modern introduction to the mathematical theory of water waves*, Cambridge University Press, 437 pp

|		
Karsten, R. H. and Swaters, G. E.	2000	Nonlinear effects in two-layer large-amplitude geostrophic dynamics. Part I. The strong-beta case, *J. Fluid Mech.*, **412,** 125–160
Lee, S. and Held, I. M.	1993	Baroclinic wave packets in models and observations. *J. Atmos. Sci.*, **50,** 1413–1428
Nakamura, M. and Plumb, R. A.	1994	The effects of flow asymmetry on the direction of Rossby wave breaking, *J. Atmos. Sci.*, **51,** 2031–2045
Newell, A. C.	1969	Rossby wave packet interactions, *J. Fluid Mech.*, **35,** 255–271
Nycander, J., Dritschel, D. G. and Sutyrin, G. G.	1993	The dynamics of long frontal waves in the shallow-water equations. *Phys. Fluids A*, **5,** 1089–1091.
Peters, D., and Waugh, D. W.	1996	Influence of barotropic shear on the poleward advection of upper tropospheric air. *J. Atmos. Sci.*, **53,** 3013–3031
Pierrehumbert, R. T. and Swanson, K. L.	1995	Baroclinic Instability. *Ann. Rev. Fluid Mech.*, **27,** 419–467
Plumb, R. A.	1977	The stability of small amplitude Rossby waves in a channel. *J. Fluid Mech.*, **80,** 705–720
Polvani, L. M. and Plumb, R. A.	1992	Rossby wave breaking, microbreaking, filamentation and secondary vortex formation: the dynamics of a perturbed vortex. *J. Atmos. Sci.*, **52,** 462–476
Pratt, L. J.	1988	Meandering and eddy detachment according to a simple (looking) path equation. *J. Phys. Oceanogr.*, **18,** 1627–1640.
Pratt, L. J. and Stern, M. E.	1986	Dynamics of Potential Vorticity Fronts and Eddy Detachment. *J. Phys. Oceanogr.*, **16,** 1101–1120
Rhines, P. B.	1994	Jets. *Chaos*, **4,** 313–341
Sinha, B., and Richards, K. J.	1999	Jet structure and scaling in Southern Ocean Models. *J. Phys. Oceanogr.*, **29,** 1143–1155
Slomp, C. G. and Swaters, G. E.	1997	Finite-amplitude perturbations and modulational instability of a stable geostrophic front. *Geophys. Astro. Fluid Dyn.*, **86,** 149–172
