J. Fluid Mech. (1989), *vol.* 207, *pp.* 231–266
Printed in Great Britain

The effect of barotropic instability on the nonlinear evolution of a Rossby-wave critical layer

By PETER H. HAYNES

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge, CB3 9EW, UK

(Received 27 August 1987 and in revised form 16 December 1988)

A study of the flow within the critical layer of a forced Rossby-wave is made, using a high-resolution numerical model. The possibility of growth of disturbances through barotropic instability and the extent to which these disturbances modify the subsequent time evolution is of particular interest. The flow is characterized by a parameter μ, equal to the cross-stream lengthscale divided by a downstream wavelength. In the long-wavelength case, $\mu \ll 1$, where there is a clear conceptual division between the instability and the basic flow, the results of the simulation confirm the importance of the growing and saturating disturbances in rearranging the vorticity within the critical layer. When the wavelength is not so long, the distinction between the instability and the straightforward time evolution of the basic flow is less clear. Nonetheless for $\mu < 0.25$ the ultimate evolution is still sensitive to the details of the initial perturbations and in this sense the flow may be regarded as being unstable. The time-integrated absorptivity of the critical layer may be considerably increased by the effects of the instability, sometimes to three or four times that predicted by the Stewartson–Warn–Warn solution. The nature of the flow, at least during the period in which the dynamics are essentially inviscid, does not seem to change when higher harmonics to the forced wave are resonant. The behaviour seen in Béland's (1976) numerical model is re-examined in the light of these findings. A simple model of the redistribution of vorticity by the unstable disturbances is formulated, and its predictions are shown to agree well with the numerical simulations.

1. Introduction

This paper will be concerned with Rossby-wave critical layers in which the role of dissipative processes is relatively minor, the flow in the critical layer is highly nonlinear and the vorticity field is evolving in time. This is the parameter regime that seems most relevant to many geophysical flows, at least in the sense that the fluid-dynamical processes found to take place in the critical layer are also believed to be those that are often important on large scales in the atmosphere and the ocean. State-of-the-art reviews of the theory of critical layers, including summaries of the physical background which motivates such theory, have been given by Stewartson (1981), and Maslowe (1986). Here it is simply noted that continuing interest in the Rossby-wave problem seems justified because the nonlinear Rossby-wave critical layer provides a paradigm for the phenomenon of Rossby-wave breaking, which is proving a valuable organizing concept in the search for understanding of realistic, and highly complicated, large-scale atmospheric flows (see McIntyre & Palmer 1985 and references). It is being appreciated that many such flows are highly

inhomogeneous, with regions in which the flow appears wave-like adjoining others in which it appears essentially nonlinear. Whilst there have been a number of successful attempts to simulate such flows in numerical models (see, for example, Held & Phillips 1987; Juckes & McIntyre 1987) nonlinear critical-layer theory, which applies to a limiting class of such flows, remains attractive because it includes a natural quantitative description of the dynamical interaction between the different regions.

The work of Stewartson (1978) and Warn & Warn (1978) has allowed considerable progress in our understanding of time-dependent, nonlinear, Rossby-wave critical layers for two reasons. First, they were able to apply the method of matched asymptotic expansions to the problem, treating the critical layer as an inner region. The equations of motion for the whole flow were then reduced to a single hyperbolic equation for the vorticity within the critical layer. This, by itself, was a considerable advance, since the evolution has a straightforward physical interpretation and, although the new evolution equation required numerical methods for its solution, numerical resources could be concentrated in the critical layer. There was no need to use numerical methods to solve throughout the much larger flow domain away from the critical layer as well, as had been the case in earlier investigations by Béland (1976, 1978), for example. Nonetheless, the solution of the evolution equation in the critical layer still represented a formidable numerical problem, and Warn & Warn (1978) found that it was possible to integrate only for a relatively short time before features in the vorticity field of scale comparable with the numerical resolution were generated. The second important development was therefore the discovery jointly by Stewartson (1978) and Warn & Warn (1978) of special cases permitting analytical solution of the critical-layer equations. The analytical solution (noted by Stewartson) provides a self-consistent description of the flow for a particular set of external parameter values (noted by Warn & Warn). The Stewartson–Warn–Warn solution (hereinafter termed the SWW solution) allowed, for those special cases, firm and reliable predictions about the time evolution of the critical layer, and was immune to any criticism of insufficient numerical resolution, unlike some of the earlier work on this problem.

