J. Fluid Mech. (1985), *vol.* 161, *pp.* 493–511
Printed in Great Britain

Nonlinear instability of a Rossby-wave critical layer

By PETER H. HAYNES

Department of Applied Mathematics and Theoretical Physics, University of Cambridge†

(Received 4 April 1985 and in revised form 6 August 1985)

The vorticity distribution in a Rossby-wave nonlinear critical layer, given by the Stewartson–Warn–Warn solution, may be strongly modified by the action of shear instability. In a companion paper (Killworth & McIntyre 1985) it was shown, using linear theory, that unstable modes indeed existed. Here, using numerical methods, the nonlinear evolution of unstable disturbances is followed up to the time at which their growth ceases. By such a time there has been considerable redistribution of vorticity in the critical layer.

1. Introduction

The idea of a critical line is well established in hydrodynamics and frequently arises in problems involving monochromatic wave disturbances to a shear flow. In the neighbourhood of the critical line, where the phase speed of the wave is equal to the basic-flow speed, steady linear theory is inadequate to describe the dynamical balance governing the waves. In this paper we shall be concerned with the Rossby-wave case (where this term is used not just in the geophysical context, but includes any wave which owes its existence to a basic-state vorticity gradient) reviewed by Stewartson (1981). Our understanding of the nonlinear and time-dependent processes that may take place near a Rossby-wave critical line, if the viscosity is small enough, has been considerably enhanced by the numerical work of Béland (1976, 1978) and the analytical theory presented by Stewartson (1978) and Warn & Warn (1978) (hereinafter referred to as SWW).

The problem that SWW considered was one in which small-amplitude monochromatic waves are forced on a zonal shear flow by corrugations in a rigid wall. The corrugations are switched on at some instant and in the subsequent evolution transient effects become confined to a region centred on the critical line and decreasing in thickness with time (Dickinson 1970), while elsewhere the wave pattern settles down to a steady state. However small the wave amplitude, nonlinear effects must eventually become important in the thin region of transient motion surrounding the critical line (Warn & Warn 1976). This region is customarily referred to as the critical layer.

The problem therefore exhibits a two-scale structure; a thin critical layer where the motion is nonlinear and transient and a much broader region where, to a first approximation, the motion consists of quasi-steady waves. This structure was exploited mathematically by SWW, using the method of matched asymptotic expansions. In such a formulation, interaction between the motion in the two regions

† Also J.I.S.A.O. Contribution no. 17, University of Washington, Seattle, Washington 98195.

occurs through the matching conditions, which require that the solution valid in one region smoothly blends into that valid in the other.

The details of the SWW analysis are reviewed in the companion paper to this, Killworth & McIntyre (1985), hereinafter referred to as KM. One particularly interesting result emerges when the forced waves are assumed to have long zonal wavelength. Then, as noted by Warn & Warn (1978), in certain special cases of the distance between the rigid wall and the critical line, the equation for the absolute vorticity in the critical layer has a closed-form analytic solution found by Stewartson (1978). The absolute vorticity $\hat{Q}$ may therefore be written explicitly as a known function of the suitably scaled and non-dimensional variables $\hat{x}$, Y and $\hat{T}$, respectively representing distance along and across the critical layer, and time. The distribution of $\hat{Q}$ at various times $\hat{T}$ during the evolution of the nonlinear critical layer is shown in figure 2 in KM and in figures 5–7 appearing later in this paper. As these figures make apparent, contours of $\hat{Q}$, which coincide with material contours, are simply deformed by flow in the critical layer. The streamlines of the flow take the form of Kelvin cat's eyes of constant size and shape in this special case.

The SWW solution serves as a valuable benchmark that reveals many aspects of nonlinear critical-layer behaviour. Of particular interest is the matching condition across the layer, which controls whether it acts as a wave absorber, reflector, or over-reflector. According to the SWW solution, the critical layer evolves from an initial absorbing state, through reflection, to an over-reflecting state. This progression is then reversed, and subsequently the critical layer continually changes from over-reflection to absorption and back again. The oscillations are of decreasing amplitude so that ultimately the critical layer tends toward a perfectly reflecting state.

However, it seems that the SWW solution does not tell the whole story in that the flow it predicts is unstable. In fact there is a hint of the unstable behaviour in Béland's (1976) numerical simulation (see especially his figure 5). The instability is possible because of the reversal of the meridional vorticity gradient at certain locations in the critical layer as may be seen in figure 2 of KM. In a parallel flow this would be a necessary condition for instability: it turns out that here, because of the assumption made by SWW that the forced waves have long zonal wavelength, which is essential for the analytic solution to be valid, the flow may be regarded as locally parallel. It is useful to express the long-wavelength condition quantitatively by defining a small parameter μ equal to the cross-stream lengthscale of the waves in the outer flow divided by the wavelength.

KM show that the scale of the disturbances to which the flow is unstable is much smaller than the wavelength of the forced waves, and the timescale on which they develop is much shorter than the evolution time for the critical layer as a whole. Thus, even though the vorticity distribution in the critical layer varies in the zonal direction, and changes with time, the stability problem may be tackled by considering the evolution of disturbances to a zonally uniform and time-independent basic state, which corresponds to the cross-stream vorticity distribution at a particular zonal location $\hat{x}$ and at a particular time $\hat{T}$. To describe the zonal structure and the variation with time of the disturbance it is necessary to define the new variables $x = \mu^{-1}\hat{x}$ and $T = \mu^{-1}\hat{T}$. The smallness of the parameter μ allows $\hat{x}$ and $\hat{T}$ to be held constant as x and T vary.

The starting point for the work presented in this paper is taken to be the equations which describe the evolution of a disturbance to the flow represented by the SWW solution. The derivation of these equations is set out in KM; however it is useful to

recall that they are obtained at leading order in a perturbation analysis based on the small parameter ϵ, which is a measure of the wave amplitude.

The equation for the leading-order disturbance vorticity $Q(x, Y, T)$ in the critical layer at particular values of $\hat{x}$ and $\hat{T}$ is given by

$$\left\{\frac{\partial}{\partial T}+Y\frac{\partial}{\partial x}+\frac{\partial C}{\partial x}\frac{\partial}{\partial Y}\right\}Q(x, Y, T)+\frac{\partial C}{\partial x}\frac{\partial \hat{Q}}{\partial Y}(\hat{x}, Y, \hat{T})=0. \tag{1.1}$$

This simply describes the advection of absolute vorticity $Q(x, Y, T)+\hat{Q}(\hat{x}, Y, \hat{T})$, being the disturbance vorticity plus the absolute vorticity given by the SWW solution, around streamlines described by the curves $\frac{1}{2}Y^2-C(x, T)=$ constant. These streamlines take the form of secondary cat's eyes, of the same width as the primary, forced cat's eye, but with much shorter length. The function $C(x, T)$ is related to the vorticity Q through the relation

$$\int_{-\infty}^{\infty}\tilde{Q}(k, Y, T)\,\mathrm{d}Y=-\pi\cot\frac{\pi}{2|k|}\tilde{C}(k, T), \tag{1.2}$$

where the notation $\tilde{}$ represents the Fourier transform with respect to x, so that

$$\tilde{C}(k, T)=\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx}C(x, T)\,\mathrm{d}x, \tag{1.3a}$$

$$\tilde{Q}(k, Y, T)=\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx}Q(x, Y, T)\,\mathrm{d}x. \tag{1.3b}$$

The equation (1.2) is an expression of the matching condition between the disturbance in the critical layer and that in the outer flow. Together with (1.1) it provides a closed mathematical system, which may be solved given an initial distribution of vorticity in the critical layer.

A linearized version of (1.1), which is valid when the disturbance is small so that the term $(\partial C/\partial x)(\partial Q/\partial Y)$ may be ignored, is solved in KM. They calculated the growth rates and spatial structures of the unstable modes and by finding fastest-growing modes confirmed that the assumptions made concerning the scale of the disturbances were correct. A short account of the linear stability characteristics, including some results not presented by KM, is given in §2.