Sutherland, B. R.	2001	Finite-amplitude internal wavepacket dispersion and breaking. *J. Fluid Mech.*, **429,** 343–380
Swanson, K. L.	2000	Stationary wave accumulation and the generation of low-frequency variability on zonally varying flows. *J. Atmos. Sci.*, **57,** 2262–2280
Swanson, K. L.	2001	Blocking as a local instability to zonally varying flows. *Q. J. R. Meteorol. Soc.*, **127,** 1341–1355
Swanson, K. L., Kushner, P. J. and Held, I. M.	1997	Dynamics of barotropic storm tracks. *J. Atmos. Sci.*, **54,** 791–810
Thorncroft, C. D., Hoskins, B. J. and McIntyre, M. E.	1993	Two paradigms of baroclinic life cycle behaviour. *Q. J. R. Meteorol. Soc.*, **119,** 17–35
Whitham, G. B.	1974	*Linear and Nonlinear waves.* John Wiley and Sons, 636 pp
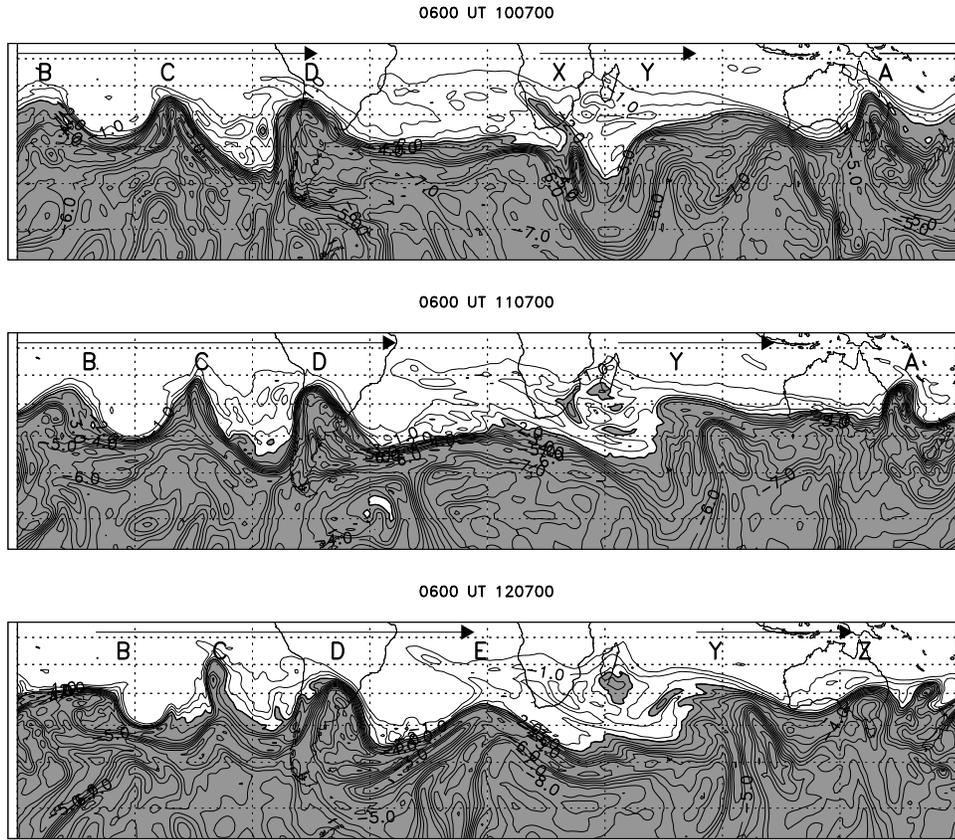
Figure 1. Ertel's Potential Vorticity $Q = (1/\rho)\nabla\theta.\zeta$ where ρ is density, θ is potential temperature and ζ is absolute vorticity on the $\theta = 340$ K isentropic surface on three consecutive days during 10-12 July 2000 (at 0600UT). The PV field is derived from ECMWF operational analysis. The domain shows the southern hemisphere with latitude range -5° to -65°. The contour interval is 0.5 potential vorticity units (PVU, 1 PVU=1×10^{-6} K kg^{-1} m^2 s^{-1}). Stratospheric air, defined as having PV less than -2 PVU, is shaded grey.
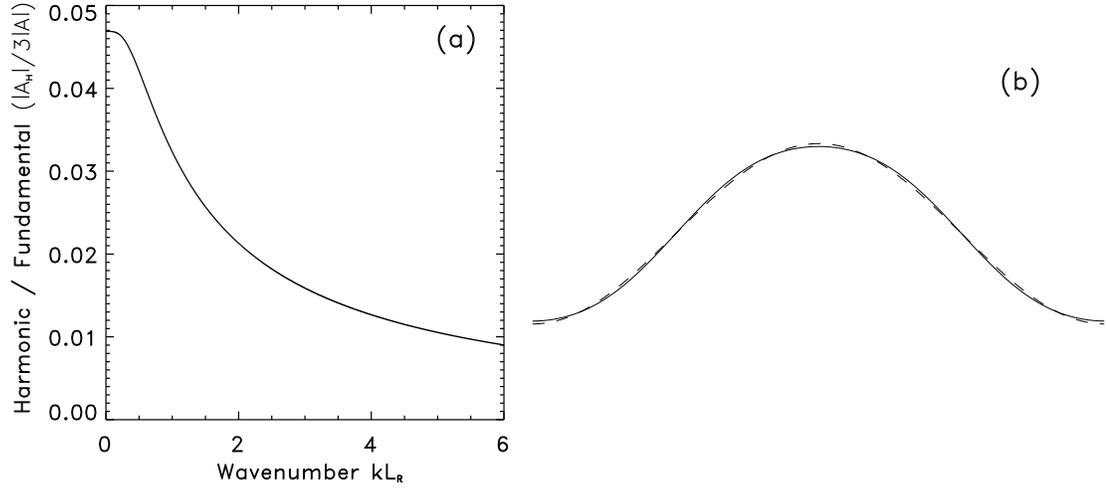
Figure 2. (a) The third harmonic amplitude divided by fundamental amplitude for the weakly nonlinear Rossby waves as a function of wavenumber kL_R. (b) Showing the leading order correction to the shape of the potential vorticity contour (solid line) $\eta_0 + \epsilon^2 \eta_2$ predicted by equations (14), (26) and (36) for $kL_R = 1$ (with $\epsilon|A|=1$). The dashed line shows the sinusoidal wave η_0.