One of the interesting predictions of the SWW solution is that the critical layer evolves from a state in which it absorbs incident Rossby waves, through a perfectly reflecting state, and into an over-reflecting state. The earlier numerical results of Geisler & Dickinson (1974) and Béland (1976, 1978) are in broad agreement with this, indicating that the numerical models used by these authors were at least adequate to reproduce this aspect of the behaviour. While Béland's model included up to seven zonal harmonics, that of Geisler & Dickinson included the zonal mean and only one harmonic. The success of the latter might therefore be considered surprising and a recent paper by Haynes & McIntyre (1987) discusses this point in some detail. The SWW solution further indicates that, as time increases, the critical layer oscillates from an over-reflecting state to an absorbing state and back again. These oscillations decrease in amplitude, so that, after a long time, the critical layer approaches a state of perfect reflection. (This decay is not reproduced by single-harmonic models.) At this stage the vorticity pattern in the critical layer is predicted (in the absence of dissipation) to have extremely fine spatial structure. For this reason such large times are completely beyond the range of validity of any numerical model which is effectively dissipationless, whether or not it exploits the matched asymptotic formalism.

The absorbing and reflecting properties of the critical layer may be understood in quite general terms, without reference to any particular solution for the flow, by

considering the behaviour of a certain conservable wave quantity A, that satisfies a conservation relation of the form

$$\frac{\partial A}{\partial t} + \nabla \cdot \boldsymbol{F} = S, \tag{1.1}$$

which is valid at finite amplitude. The expressions for the pseudomomentum A and the flux $\boldsymbol{F}$ are given by Killworth & McIntyre (1985) and the dissipative term S is given by Haynes (1988). Their explicit forms are not required for this discussion. The conservation relation (1.1) does allow us, within certain well-known limitations, to make precise such ideas as absorption and reflection. For instance, if the integral of the normal component of $\boldsymbol{F}$ over a surface bounding a region has the appropriate sign, then we may say that such a region is acting as a net absorber of waves. In the absence of dissipation and if a wave train is incident on the critical layer, it acts as an absorber initially for the simple reason that, as the waves arrive, the amount of wave activity in the vicinity of the critical line increases. In the presence of dissipation there is the possibility that the rate of dissipation of wave activity can subsequently balance the convergence of $\boldsymbol{F}$. Thus dissipative critical layers could in principle act as absorbers in the steady state.

On the other hand, if there is no overt dissipation, then long-term absorption can occur only if the amount of wave activity in the critical layer continues to increase. Of course, there is no *a priori* reason why this should not be the case, but, at least for the case of Rossby waves on a zonally symmetric flow, Killworth & McIntyre (1985) showed that the integral of the wave activity over a certain region may be bounded by a function of the width of that region and the vorticity field which pre-existed in it. They deduced that the time integral of the flux $\boldsymbol{F}$ into this region, which they called the time-integrated absorptivity, was finite. Moreover they showed that this conclusion was not changed by introducing dissipation in a certain, fairly wide class of cases, including some of geophysical interest.

It should be emphasized that the preceding discussion is independent of the details of the flow in the critical layer; the value of the SWW solution is that it provides an illuminating and specific example of the absorption–reflection behaviour and its dynamical mechanism.

Given the SWW solution as a benchmark case, it is natural to ask what aspects of critical-layer evolution it does *not* capture. This line of enquiry has recently led to the following two questions being posed. First, to what extent is the behaviour which it predicts representative of that of flows for which the external parameters do not take the special values required for the solution to be valid? Secondly, is the flow which it represents physically realizable?