The investigation described here has a well-defined goal: to demonstrate convincingly that the unstable disturbances grow to sufficient amplitude to disrupt completely the vorticity distribution predicted by SWW. This is accomplished by examining the behaviour of disturbances on zonally uniform basic states corresponding to those given by the SWW solution at particular times, and locations along the critical layer. Rather than solving the linear problem, the relevant *nonlinear* equations are integrated numerically and the disturbance is allowed to modify the vorticity distribution on which it grows. This task is made feasible by the fact that numerical calculation is necessary only in the critical layer, the whole effect of the outer flow being dealt with through the condition (1.2). Since any conclusions made in this paper are based on numerical results it is important to establish the reliability of the numerical methods used. These are described in §3, in some detail where the method is non-standard. The results themselves are then presented in §4. They demonstrate conclusively that the instability is strong enough to lead to a drastic rearrangement of the vorticity in the critical layer.

2. The linear instability problem

As reported in KM, some analytical progress may be made with the instability problem if it is assumed that the disturbance is weak, so that it does not modify the absolute vorticity gradient. In such a case the linearized version of (1.1) given by

$$\left\{\frac{\partial}{\partial T}+Y\frac{\partial}{\partial x}\right\}Q(x,Y,T)+\frac{\partial C}{\partial x}\frac{\partial \hat{Q}}{\partial Y}(\hat{x},Y,\hat{T})=0 \tag{2.1}$$

is appropriate.

Attention is confined to normal-mode solutions of the form

$$Q(x,Y,T)=\mathrm{Re}\{Q_c(Y)\,\mathrm{e}^{\mathrm{i}k(x-ct)}\}. \tag{2.2}$$

Substituting this into (2.1) and the matching condition (1.2) leads to the eigenvalue condition

$$\int_{-\infty}^{+\infty}\frac{\partial \hat{Q}}{\partial Y}(\hat{x},Y,\hat{T})\frac{\mathrm{d}Y}{Y-c}=\pi\cot\frac{\pi}{2|k|}, \tag{2.3}$$

as obtained by KM. The integral is evaluated along a contour passing below $Y=c$ in the complex-Y plane.

KM demonstate by example that, for values of $\hat{x}$ and $\hat{T}$ such that the absolute vorticity gradient $\partial\hat{Q}/\partial Y$ changes sign, unstable modes exist with growth rates that are quite sizeable. These results are presented in their table 1. Note that the results indicate that the stability characteristics of the vorticity pattern are asymmetric about the centre of the cat's eye. This is to be expected since (2.1) is not invariant under the transformation $\hat{x}\to-\hat{x}$, $Y\to-Y$. Here we note the further result, not shown in KM, that the condition that $\partial\hat{Q}/\partial Y$ is zero for some Y is both necessary and sufficient for the existence of unstable modes. An analogue of Rayleigh's theorem may be proved by considering the imaginary parts of each side of (2.3), giving the condition

$$\mathrm{Im}\,(c)\int_{-\infty}^{+\infty}\frac{\partial Q}{\partial Y}(\hat{x},Y,\hat{T})\frac{\mathrm{d}Y}{|Y-c|^2}=0. \tag{2.4}$$

Modes with $\mathrm{Im}\,(c)>0$ exist only if $\partial Q/\partial Y$ changes sign for some Y.

In order to prove the sufficiency of this condition we may appeal to a version of the Tollmien–Lin perturbation analysis (e.g. Drazin & Reid 1981, pp. 134–136), which predicts that every neutral mode has a neighbouring unstable mode. All that needs to be shown is that if $\partial\hat{Q}/\partial Y=0$ for some Y then at least one neutral mode exists. Consider the left-hand side of (2.3) as a function of the complex variable c. If we allow $\mathrm{Im}\,(c)\to0^+$ we obtain the result

$$\int_{-\infty}^{+\infty}\frac{\partial \hat{Q}}{\partial Y}\frac{\mathrm{d}Y}{Y-\mathrm{Re}\,(c)}+\pi\mathrm{i}\frac{\partial \hat{Q}}{\partial Y}(\hat{x},\mathrm{Re}\,(c),\hat{T})=\pi\cot\frac{\pi}{2|k|}. \tag{2.5}$$

The imaginary part of this equation may be satisfied by taking $\mathrm{Re}\,(c)$ equal to that value of Y for which $\partial\hat{Q}/\partial Y$ is zero – this then fixes the value of the real part of the left-hand side. However the function of k on the right-hand side takes all possible real values as k increases from zero to infinity. Thus a value of k certainly exists for which the real part of (2.5) is satisfied and so a neutral mode, and therefore a growing mode, exists.

3. Numerical methods

In order to integrate the nonlinear equations efficiently and accurately it was necessary to use a variety of numerical techniques, some standard and some non-standard. To support the claim that the numerical results presented in the next section are reliable, these techniques will be briefly described here.

It seemed natural to use a spectral representation of the variables Q and C in the x-direction, writing them as truncated Fourier series thus:

$$Q = \tfrac{1}{2}Q_0(Y,T) + \sum_{n=1}^{N} \{Q_n(Y,T)\cos knx + Q_{n+N}(Y,T)\sin knx\}; \tag{3.1}$$

$$C = \sum_{n=1}^{N} \{C_n(T)\cos nkx + C_{n+N}(T)\sin nkx\}. \tag{3.2}$$

Substitution of these expressions into (1.1) gives a set of coupled partial differential equations in Y and T for the Fourier coefficients Q_n. Similarly, each quantity C_n is related to the Y-integral of Q_n through (1.2), indeed the spectral representation seems a natural way to deal with this equation. In general there are a number of advantages to using a spectral method, the main one being the high spatial accuracy possible with a limited number of spectral components (see Gottlieb & Orszag 1977, and references). The possibility of using a spectral representation in the Y-direction was also considered, but no natural set of functions on which to base such a representation could be found.

The main disadvantage of using spectral methods is that multiplication is a 'non-local' operation in wavenumber space, in the sense that a convolution series over all harmonics must be evaluated. However, as proposed by Orszag (1969), a more economical method of evaluating products is to apply a Fourier transform to the spectral coefficients, perform the multiplication at a set of points in physical space, and then transform back. If fast transform algorithms are used this is more economical than direct convolution, and so this method was used here. 'Aliased' products were used throughout and no problems with this were encountered. (For a full discussion of aliasing see Orszag 1971.)

Explicit time integration of the partial differential equations for the Fourier coefficients was confined to a number of grid points at equally spaced values of Y between $+a$ and $-a$. However it was necessary to estimate the vorticity Q outside the domain $(-a, a)$ for two reasons: (i) to take account of the contribution to the integral expressions (1.2) for the C_n, which had to be evaluated at each step in the time integration; and (ii) to evaluate finite-difference expressions at the edges of the computational domain.

The required estimate was obtained by an approximate analysis of (1.1), on the basis that the dominant balance at large values of Y was between zonal advection of disturbance vorticity by the basic-state shear $Y(\partial q/\partial x)$ and cross-stream advection of the basic-state vorticity by the disturbance $(\partial C/\partial x)(\partial \hat{Q}/\partial Y)$. This is consistent with the proposed outer behaviour – both depend on the assumption that the disturbance vorticity in the outer region varies only on the inner timescale. It is therefore expected that this approximation scheme will be able to deal with the motion arising from any vorticity perturbation initially confined to the critical layer.

Assuming this dominant balance, it is straightforward to obtain an expression for Q as $|Y| \to \infty$ by substitution of successively higher approximations into the other terms in (1.1). The only further calculation needed is to determine the x-independent

part of Q, by considering the x-average of (1.1), and integrating with respect to time. It is found that Q takes the form

$$Q = -\frac{C}{Y}\frac{\partial \hat{Q}}{\partial Y} + \int^x \frac{\mathrm{d}x'}{Y^2}\frac{\partial C}{\partial T}(x', T)\frac{\partial \hat{Q}}{\partial Y} - \frac{\int^x \mathrm{d}x' \int^{x'} \mathrm{d}x''}{Y^3}\frac{\partial^2 C}{\partial T^2}(x'', T)\frac{\partial \hat{Q}}{\partial Y}$$
$$-\frac{1}{2}\frac{C^2}{Y^3}\frac{\partial \hat{Q}}{\partial Y} - \frac{1}{2}\frac{C^2}{Y^2}\frac{\partial^2 \hat{Q}}{\partial Y^2} - \frac{1}{2}\frac{\bar{C}^2}{Y^3}\frac{\partial \hat{Q}}{\partial Y} + o\left(\frac{1}{Y^3}\right) \quad (3.3)$$

for large $|Y|$, where $(\bar{\ })$ denotes the mean over x.