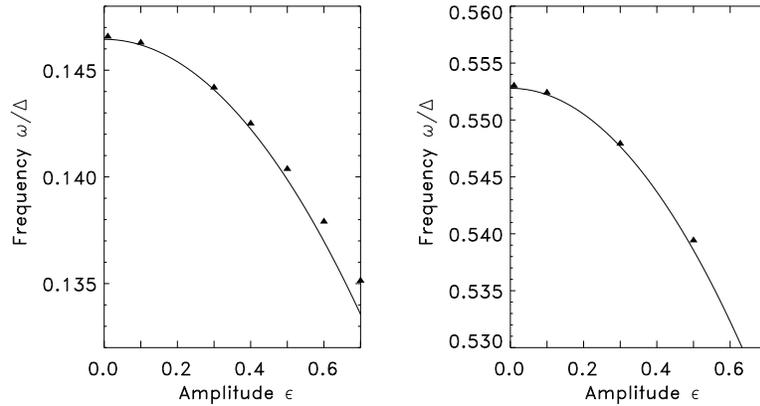

Figure 3. Non-dimensional wave frequency against wave amplitude ϵ. Model calculations (triangles) are compared with weakly nonlinear predictions from (37), (curves). The left panel shows waves with $kL_R = 1$, with numerical experiments for $\epsilon = 0.01, 0.1, 0.3, 0.4, 0.5, 0.6$ and 0.7 and in the right panel $kL_R = 2$, $\epsilon = 0.01, 0.1, 0.3,$ and 0.5.
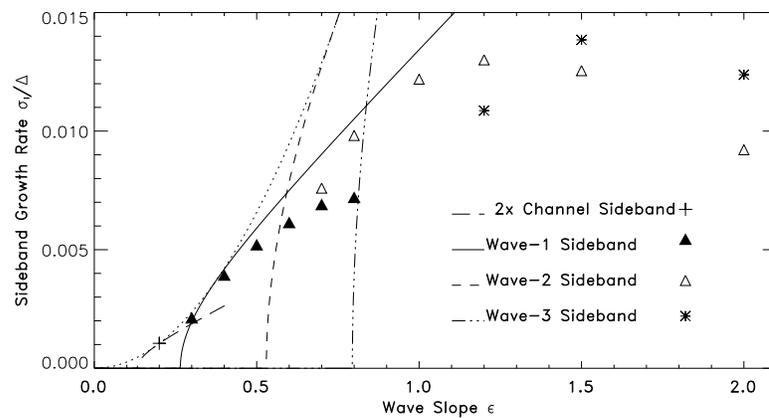
Figure 4.   Sideband growth rate against wave amplitude ϵ. Model calculations (triangles) are compared with weakly nonlinear predictions from (34). In the calculations the initial wavelength $kL_R = 1$ and the domain is periodic with length scale $L_x = 12\pi k^{-1}$. The predicted growth rates of wave-1, wave-2 and wave-3 sideband modulations (see text) are given by the solid, dashed and dot-dashed curves respectively. The curve with long dashes shows the predicted wave-1 sideband growth rates in a domain with double the size, and