Regarding the first point, it was made clear in Stewartson (1978) and Warn & Warn (1978) (hereinafter referred to as SWW) that the analytic solution was only valid for certain special configurations of the flow boundaries. Indeed, Warn & Warn integrated the critical-layer equations numerically in other cases and showed that the flow evolution could be considerably different (see their figure 4, for example). A larger range of cases has recently been investigated by Ritchie (1985), concentrating in particular on flow configurations for which the forced wave, or one of its higher harmonics, is close to resonance, unlike the SWW case. In some circumstances Ritchie found that the streamlines in the critical layer become considerably distorted, often forming multiple closed cat's eyes.

Regarding the second point, it has recently been shown by Killworth & McIntyre

(1985) that the flow described by the SWW solution is unstable. They show that the fastest-growing disturbances have much shorter wavelength than the forced wave and consequently grow very rapidly compared with the time development of the SWW flow. This scale separation in time and space allows the instability problem to be reduced, at each position along the critical layer, and at each time in its evolution, to the classical one of that for disturbances to a steady unidirectional flow. The appropriate flow to consider is that with velocity profile equal to the basic-state shear and absolute vorticity profile given by the SWW solution at the position and time of interest. On this basis Killworth & McIntyre calculated the growth rates of small-amplitude unstable disturbances and Haynes (1985) used numerical methods to follow the growth very accurately to finite amplitude, thereby demonstrating that the disturbances could drastically rearrange the absolute vorticity field within the critical layer and thus drastically change the absorption–reflection properties.

The objective of this paper is partly to assess the effect of the instability on the critical-layer flow as a whole, particularly regarding the absorption–reflection behaviour. 'Local' calculations of the type described in the preceding paragraph cannot make quantitative predictions about such effects, beyond the fact that they could be considerable. Instead a numerical model which can successfully simulate the evolution of the entire critical layer must be used. A model similar to those used by Warn & Warn (1978) and Ritchie (1985), which exploit the economy of the matched-asymptotics approach but allow the flow to have quite general structure in the streamwise direction, is appropriate. Indeed, the fact that the instability did not appear to arise in either of the latter investigations provides a strong case for further work to find out why this was so. The second objective of this paper is therefore to set the instability in the context of other critical-layer work, and to assess the parameter ranges in which it is likely to be dominant and those in which other effects such as resonance might be more important.

In §2 a brief description is given of the analysis that leads to a closed set of equations for the motion within the critical layer, which forms an inner region for the purposes of the matched-asymptotics formulation. In §3 the numerical method used for solving the critical-layer equations is set out, in some detail for those aspects which are non-standard. This leaves the way open for an account of the main series of numerical experiments in §4. Beland's (1976) results showed behaviour which, at first sight could have been the result of instability, or of resonance. His numerical experiments are re-examined in §5 in order to determine which of these was most likely to be important.

The emphasis in this paper is on numerical simulations which resolve the fluid-dynamical details of the flow in the critical layer. However, given the results of these simulations, it is interesting to compare them with results obtained from a model in which mixing by the unstable eddies is represented in a very simple way. One possible model is discussed in §6 and is applied to the flow predicted by the SWW solution. Concluding remarks are made in §7.

2. Mathematical formulation

2.1. *The critical-layer equations*

We shall consider the problem, originally studied by Stewartson (1978) and Warn & Warn (1978), of disturbances to a two-dimensional shear flow on a beta-plane. The geometry is described by Cartesian coordinates x, y, with the x-axis being aligned

parallel to the direction of the basic flow. The time evolution of the flow is taken to be described by the barotropic vorticity equation

$$\frac{\mathrm{D}q}{\mathrm{D}t} = 0, \tag{2.1}$$

where $\mathrm{D}/\mathrm{D}t = \partial/\partial t + u\,\partial/\partial x + v\,\partial/\partial y$ and denotes the rate of change following a fluid particle, with u and v being velocity components in the x- and y-directions respectively. The quantity q is the absolute vorticity, defined by