The integral of Q over Y that appears in the expression for C may be split into two parts, being the contributions from $(-a, a)$ and from outside this interval. Since Q is known at each gridpoint the first contribution may be calculated directly using Simpson's rule. The second contribution may be evaluated in terms of the (unknown) $\partial C/\partial T$ and $\partial^2 C/\partial T^2$ by integrating (3.3) over the intervals $(-\infty, -a)$ and (a, ∞). Then the equation for C_n in terms of Q_n may be rearranged to give an expression involving the unknown quantities on one side, and the (known) integral of Q_n over $(-a, a)$ on the other. This equation may be solved for C_n provided that the time derivatives $\partial C_n/\partial T$ and $\partial^2 C_n/\partial T^2$ can be eliminated.

Now, in the course of a numerical integration, one might expect that these could be written as finite-difference expressions in time, using information from previous time steps. However, when this was done, it was found that a rapidly growing numerical instability disrupted the solution. This instability was caused by the extra degree of freedom introduced into the equations by the time-difference terms, increasing their order by one, and allowing the growth of a parasitic solution. In fact the information necessary to calculate the time derivatives is present in (1.1) as it stands and may be calculated from the instantaneous values of C and Q. The details of this calculation, and the resulting estimate for C are given in Appendix A. It was this procedure that was followed in the numerical integration. Once the estimate for C had been obtained this could then be substituted back into (3.3), and used to find the values of Q needed to evaluate centred-difference expressions at the edges of the computational domain.

Finally, having dealt with the discretization process in x and Y, it is now appropriate to consider the method used to integrate (2.5) in time. One difficulty is the fact that the speed of the basic shear flow increases towards the edges of the computational domain. The equations for Q_n, obtained by substituting (3.1) and (3.2) into (1.1), are of the form

$$\frac{\partial Q_m}{\partial T} + mkYQ_{m+N} + mkC_{m+N}\frac{\partial \hat{Q}}{\partial Y} + N_m = 0 \quad (1 \leqslant m \leqslant N), \quad (3.4)$$

$$\frac{\partial Q_{m+N}}{\partial T} - mkYQ_m - mkC_m\frac{\partial \hat{Q}}{\partial Y} + N_{m+N} = 0 \quad (1 \leqslant m \leqslant N), \quad (3.5)$$

$$\frac{\partial Q_0}{\partial T} + N_0 = 0, \quad (3.6)$$

where N_m denotes a nonlinear term. It is clear that, when Y is large, the time steps taken will need to be quite small so that there is no numerical instability, due to the $Y(\partial Q/\partial x)$ term, near the edge of the computational domain. In order to allow the use of a larger time step these terms were dealt with implicitly, since the equations that result are easy to solve. The other terms are evaluated using an explicit Adams–

Bashforth method. The integration scheme, described in Appendix B, was then second-order accurate in time.

Any results of numerical simulations are valueless unless there is an independent check on the methods used. Three such checks were applied here.

(i) The growth rates of disturbances were compared with that predicted by the linear stability theory. However, this could give no information about the correctness of the nonlinear calculation.

(ii) The equation (1.1) also has an invariant related to the enstrophy. If (1.1) is multiplied by $\hat{Q}+Q$, and the mean with respect to x (but *not* over $\hat{x}$) is taken, it is found that

$$\frac{\partial}{\partial T}\{\overline{2\hat{Q}Q+Q^2}\}+\frac{\partial C}{\partial x}\frac{\partial}{\partial Y}\{\overline{2\hat{Q}Q+Q^2}\}. = 0. \tag{3.7}$$

It is not immediately clear that this expression may be integrated with respect to Y on $(-\infty, \infty)$. However, if Q is of the asymptotic form (3.3), the integral is well defined and so the result

$$\frac{\mathrm{d}}{\mathrm{d}T}\int_{-\infty}^{\infty}\overline{2\hat{Q}Q+Q^2}\,\mathrm{d}Y = 0 \tag{3.8}$$

may be obtained. The conserved quantity, Z say, was evaluated at a number of stages during the numerical calculation, using an asymptotic estimate for the contribution to the integral outside the computational domain.

(iii) A feature of (1.1) is that fluid particles moving with velocity $(Y, \partial C/\partial x)$ conserve the quantity $\hat{Q}+Q$. The motion of a number of such particles was followed in the calculation, and the quantity $\hat{Q}+Q$ calculated at each of their positions. The requirement that this quantity be conserved is an extremely demanding test of the local accuracy of the vorticity field predicted by the model.

4. Numerical experiments

The evolution equation (1.1) describes the nonlinear deformation of vorticity contours by the flow associated with the disturbance. Inherent in this process is a rapid cascade of the vorticity field to smaller and smaller spatial scales. The time taken for the cascade to reach the smallest spatial scale resolved in a numerical model provides an obvious limit on the time for which the vorticity field can be followed by that model. However it was found possible, despite this limit, to follow the evolution of the disturbance up to such time as the associated secondary cat's eye ceased to grow. This gave a first idea of the effectiveness of the disturbance in redistributing, through advection around the growing cat's eye, the vorticity in the critical layer.

In each case studied the initial form of the disturbance was set to be that of the linear normal-mode solution, with small amplitude. For each position $\hat{x}$ along the critical layer and time $\hat{T}$ in its evolution considered, the wavenumber of the disturbance in the x-direction was assumed to be that of the fastest-growing mode according to linear theory. As the amplitude of the disturbance grew, higher harmonics to that initially present were forced by the nonlinear terms in the equations.

Some preliminary experiments were carried out to assess suitable choices for the size of the computational domain and the resolution of the numerical model. It was found that taking the boundaries of the domain at Y equal to $+5$ and -5 and using 201 grid points evenly spaced across this domain gave satisfactory results. Indications

were that these choices gave errors of less than 3% (in terms of initial growth rates, width of regions over which the vorticity was drastically rearranged, etc.) and this aspect of the errors associated with the computational scheme will not be discussed further. Attention will instead be focused on the effects of the finite number of degrees of freedom in the x-direction.

One would expect the details of the vorticity field to be sensitive to lack of spatial resolution. However, one might also expect that the stream function induced by this vorticity field would be less sensitive, since the operation of calculating one from the other involves an inverse Laplacian, which acts as a spatial low-pass filter. In fact, in this particular case the quantity C, representing the Fourier transform of the stream function, is given by (1.2). Since the wavenumber-dependent factor is proportional to k when k is large, the disturbance stream-function pattern should be relatively insensitive to the small-scale vorticity field.

This intuition was confirmed by running the numerical model with parameters appropriate to a particular location in the critical layer and a particular stage in its evolution, but varying the number of harmonics included in the x-direction. Table 1 shows the evolution in time with the number of the component of stream function in the first harmonic. The case displayed is that of $\hat{T} = 2\sqrt{2}$, corresponding to KM's figure 2(*c*), with $\hat{x} = \frac{1}{4}\pi$. It may be seen that there is little difference in the results if seven or more harmonics are present.

It could be argued that a higher-harmonic component would be more sensitive to the number of harmonics. However, the first harmonic is the dominant part of the stream function, and it appears that small changes in other harmonics are unimportant to the rearrangement of vorticity. This is supported by the evidence in figure 1, which shows the spectrum of C at various times during one of the numerical experiments represented in table 1, for the particular case of 15 harmonics. Although the amplitudes of higher harmonics do increase in time relative to that of the first, they clearly never become as important.

In contrast, evidence from these numerical experiments suggests that the vorticity field *is* sensitive to the number of harmonics. The enstrophy spectrum, shown in figure 1, fills rather rapidly at the higher wavenumbers resolved. In addition, direct comparison of the vorticity fields predicted by models with varying zonal resolution reveals strong differences. Figure 2 shows the vorticity field at a particular time in the evolution of the disturbance, for runs with 7 and 31 harmonics. The contours correspond to material lines, which are aligned in the x-direction when the disturbance amplitude is very small and are subsequently deformed by the flow. The appearance of closed contours is therefore an indication that the model resolution is insufficient to reproduce the true vorticity field.

Also shown in each figure, as a solid line, is the mean-vorticity profile, being that obtained by averaging over a number of wavelengths of the disturbance (on the x-scale), rather than that over the wavelength of the forced wave (on the $\hat{x}$-scale). The initial profile, present before the disturbance started to grow and so corresponding to a local profile at a particular time $\hat{T}$ in the SWW solution, is shown by the dashed line. In addition a linear profile, corresponding to the β-plane distribution of absolute vorticity, is shown by the dotted line.