the dotted curve show the predicted sideband growth rates for an infinite domain. The solid triangles show the wave-1 sideband growth rates for the numerical experiments, and the unfilled triangles shows the wave-2 sideband growth rate. For $\epsilon = 0.7$ and higher the wave-2 sideband growth dominates.
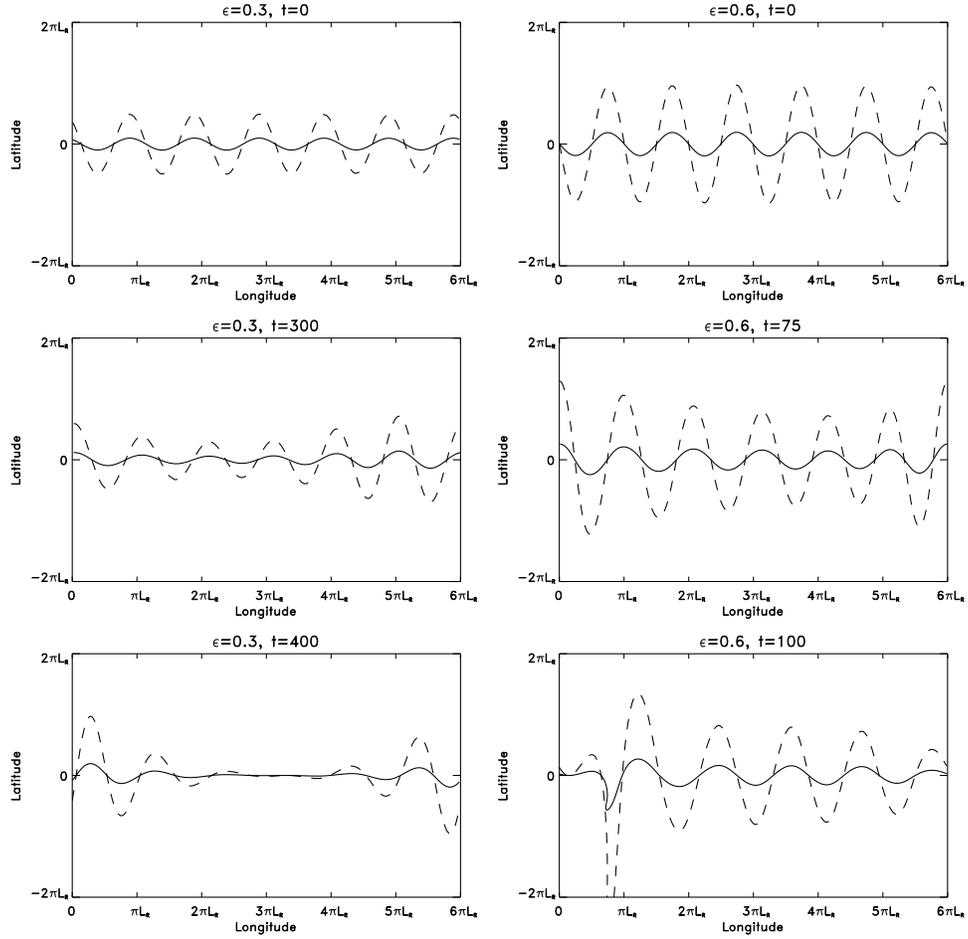
Figure 5.   Snapshots of the contour evolution in the $\epsilon = 0.3$ experiment (left panels) and the $\epsilon = 0.6$ experiment (right panels). In the $\epsilon = 0.3$ case the snapshots are for times $t = 0, 300$ and $400 \times 2\pi\Delta^{-1}$, and for $\epsilon = 0.6$ they are for $t = 0, 75$ and $100 \times 2\pi\Delta^{-1}$. The dashed curves show the same waves plotted with $5\times$ their actual amplitude for clarity.
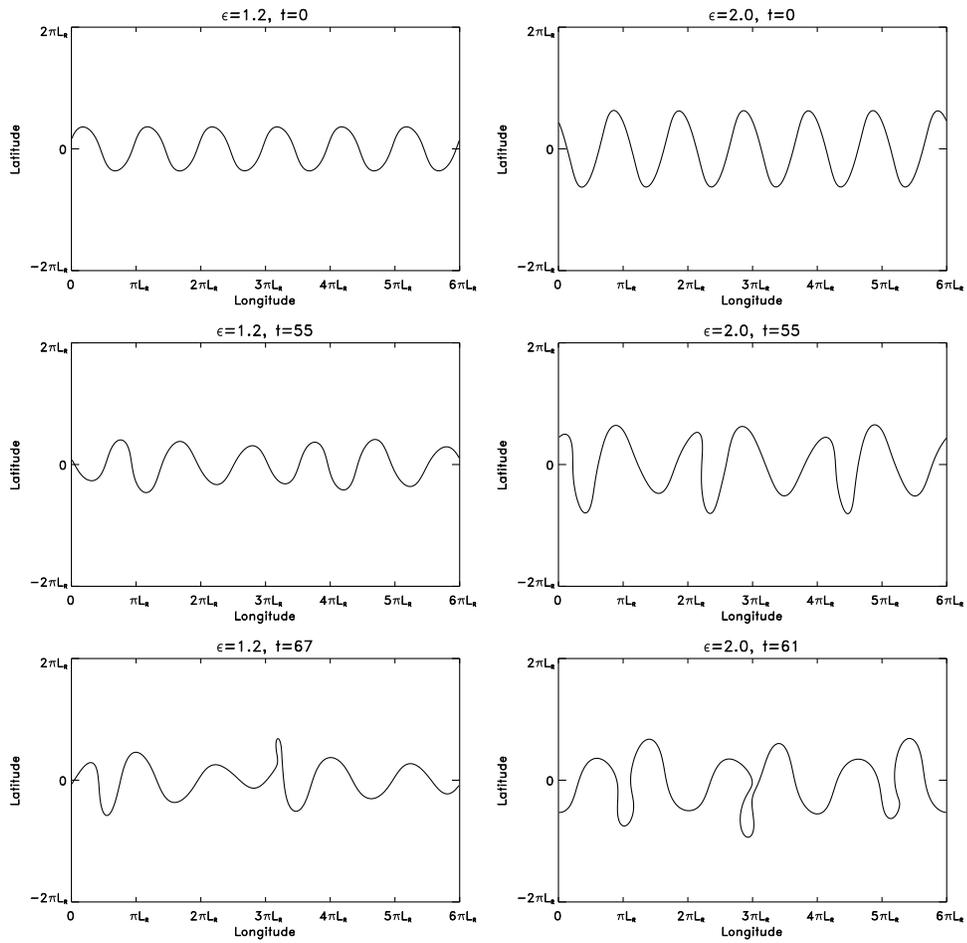
Figure 6.   Snapshots of the contour evolution for $\epsilon = 1.2$ (left) and $\epsilon = 2.0$ (right) experiments. Time units are again $2\pi\Delta^{-1}$.