$$q = f_0 + \beta y + \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}, \tag{2.2}$$

where $f_0 + \beta y$ is the planetary vorticity, f_0 and β being constants, with β representing the gradient of planetary vorticity in the y-direction. It will be taken that the flow is incompressible, allowing the definition of a stream function, ψ, such that $u = -\partial\psi/\partial y$ and $v = \partial\psi/\partial x$. The absolute vorticity may therefore be written in the form

$$q = f_0 + \beta y + \nabla^2\psi. \tag{2.3}$$

This may be regarded as an equation for ψ in terms of q, which has a unique solution, given suitable boundary conditions, so that in the usual way we may speak of the ψ-field induced by a given q-field. The equations (2.1) and (2.3) are therefore, together with the boundary conditions, completely sufficient to describe the flow.

The basic state is taken to be a flow in the x-direction, $(u, v) = (\Lambda y, 0)$, with Λ constant. The quantity Λ/β is then a natural cross-stream lengthscale. The flow domain is unbounded in $y < 0$, but bounded at $y = y_b\,(> 0)$ by a rigid wall. The flow is disturbed from an x-independent state by spatially periodic corrugations in this wall, of small amplitude, $O(\epsilon\Lambda/\beta)$ and wavenumber k. The corrugations have the effect of imposing v, and therefore $\partial\psi/\partial x$, at $y = y_b$ and so force disturbances to the flow.

The presence of the small parameter ϵ allowed SWW to make use of asymptotic techniques in solving the evolution equations (2.1) and (2.3) and their method will be outlined below, since it leads to the reduced form of the problem to be investigated numerically in this paper.

It is first natural to write the total stream function as

$$\psi = -\tfrac{1}{2}\Lambda y^2 + \epsilon\phi, \tag{2.4}$$

so defining a disturbance stream function, ϕ.

It also proves convenient to redefine the variables ψ, ϕ, x, y and t in non-dimensional form, scaled respectively by the quantities Λ^3/β^2, Λ^3/β^2, k^{-1}, Λ/β and $\beta/k\Lambda^2$ respectively. The equation satisfied by ϕ may then be shown to be

$$\left(\frac{\partial}{\partial t} + y\frac{\partial}{\partial x}\right)\left(\frac{\partial^2\phi}{\partial y^2} + \mu^2\frac{\partial^2\phi}{\partial x^2}\right) + \frac{\partial\phi}{\partial x} + \epsilon J\left(\phi, \frac{\partial^2\phi}{\partial y^2} + \mu^2\frac{\partial^2\phi}{\partial x^2}\right) = 0, \tag{2.5}$$

where J is the Jacobian with respect to x and y, and the important dimensionless quantity

$$\mu = k\Lambda/\beta \tag{2.6}$$

is the ratio of the lengthscale of the forced wave in the x-direction to the cross-stream lengthscale Λ/β. In this paper μ will be taken to be formally $O(1)$. In addition ϕ is taken to satisfy the boundary condition

$$\phi \to 0 \quad \text{as } y \to \infty, \tag{2.7a}$$

and the condition

$$\phi = w(t)\cos x \tag{2.7b}$$

at the wall, where the function w describes the amplitude of the corrugations with time. Given an initial condition such as $\phi = 0$ everywhere before some time t_0, (2.5), together with the boundary conditions (2.7a, b), is sufficient to determine ϕ for all $t > t_0$.