Superimposed on each vorticity plot is the secondary 'cat's eye' produced by the first harmonic part of the disturbance stream function only. For reasons already stated this is believed to be a good approximation to that which would be obtained by including all the harmonics. As can be seen from figure 2(*b*), at $T = 5.5$, when the disturbance had almost grown to its maximum size (which is attained at $T = 5.7$),

T	$N = 3$	$N = 7$	$N = 15$	$N = 31$
4.0	0.0115	0.0167	0.0167	0.0167
4.2	0.1327	0.1341	0.1341	0.1341
4.4	0.2391	0.2403	0.2403	0.2403
4.6	0.3329	0.3339	0.3340	0.3339
4.8	0.4125	0.4131	0.4131	0.4131
5.0	0.4768	0.4766	0.4766	0.4766
5.2	0.5256	0.5231	0.5231	0.5230
5.4	0.5594	0.5526	0.5526	0.5526
5.8	0.5841	0.5632	0.5629	0.5629
6.0	0.5723	0.5472	—	0.5464

TABLE 1. Values of $(C_1^2 + C_{N+1}^2)^{\frac{1}{2}}$. The value of the first-harmonic component of the stream function $(C_1^2 + C_{N+1}^2)^{\frac{1}{2}}$ attained at time T, for models with differing numbers of harmonics N included. These results are for $\hat{x} = \frac{1}{4}\pi$ and $\hat{T} = 2\sqrt{2}$ (corresponding to figure 6).

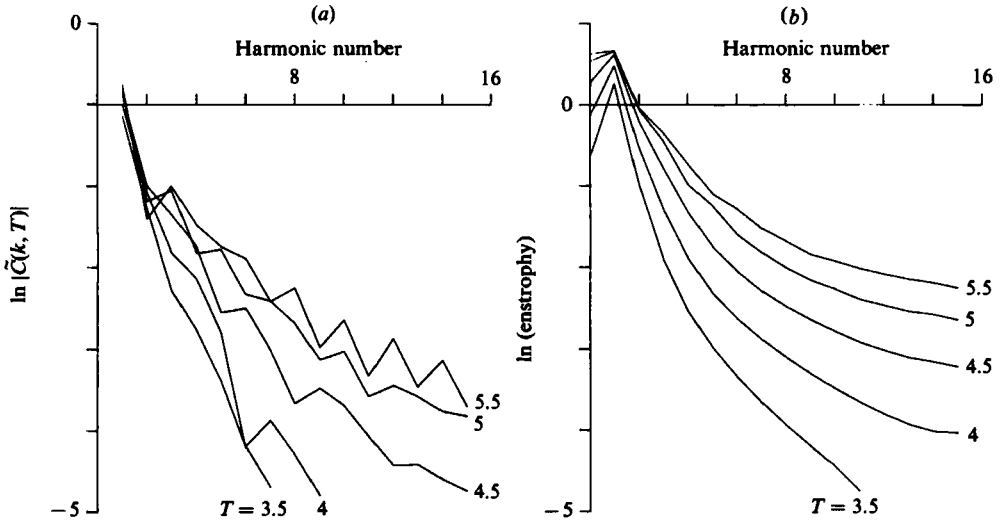

FIGURE 1. Amplitudes of (*a*) the zonally varying part of the stream function and (*b*) the enstrophy, in each wavenumber component at times T of 3.5, 4.0, 4.5, 5.0 and 5.5, in a case with 15 harmonics included in the calculation.

the mean vorticity had effectively been redistributed across the whole region defined by the limits of the cat's eye. A very important result, manifest in table 1 and in figure 2 (*a*), was that the width of this region was predicted rather well by the model with only seven zonal harmonics, even though this was unable to reproduce accurately the details of the vorticity field.

It is again stressed that the wavelength of the disturbance is very much less than that of the forced wave, so that if a diagram of the whole cat's eye, e.g. KM figure 2 or figures 5–7 in this paper, were drawn to be in proportion with figure 2, it would have to be stretched an enormous amount in a direction parallel to the x-axis. Thus, although the vorticity pattern in figure 2 looks superficially the same as that in KM figure 2 (*c*) for example, it would not necessarily be unstable in the same way because its wavelength is comparable with the relevant outer-flow lengthscale and it cannot be considered as a locally parallel flow.

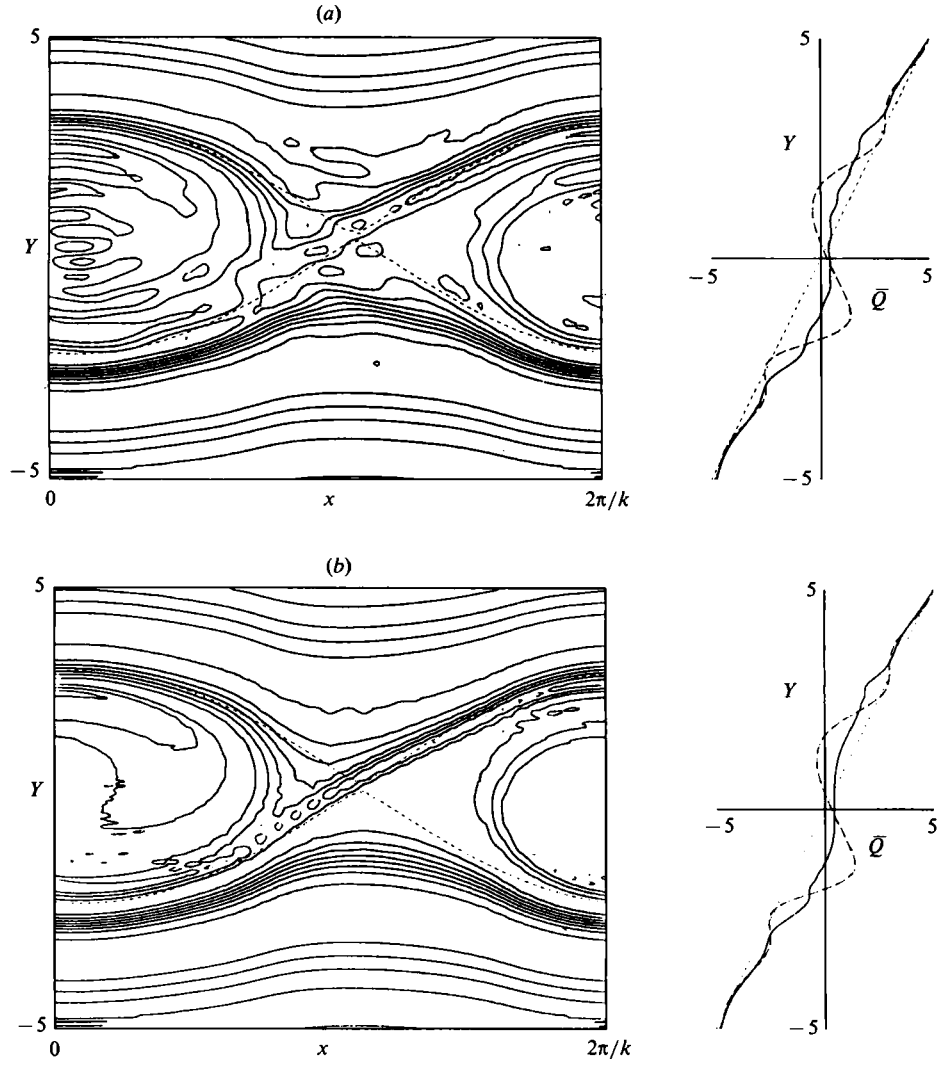

FIGURE 2. The absolute vorticity field at a late stage of a model run with $\hat{T} = 2\sqrt{2}$, corresponding to KM's figure 2, and $\hat{x} = -\frac{1}{4}\pi$. Each of the upper ($a$) and lower ($b$) figures consists of two parts. On the left absolute vorticity contours are shown at intervals of 0.5 over the entire computational domain. Superimposed on the contours is the instantaneous shape of the growing cat's eye, shown as a finely dashed line, around which material is being advected.

On the right the meridional profile of absolute vorticity is shown. The solid line is the instantaneous mean (over x) profile. The coarsely dashed line is that which was present when the disturbance had infinitesimal amplitude and is identical to the local profile, at a particular value of $\hat{x}$, predicted by the SWW solution. The finely dashed line is the original beta-plane profile, which existed before the forcing was switched on.

In both cases the model included 201 grid points evenly spaced in Y across the domain. However in (a) the model contained 7, and in (b) 31, harmonics in the x-direction.

The three checks on the reliability of the numerical method which were described in §3 were applied to this case. It was found that:

(i) the growth rates predicted by linear theory and observed in the model behaviour differed by less than 0.2%;

(ii) the invariant quantity Z varied by less than 2% of the amount by which its component parts (e.g. the disturbance enstrophy) changed during the course of the numerical integration.

The results of applying the most demanding test, (iii), are most easily shown graphically. Figures 3(*a*), (*b*), (*c*) and (*d*) show the change in position of thirty particles, denoted by crosses, as the disturbance grew.

These particles all lie on a single material line, which was originally oriented zonally. This line was then twisted up, by the original cat's eye velocity field due to the forced wave and, as shown, is now being twisted up again, on a smaller x-scale, by the secondary cat's eye due to the growing disturbance.

It may be seen that, even with 31 harmonics (and 63 degrees of freedom in the x-direction), particles do not conserve their vorticity exactly, up to the time $T = 5.7$. However, the fact that the vorticity remained more or less invariant for each particle, at least as the growth rate began to decrease, gave confidence in the accuracy of the numerical methods and the program code. For reasons already cited it is believed that the discrepancy is not important as regards the ability of the model to predict the evolution of the disturbance. It is on small scales and therefore is likely to represent vorticity that is 'passive' rather than 'active'.

At different locations in the critical layer, disturbances have different growth rates, and reach different ultimate amplitudes. For example, figures 4(*a*) and (*b*) show the vorticity fields, leading-harmonic cat's eye outlines and mean-vorticity profiles, as the fastest-growing disturbance reaches its maximum amplitude at the locations $\hat{x} = \frac{1}{2}\pi$ and $\hat{x} = -\frac{1}{2}\pi$. These figures should reinforce the idea that the width of the secondary cat's eye is a good indication of the size of the region over which vorticity will be redistributed.

To get an overall picture of the effect of the instability on the vorticity distribution in the critical layer, a number of low-resolution (7 harmonic) calculations were then performed, for various locations in the critical layer, and for different stages in its evolution. Using the maximum value attained by the first-harmonic component of the stream function, the maximum width attained by the secondary 'cat's eyes' could be calculated to good approximation. Then, assuming that these were centred where the mean flow was equal to the real part of the phase speed of the growing disturbance as predicted by the linear theory, the values of Y between which vorticity would be stirred up were calculated. In figures 5–7 these values have been marked on the corresponding plots of the absolute vorticity as predicted by the SWW theory. Recall that the left-to-right extent of these pictures is very much greater than that in figures 2–4. The vertical bars superimposed on the plots for $\hat{T} = 3\sqrt{2}/2$, $2\sqrt{2}$ and $3\sqrt{2}$, therefore, suggest the extent to which the instability is likely to rearrange the vorticity in the critical layer. In particular, it is clear that the regions affected do not necessarily increase in size as $\hat{T}$ increases, even though the vorticity gradients become stronger. The asymmetry of the regions, both from top to bottom and left to right of the figures, is also striking, particularly in figure 7.

5. Discussion

The main result of this paper is the confirmation that the instability discovered by KM is effective in substantially rearranging the vorticity in the nonlinear critical layer. Indeed figures 5–7 indicate that such rearrangement will drastically change the vorticity distribution and therefore, after differential advection by the basic shear has had time to act, the matching condition to the outer flow, through (1.2). This

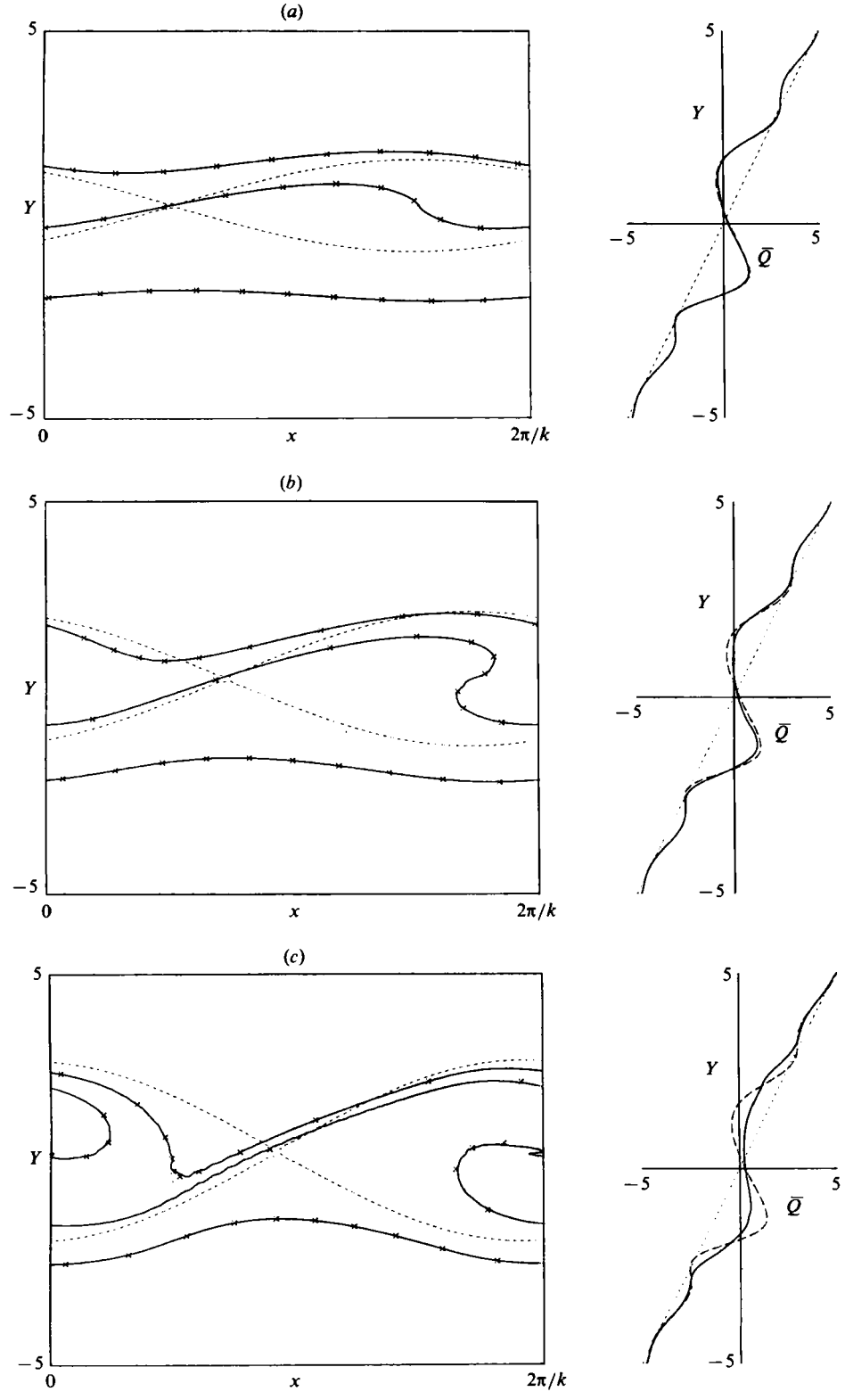

FIGURE 3 (*a–c*). For caption see facing page.

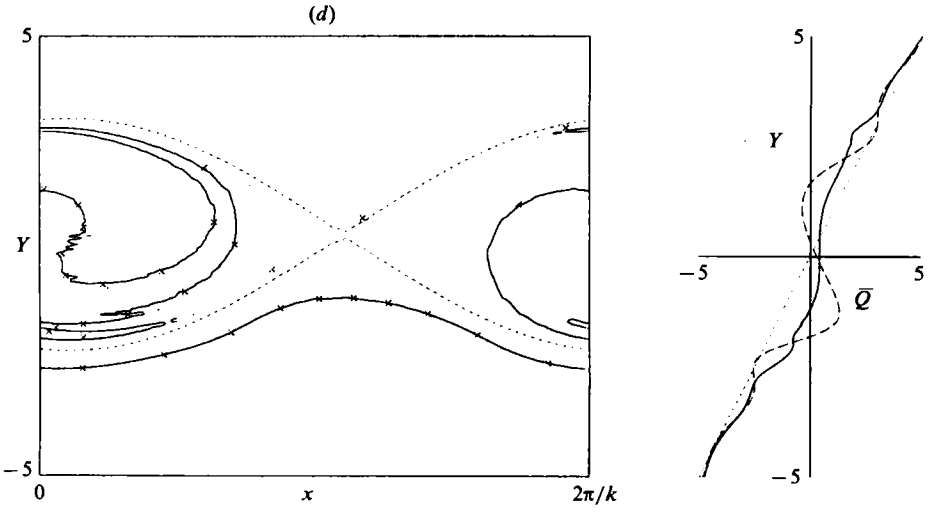

FIGURE 3. Each figure shows stages in the evolution of a single absolute-vorticity contour (value -0.1) as it is twisted up by the growing disturbance. The changes in position of 30 particles, initially lying on this contour, were calculated independently of the time integration of the vorticity equation. The extent to which the particles remain on the contour is a demanding test of the accuracy of the numerical integration.