This problem, with the function $w(t)$ taking the value 0 for all $t < 0$ and the value 1 for all $t > 0$, has been studied by a number of previous authors (Dickinson 1970; Warn & Warn 1976, 1978; Stewartson 1978; see also Killworth & McIntyre 1985). For $t = O(1)$ the leading-order solution is simply the solution of the linear equation obtained by neglecting the term multiplied by ϵ in (2.5). In physical terms this equation describes the evolution of the flow through the propagation of linear Rossby waves and the associated irrotational flow. Its solution was studied in the case $\mu \ll 1$ by Dickinson (1970) and reveals that away from the critical line at $y = 0$ the wave pattern settles down to a steady state, but that the motion remains strongly time dependent in a region about the critical line. The thickness of this region, and the characteristic lengthscale of vorticity variations within it, decrease as t^{-1}. Warn & Warn (1976) showed that this decrease in scale would, in the absence of dissipation, inevitably lead to the nonlinear terms neglected in (2.5) becoming as large as those terms retained when $t = O(\epsilon^{-\frac{1}{2}})$. The thickness of the region of time-dependent motion, would then be $O(\epsilon^{\frac{1}{2}})$. This region is referred to as the nonlinear critical layer. For $t \gtrsim \epsilon^{-\frac{1}{2}}$ it is a region of substantial vorticity rearrangement which can be regarded as an example of the breaking of Rossby waves.

Because $\epsilon \ll 1$ the problem contains two divisions of scale, which allowed SWW to apply asymptotic techniques and thereby simplify the problem considerably. One division is between the timescale for linear Rossby-wave propagation and the timescale for nonlinear effects to become important. The second is between the cross-stream lengthscale of the disturbances away from the critical layer and the cross-stream lengthscale within the critical layer.

The steps in the SWW formulation which reduce (2.5) to a closed set of equations for the motion within the critical layer will now be briefly described. The calculation is set out in more detail in SWW, Ritchie (1985) and Killworth & McIntyre (1985). Here we follow the notation used by Killworth & McIntyre (1985, §2) as far as possible, but cover the cases studied by Warn & Warn (1978) and Ritchie (1985) where account must be taken, at leading order, of the changing flow within the critical layer. It is worth recalling that, at leading order, this flow change is not induced by the vorticity anomalies in the critical layer itself, but by the vorticity in the quasi-steady Rossby wave field associated with such anomalies, which extend well away from the critical layer.

We follow SWW by defining a slow time variable

$$T = \epsilon^{\frac{1}{2}}t \tag{2.8}$$

which changes by $O(1)$ during the nonlinear evolution, and a stretched y-variable

$$Y = \epsilon^{-\frac{1}{2}}y \tag{2.9}$$

which changes by $O(1)$ within the critical layer.

Within the slow time formulation the leading-order stream function outside the critical layer may be written as

$$\psi(x,y,t) = -\tfrac{1}{2}y^2 + \epsilon\phi_1(x,y,T) + O(\epsilon^{\frac{3}{2}}\ln\epsilon). \tag{2.10}$$

The function $\phi_1(x,y,T)$ satisfies the steady linear equation for Rossby waves on a shear flow,

$$y\frac{\partial}{\partial x}\left\{\frac{\partial^2\phi_1}{\partial y^2} + \mu^2\frac{\partial^2\phi_1}{\partial x^2}\right\} + \frac{\partial\phi_1}{\partial x} = 0 \tag{2.11}$$

in the two regions $y > 0$ and $y < 0$. The general solution to this equation which is 2π-periodic in the x-direction and matches with the boundary condition (2.7*b*) may be written as a Fourier series in the form

$$\phi_1(x,y,T) = \mathrm{Re}\left\{\sum_{n=0}^{\infty}\phi_{1n}(y,T)\,\mathrm{e}^{\mathrm{i}nx}\right\}, \tag{2.12a}$$

where
$$\phi_{1n}(y,T) = A_n(T)f(\mu n,y) + B_n(T)\,g(\mu n,y) \quad (y>0) \tag{2.12b}$$