In the figures the positions of the particles are shown by crosses, otherwise the graphical conventions used are those described in the caption to figure 2. The times T corresponding to each figure are (*a*) 4.0, (*b*) 4.5, (*c*) 5.0, (*d*) 5.5. In the case displayed 31 harmonics were used to represent the x-dependence of the vorticity field.

represents an important modification to the SWW theory. The precise consequences for the long-time evolution are unknown at present. All we know is that, as proved in KM, the instability cannot cause the critical layer to remain a wave absorber (or become an over-reflector) in the time mean.

The results of the linear stability analysis suggested that the instability might have a stronger effect in the left-hand half of the cat's eye, as it is shown in figures 5–7, insofar as growth rates were larger there (recall KM's table 1). On the basis of this result and the fact that the real parts of the phase speeds of growing modes were generally positive, one might speculate that the vorticity would be redistributed most effectively somewhat above and to the left of the centre of the cat's eye. The results of the nonlinear calculations presented here are a reminder that, in general, 'linear' growth rates are a poor guide to the ultimate amplitude that an unstable disturbance will reach. The regions shown in figures 5–7 are not wider on the left-hand side of the cat's eye, although they are generally centred in its upper half.

There is one obvious criticism of the approach taken in this paper to the problem of the evolution of a nonlinear critical layer: that the basic states on which the disturbances were allowed to grow would be unlikely to be achieved. The instability would become effective almost as soon as the first reversal of the vorticity gradient appears (at $\hat{T} = \frac{1}{2}\pi$). What would presumably happen is that the instability would lead to regions in which the vorticity had been considerably rearranged and which slowly changed in size and shape. No attempt has been made in this paper to solve such a problem.

The regions shown in figures 5–7 might be thought of as resulting from an extremely careful experiment in which the nonlinear critical layer evolved, undisturbed, until

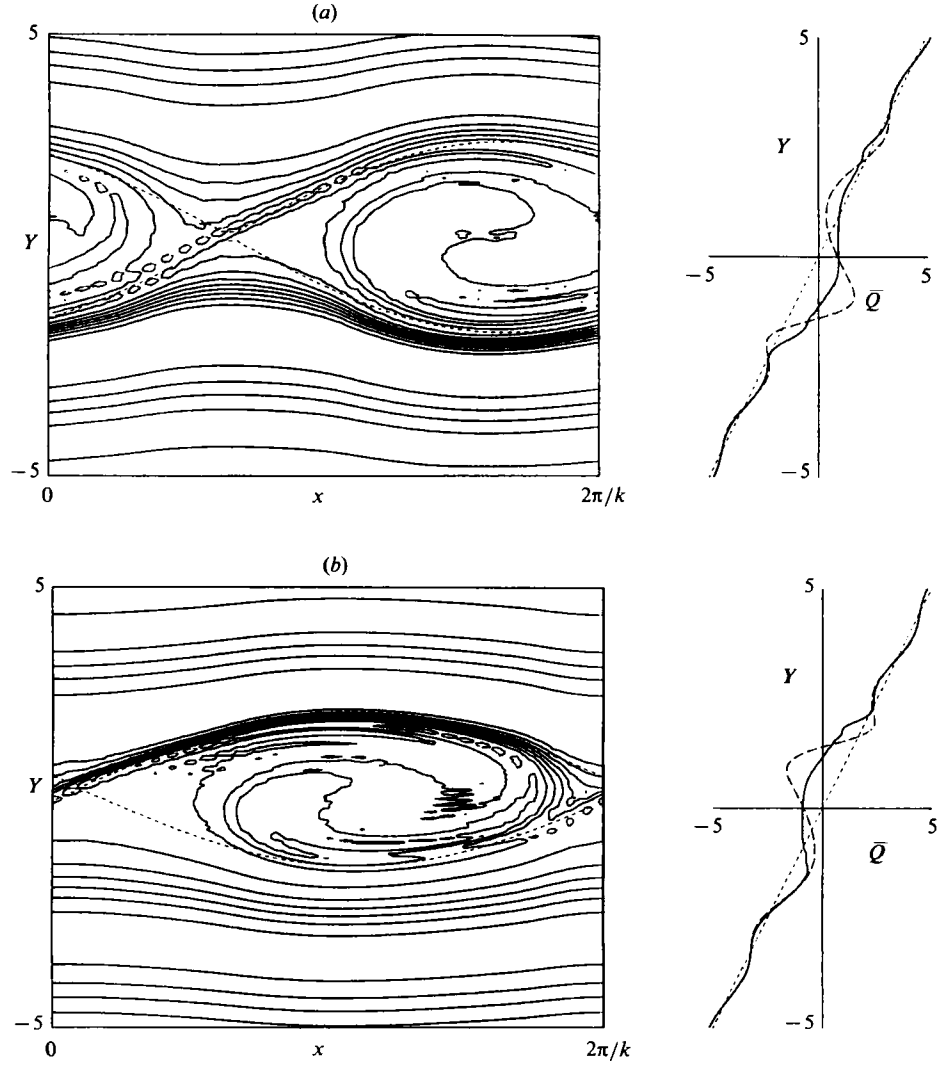

FIGURE 4. This figure shows the absolute vorticity field at the instant the growing disturbance reached its maximum amplitude. Graphical conventions are as for figure 2, and 31 harmonics were used to represent the x-dependence of the vorticity field. Results are displayed for basic states corresponding to the time $\hat{T} = 2\sqrt{2}$ in the evolution of the nonlinear critical layer at values of $\hat{x}$ of (*a*) $\frac{1}{4}\pi$ and (*b*) $-\frac{1}{4}\pi$.

a certain time when a small amount of 'noise' was introduced. The amplitude of the disturbance triggered by such noise would then grow extremely rapidly, and the vorticity field would be stirred up.

The difference in timescale between that for the eddies, and that for the critical layer as a whole, is a consequence of the long-wave assumption which allows the SWW solution. It provides an important simplification that makes the instability problem much more tractable, and is difficult to say exactly what would happen if it were relaxed. Here the instability is spectacularly efficient at rearranging the vorticity because the time which it takes to develop is so short compared with that for the

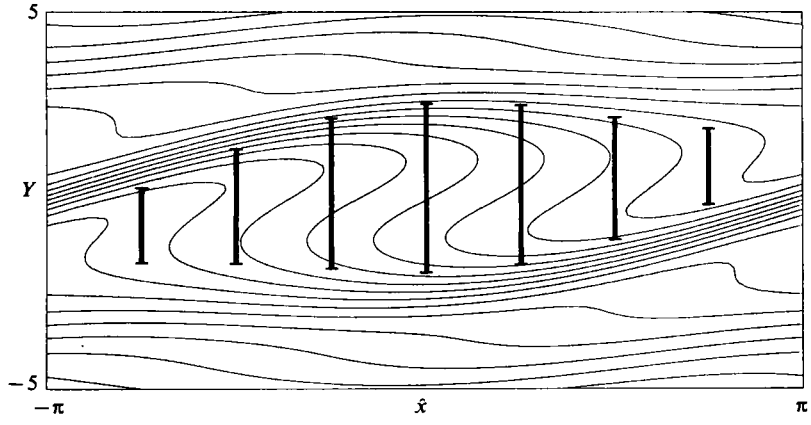

FIGURE 5. Contours of absolute vorticity in the SWW nonlinear critical layer at $\hat{T} = 3\sqrt{2}/2$. Superimposed are vertical lines which show the maximum extent of the disturbance cat's eye at the corresponding value of $\hat{x}$.

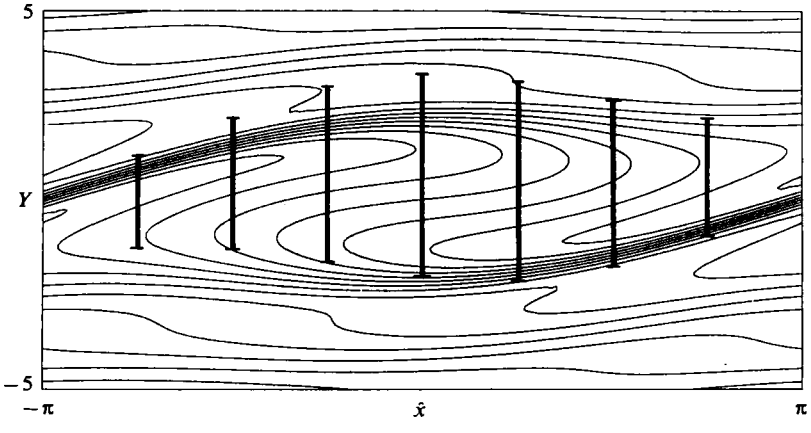

FIGURE 6. Contours of absolute vorticity in the SWW nonlinear critical layer at $\hat{T} = 2\sqrt{2}$. Superimposed are vertical lines which show the maximum extent of the disturbance cat's eye at the corresponding value of $\hat{x}$.

evolution of the critical layer as a whole. With shorter wavelength for the forced waves it is possible that the instability would not be so effective in rearranging the vorticity field.

The only concrete example we have is in the numerical simulation of Béland (1976). There the value of μ is 0.25 whereas here it has been assumed to be very small. Consequently, in Béland's case, the instability would not be expected to grow so rapidly on the timescale of evolution of the whole nonlinear critical layer. The fact that disturbances grow more rapidly in a particular location might then be more important. Indeed this might explain why the disturbance that could be growing as a result of the critical-layer instability seems strongest in the upper left-hand part of the cat's eye. The possibility also remains that the cat's eye distortions seen in Béland's simulation are simply a result of the feedback between the changing vorticity distribution in the critical layer and the shape and size of the cat's eyes. This feedback is almost always present, except in the special case in which the SWW

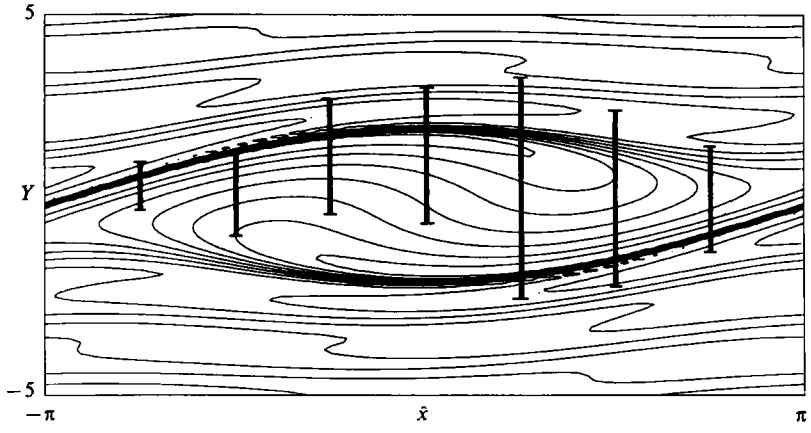

FIGURE 7. Contours of absolute vorticity in the SWW nonlinear critical layer at $\hat{T} = 3\sqrt{2}$. Superimposed are vertical lines which show the maximum extent of the disturbance cat's eye at the corresponding value of $\hat{x}$.

solution is valid, and may be enhanced if the outer flow is in a configuration for which the forced wave or one of its higher harmonics is near resonance (Ritchie 1985).

It is likely that a 'local' analysis, studying the instability near $\hat{T} = \frac{1}{2}\pi$, could have been performed. There is presumably some asymptotic regime in which the growth rate of disturbances is comparable with the rate of change of the vorticity profile and the instability problem is an explicitly time-dependent one. However, it would have been impossible to follow the evolution of the critical layer for very long with such an approach, and it would be unlikely to give insight about where the instability would stir up the vorticity. Further, it would have been difficult to extend the scope of the methods used in such a calculation.

Although the analytic solution of Stewartson (1978) represents an important step forward in our understanding of nonlinear critical layers, it can be applied only to the case of long wavelength and to particular forcing configurations. Probably the only realistic approach to the whole problem of the evolution of the nonlinear critical layer, including the effects of the instability and possible resonances, will be a high-resolution numerical simulation. However, rather than using Béland's (1976, 1978) approach in which the entire problem was solved numerically, it is likely to be much more efficient and effective to use the method of matched asymptotic expansions, and solve numerically in an inner region only. In fact Warn & Warn (1978) did this successfully, but only for the long-wavelength limit ($\mu \to 0$). Work on the numerical solution of the critical-layer equations for finite μ, allowing the possibility of instability and resolving the resulting vigorous eddy motion, is in progress and will be reported elsewhere.

This work formed part of the author's Ph.D. thesis, completed at the University of Cambridge with financial support from the UK Natural Environment Research Council. Advice from Drs P. D. Killworth and M. E. McIntyre is gratefully acknowledged. Preparation of the manuscript was supported by the Joint Institute for the Study of the Atmosphere and Ocean.

Appendix A. Method of calculating C

The limits of the computational grid are set to be $Y = +a$ and $Y = -a$ respectively. Then, using the asymptotic form (3.3), it may be shown that

$$\int_{-\infty}^{\infty} \tilde{Q}\,\mathrm{d}Y = \int_{-a}^{a} \tilde{Q}\,\mathrm{d}Y + \frac{2}{aik}\frac{\partial \tilde{C}}{\partial T} + \frac{[\hat{Q}(\hat{x}, a, \hat{T}) + Q(\hat{x}, -a, \hat{T})]}{a}\tilde{C}, \tag{A 1}$$

where $$Q = \frac{1}{2\pi}\int \mathrm{e}^{\mathrm{i}kx}\tilde{Q}(k, Y, T)\,\mathrm{d}k \quad \text{and} \quad C = \frac{1}{2\pi}\int \mathrm{e}^{\mathrm{i}kx}\tilde{C}(k, T)\,\mathrm{d}k.$$

Now, using (1.2) and writing $-\pi\cot\pi/2k = \Theta(k)$, the result

$$\tilde{C}(k, T) = \Theta(k)\int_{-a}^{a} \tilde{Q}\,\mathrm{d}Y + \frac{2\Theta(k)}{aik}\frac{\partial \tilde{C}}{\partial T} + \Theta(k)\tilde{C}\frac{[Q(\hat{x}, a, \hat{T}) + \hat{Q}(\hat{x}, -a, \hat{T})]}{a} \tag{A 2}$$

is obtained, which serves as a basis for the estimation of $\tilde{C}$.

The first approximation to $\tilde{C}$ is defined by

$$\tilde{C}^{(1)} = \Theta(k)\int_{-a}^{a} \tilde{Q}\,\mathrm{d}Y \tag{A 3}$$

and is correct to $O(a^{-1})$. This is now used to estimate the $\partial\tilde{C}/\partial T$ appearing in (A 2) and so provide an approximation to $\tilde{C}$ correct to $O(a^{-2})$.

Differentiating (A 3) with respect to T, and then substituting for $\partial\tilde{Q}/\partial T$ using the Fourier transform of (1.1), leads to

$$\frac{\partial \tilde{C}^{(1)}}{\partial T} = -\Theta(k)\int_{-a}^{a}\left\{\mathrm{i}kY\tilde{Q} + \mathrm{i}k\tilde{C}\frac{\partial \hat{Q}}{\partial Y} + \tilde{N}\right\}\mathrm{d}Y, \tag{A 4}$$

where $N = (\partial C/\partial x)\,\partial Q/\partial Y$. The nonlinear term N only gives an $O(a^{-1})$ contribution when integrated with respect to Y, and so is discarded at this stage in the approximation procedure. However, the quantity $\tilde{C}$ is effectively multiplied by a when the integration is performed, and so this must not be replaced by the approximation $\tilde{C}^{(1)}$.