and
$$\phi_{1n}(y,T) = C_n(T)\,(\mu n,y) \quad (y<0), \tag{2.12c}$$

the functions f, g, and h being defined by

$$f(\mu n,y) = \left\{\Gamma\left(1-\frac{1}{2\mu n}\right)U\left(-\frac{1}{2\mu n},0,2\mu n y\right) - \pi\cot\left(\frac{\pi}{2\mu n}\right)M^*\left(-\frac{1}{2\mu n},0,2\mu n y\right)\right\}\mathrm{e}^{-\mu n y}, \tag{2.13a}$$

$$g(\mu n,y) = -M^*\left(-\frac{1}{2\mu n},0,2\mu n y\right)\mathrm{e}^{-\mu n y} \tag{2.13b}$$

and
$$h(\mu n,y) = \Gamma\left(1+\frac{1}{2\mu n}\right)U\left(\frac{1}{2\mu n},0,-2\mu n y\right)\mathrm{e}^{\mu n y}. \tag{2.13c}$$

The notation $M^*(a,b,z)$ is used here for the function $\{\Gamma(b)\}^{-1}M(a,b,z)$. $M(a,b,z)$ and $U(a,b,z)$ are respectively the first and second confluent hypergeometric functions, in the notation of Abramowitz & Stegun (1965). The function $M^*(a,b,z)$ is defined for all values of b, including non-positive integers. Near $y = 0$ the functions f, g, and h have the asymptotic forms

$$f(\mu n,y) = 1 - y\ln y + a_1 y + O(y^2\ln y), \tag{2.14a}$$

$$g(\mu n,y) = y + O(y^2) \tag{2.14b}$$

and
$$h(\mu n,y) = 1 - y\ln|y| + a_1 y + O(y^2\ln y). \tag{2.14c}$$

The constant a_1 is given by

$$a_1 = \nu(2) + \nu(1) - \nu\left(1+\frac{1}{2\mu n}\right) - \ln(2\mu n) + \mu n, \tag{2.15}$$

where $\nu(\cdot)$ (normally referred to as the psi-function) is the logarithmic derivative of the gamma function. Published tables of the functions $M^*(a,b,z)$ and $U(a,b,z)$ could not be found for the values of a and b required here. The functions were therefore evaluated using an algorithm based on that of Temme (1983).

It may be shown that, in the limit $\mu\to 0^+$, the functions $f(\mu n,y)$, $g(\mu n,y)$ and $h(\mu n,y)$ are identical to the f, g and h defined by Killworth & McIntyre (1985, equations 2.17*a–c*).

It is helpful, at this stage, to define the functions

$$A(x,T) = \mathrm{Re}\left(\sum_{n=0}^{\infty} A_n(T)\,\mathrm{e}^{\mathrm{i}nx}\right), \tag{2.16a}$$

$$B(x,T) = \mathrm{Re}\left(\sum_{n=0}^{\infty} B_n(T)\,\mathrm{e}^{\mathrm{i}nx}\right) \tag{2.16b}$$

and

$$C(x,T) = \mathrm{Re}\left(\sum_{n=0}^{\infty} C_n(T)\,\mathrm{e}^{\mathrm{i}nx}\right). \tag{2.16c}$$

in terms of the Fourier coefficients introduced in (2.12*b*, *c*).

In the inner region the stream function ψ is considered to be a function of x, Y and T. The evolution equation in this region may be derived by writing down an expansion for ψ based on the asymptotic sequence $\{\epsilon, \epsilon^{\frac{3}{2}}\ln\epsilon, \epsilon^{\frac{3}{2}}, \ldots\}$. The members of this sequence arise naturally if $\epsilon^{\frac{1}{2}}Y$ is substituted for y in the expression (2.12*a*), taking account of the asymptotic forms of the functions f, g and h. Further consideration of the matching requirements on ψ between inner and outer regions shows that the leading two terms in the expression for the zonally varying part of ψ must be constant across the critical layer. This leads immediately to the result that

$$A(x,T) = C(x,T). \tag{2.17}$$

The stream function in the critical layer may now be written in the form

$$\psi(x,Y,T) = \epsilon\{-\tfrac{1}{2}Y^2 + C(x,T)(1-\tfrac{1}{2}Y\epsilon^{\frac{1}{2}}\ln\epsilon) + \epsilon^{\frac{1}{2}}\Psi_1(x,Y,T) + O(\epsilon\ln\epsilon)\}. \tag{2.18}$$

It follows that the leading-order relative vorticity in the critical layer is $O(\epsilon^{\frac{1}{2}})$.