On substituting (A 4) into (A 2), and then isolating $\tilde{C}$, the next approximation $\tilde{C}^{(2)}$, which is correct to $O(a^{-2})$, is found to be

$$\tilde{C}^{(2)} = \frac{\Theta(k)\int_{-a}^{a}\tilde{Q}\,\mathrm{d}Y - \dfrac{2\Theta(k)^2}{a}\int_{-a}^{a} Y\tilde{Q}\,\mathrm{d}Y}{1 + [2\Theta(k)^2/a]\{\hat{Q}(\hat{x}, a, T) - \hat{Q}(\hat{x}, -a, T)\}}. \tag{A 5}$$

This procedure may be repeated, differentiating (A 5) with respect to T and then substituting again from (1.1). However, at this stage the nonlinear terms must be retained, but, consistent, with the approximation, $\tilde{C}^{(2)}$ may be used instead of the next approximation, $\tilde{C}^{(3)}$ in these terms. The result is that

$$\tilde{C}^{(3)}(k, T) = \left\{\Theta(k)\int_{-a}^{a}\tilde{Q}\,\mathrm{d}Y - \frac{2\Theta(k)^2}{a\Xi}\left[\int_{-a}^{a} Y\tilde{Q}\,\mathrm{d}Y - \frac{2\Theta(k)}{a}\int_{-a}^{a} Y^2\tilde{Q}\,\mathrm{d}Y\right] - \frac{2\Theta(k)^2\tilde{P}}{aik\Xi}\right\}R^{-1}, \tag{A 6}$$

where $$\Xi = 1 + \frac{2\Theta(k)^2}{a}\{\hat{Q}(\hat{x}, a, \hat{T}) - \hat{Q}(\hat{x}, -a, \hat{T})\}$$

and

$$\tilde{P} = \{(ik_1 \tilde{C}^{(2)}(k_1, T)\} * \Big\{[\tilde{Q}(k_2, Y, T) - 2\Theta(k_2) Y \tilde{Q}(k_2, Y, T)]_{-a}^{a} + \frac{2\Theta(k_2)}{a} \int_{-a}^{a} \tilde{Q}(k_2, Y, T)\, \mathrm{d}Y \Big\}.$$

The notation $f(k_1) * g(k_2)$ is used to denote the convolution of f and g. Finally R is given by

$$R = 1 + \frac{2\Theta(k)^2}{a\Xi} \Big\{ \hat{Q}(\hat{x}, a, \hat{T}) - \hat{Q}(\hat{x}, -a, \hat{T}) - 2\Theta(k) \Big[\hat{Q}(\hat{x}, a, \hat{T}) + \hat{Q}(\hat{x}, -a, \hat{T}) + \frac{2\Theta(k)}{a} \int_{-a}^{a} \tilde{Q}(x, Y, T)\, \mathrm{d}Y \Big] \Big\} - \frac{\Theta(k)}{a} [\hat{Q}(\hat{x}, a, \hat{T}) + \hat{Q}(\hat{x}, -a, \hat{T})].$$

This approximation $\tilde{C}^{(3)}(k, T)$, correct to $O(a^{-3})$, was that used to represent $\tilde{C}(k, T)$ in the numerical integration. The asymptotic form of Q was also needed, for evaluation of the centred space differences at the edge of the computational domain. For this purpose an approximation to $\partial \tilde{C}/\partial T$ was carried along in the calculation. The approximation was evaluated by setting the left-hand side of (A 2) equal to $C^{(3)}(k, T)$ and then multipying by a, and was correct to $O(a^{-2})$.

Appendix B. The time-stepping method

In this Appendix the notation of §3 is used, with the addition that the notation $Q_m^{(T)}$ is used for the value of the variable Q_m at time T.

The equations for the quantities Q_m, regarded as continuous functions of time, are of the form

$$\frac{\partial Q_m}{\partial T} + km\, Y Q_{m+N} + km\, C_{m+N} \frac{\partial \hat{Q}}{\partial Y} + N_m = 0 \quad (1 \leqslant m \leqslant N), \tag{B 1a}$$

$$\frac{\partial Q_{m+N}}{\partial T} - km\, Y Q_m - km\, C_m \frac{\partial \hat{Q}}{\partial Y} + N_{m+N} = 0 \quad (1 \leqslant m \leqslant N), \tag{B 1b}$$

$$\frac{\partial Q_0}{\partial T} + N_0 = 0. \tag{B 2}$$

These were approximated, correct to $O(\Delta T^2)$, by the finite time-differenced forms

$$\frac{Q_m^{(T+\Delta T)} - Q_m^{(T)}}{\Delta T} + \tfrac{1}{2} km\, Y (Q_{m+N}^{(T+\Delta T)} + Q_{m+N}^{(T)}) + \frac{3}{2}\Big(C_{m+N}^{(T)} \frac{\partial \hat{Q}}{\partial Y} + N_m^{(T)}\Big) - \frac{1}{2}\Big(C_{m+N}^{(T-\Delta T)} \frac{\partial \hat{Q}}{\partial Y} + N_m^{(T-\Delta T)}\Big) = 0 \quad (1 \leqslant m \leqslant N), \tag{B 3a}$$

$$\frac{Q_{m+N}^{(T+\Delta T)} - Q_{m+N}^{(T)}}{\Delta T} - \tfrac{1}{2} km\, Y (Q_m^{(T+\Delta T)} + Q_m^{(T)}) + \frac{3}{2}\Big(N_{m+N}^{(T)} - C_m^{(T)} \frac{\partial \hat{Q}}{\partial Y}\Big) - \frac{1}{2}\Big(N_{m+N}^{(T-\Delta T)} - C_m^{(T-\Delta T)} \frac{\partial \hat{Q}}{\partial Y}\Big) = 0 \quad (1 \leqslant m \leqslant N), \tag{B 3b}$$

$$\frac{Q_0^{(T+\Delta T)} - Q_0^{(T)}}{\Delta T} + \tfrac{3}{2} N_0^{(T)} - \tfrac{1}{2} N_0^{(T-\Delta T)} = 0. \tag{B 4}$$

For each m (B 3a) and (B 3b) are simultaneous equations in the quantities $Q_m^{(T+\Delta T)}$ and $Q_{m+N}^{(T+\Delta T)}$. Their form is so simple that a trivial algebraic manipulation yields explicit expressions for these quantities.

REFERENCES

ABRAMOWITZ, M. & STEGUN, I. A. 1966 *Handbook of Mathematical Functions.* US National Bureau of Standards.

BÉLAND, M. 1976 Numerical study of the nonlinear Rossby-wave critical level development in a barotropic zonal flow. *J. Atmos. Sci.* **33**, 2066–2078.

BÉLAND, M. 1978 The evolution of a nonlinear Rossby-wave critical layer: effects of viscosity. *J. Atmos. Sci.* **35**, 1802–1815.

DRAZIN, P. G. & REID, W. H. 1981 *Hydrodynamic Stability.* Cambridge University Press.

DICKINSON, R. E. 1970 Development of a Rossby wave critical level. *J. Atmos. Sci.* **27**, 627–633.

GOTTLIEB, D. & ORSZAG, S. A. 1977 *Numerical Analysis of Spectral Methods: Theory and Applications.* SIAM.

KILLWORTH, P. D. & MCINTYRE, M. E. 1985 Do Rossby-wave critical layers absorb, reflect or over-reflect? *J. Fluid Mech.* **161**, 449–492.

ORSZAG, S. A. 1969 Numerical methods for the simulation of turbulence. *Phys. Fluids Suppl.* **12**, II 250–257.

ORSZAG, S. 1971 Numerical simulation of incompressible flows within simple boundaries: accuracy. *J. Fluid Mech.* **49**, 75–112.

RITCHIE, H. 1985 Rossby-wave resonance in the presence of a nonlinear critical layer. *Geophys. Astrophys. Fluid Dyn.* **9**, 185–200.

STEWARTSON, K. 1978 The evolution of the critical layer of a Rossby wave. *Geophys. Astrophys. Fluid Dyn.* **9**, 185–200.

STEWARTSON, K. 1981 Marginally stable inviscid flows with critical layers. *IMA J. Appl. Math.* **27**, 133–175.

WARN, T. & WARN, H. 1976 On the development of a Rossby wave critical level. *J. Atmos. Sci.* **33**, 2021–2024.

WARN, T. & WARN, H. 1978 The evolution of a nonlinear Rossby wave critical layer. *Stud. Appl. Math.* **59**, 37–71.