Substitution of (2.18) into the governing equation (2.5) leads to

$$\left\{\frac{\partial}{\partial T} + \frac{\partial C}{\partial x}\frac{\partial}{\partial Y} + Y\frac{\partial}{\partial x}\right\}\{Y + Q_1\} = 0, \tag{2.19}$$

where

$$Q_1(x,Y,T) = \frac{\partial^2 \Psi_1}{\partial Y^2} \tag{2.20}$$

is the leading-order vorticity within the critical layer. This equation describes, at leading order, the advection of the absolute vorticity around streamlines which are contours of the function

$$\Psi_0(x,Y,T) = -\tfrac{1}{2}Y^2 + C(x,T). \tag{2.21}$$

In order to close the problem it is necessary to find one more equation relating $A(x, T)$, $B(x,T)$, $C(x,T)$ and $Q_1(x,T)$. This equation may be obtained from the matching condition on ψ_y across the critical layer. It is found that the matching condition may be written in the form

$$B(x,T) = \int_{-\infty}^{\infty} Q_1\,\mathrm{d}Y, \tag{2.22}$$

where the Cauchy principal value of integral is taken, and relates the vorticity field in the critical layer to the jump across it in the x-component of the velocity.

If the expression (2.12) is now substituted into the boundary condition (2.7*b*), and

the conditions (2.13*a*), (2.17) and (2.22) used, we are left with an equation for each Fourier coefficient of the form

$$f(\mu n, y_b)\, C_n(T) + g(\mu n, y_b) \int_{-\infty}^{\infty} Q_{1n}\, \mathrm{d}Y = W_n(T) = \delta_{n1}\, W(T), \tag{2.23}$$

where $\delta_{n1} = 1$ for $n = 1$ and 0 otherwise. If it is assumed that the forcing is switched on over the outer timescale then the function $W(T)$ is defined by $W(T) = w(\infty)$ and the Fourier coefficients Q_{1n} by

$$Q_1(x, Y, T) = \mathrm{Re}\left\{ \sum_{n=0}^{\infty} Q_{1n}(Y, T)\, \mathrm{e}^{\mathrm{i}nx} \right\}.$$

The leading-order vorticity equation (2.19) and (2.23) together form a closed system and may, in principle, be integrated forward in time from a suitable initial condition. Equation (2.23), when divided by $f(\mu n, y_b)$ to give

$$C_n(T) = \rho_n\, W_n(T) + \sigma_n \int_{-\infty}^{\infty} Q_{1n}\, \mathrm{d}Y, \tag{2.24}$$

where

$$\rho_n = \frac{1}{f(\mu n, y_b)}, \quad \sigma_n = -\frac{g(\mu n, y_b)}{f(\mu n, y_b)}, \tag{2.25a, b}$$

may be regarded as an explicit inversion of the Q-field in the critical layer to give the stream function in the critical layer. It should be re-emphasized, however, that the corresponding flow field is induced not by the vorticity in the critical layer but by the vorticity associated with the quasi-steady wave pattern in the outer flow. The SWW solution is valid in the case $\mu = 0$, $g(0, y_b) = 0$ and with $W(T)$ constant.

2.2. *The case* $|\rho_n| \to \infty, |\sigma_n| \to \infty$

Note that there is a possibility that $f(\mu n, y_b)$ is zero for some n, in which case the quantities ρ_n and σ_n are infinite and some care is needed in interpreting the condition (2.24). The set of pairs μ and y_b for which this occurs is extremely small (of measure zero). Nevertheless, it is worth considering this case because it gives insight into the behaviour when $f(\mu n, y_b)$ is small and ρ_n and σ_n large for some value of n, which is important when solving (2.19) and (2.24) numerically. The minimum value of $f(\mu n, y_b)$ is especially small when μ is small, since the values of the sequence μn; $n = 0, 1, 2, \ldots$ are relatively densely spaced. As has already been indicated, this regime is of considerable interest.

In the special case where $f(\mu n, y_b) = 0$ it is best to return to the original form of the equation, being (2.23). Under these circumstances $C_n(T)$ disappears altogether from this equation, leaving the condition

$$B_n(T) = \int_{-\infty}^{\infty} Q_{1n}\, \mathrm{d}Y = \frac{\delta_{n1}\, W(T)}{g(\mu n, y_b)}. \tag{2.26}$$

It is clearly no longer possible to use this to calculate $C_n(T)$ directly, given $Q_{1n}(Y, T)$. Instead (2.26) sets a constraint on the evolution of the vorticity field and it is this constraint which determines C_n, for at each instant there is only one velocity field which advects the Q-field in such a way that the condition (2.26) remains satisfied.

As discussed by Ritchie (1984, 1985), a configuration in which $f(\mu n, y_b)$ is zero may be regarded as corresponding to a resonance in the sense that, provided the Fourier coefficient, B_1, of the velocity jump is zero, then there is a solution of the steady linear equation for ϕ_{1n} which tends to zero as $y \to -\infty$ and is also zero at $y = y_b$, and which

therefore requires no forcing. Ritchie (1984) showed by numerical experiment, directly simulating the entire flow without exploiting the matched-asymptotics formalism, that if a forcing is applied in such a configuration then the evolution of the critical layer towards a reflecting state is accompanied by growth of the forced wave, and a corresponding increase in the thickness of the critical layer. Within the matched-asymptotics theory, if the configuration is exactly resonant for the forced wave, the width of the cat's eyes, for example, grows without bound. This is how the system maintains a non-zero value of B_1 when the left-hand side of (2.26) is forced to be non-zero: physically the expanding cat's eyes keep entraining fresh q-contours and twisting them into configurations whose early stages are in the right phase to make $B_1 = 0$. Thus, although the system being considered is complicated by the nature of the linear wave structure near the critical line, it exhibits some characteristics associated with simpler resonant behaviour.

One question of considerable interest concerns the behaviour when the system is resonant for harmonics higher than that of the wave being directly forced at the wall. Ritchie (1985) investigated cases where these harmonics were close to, but not exactly at, resonance. These higher harmonics are inevitably excited through the nonlinear dynamics in the critical layer. Does the resonant harmonic then grow without bound (at least within the matched-asymptotics formalism)? The answer is almost certainly no, as may be seen by considering the equations for the quantities Q_{1r} and C_r, where r is the number corresponding to the resonant harmonic. These equations may be obtained by taking the appropriate Fourier coefficient of (2.19) and (2.24), and are found to be

$$\frac{\partial Q_{1r}}{\partial T} + \mathrm{i}rYQ_{1r} + \mathrm{i}rC_r = \mathrm{i}N_r \tag{2.27}$$

and

$$\int_{-\infty}^{\infty} Q_{1r}\,\mathrm{d}Y = \frac{1}{\sigma_r} C_r, \tag{2.28}$$

where N_r is defined by

$$-\frac{\partial C}{\partial x}\frac{\partial Q_1}{\partial Y} = \mathrm{Re}\left\{\sum_{n=0}^{\infty} \mathrm{i}N_n \mathrm{e}^{\mathrm{i}nx}\right\}. \tag{2.29}$$

The right-hand side of (2.27) therefore represents the forcing of the vorticity in harmonic r by the nonlinear terms in (2.19). Suppose we regard this forcing as a given function of time. Then (2.27) and (2.28) are linear in the quantities Q_{1r} and C_r, and in principle may be solved, for instance by taking the Laplace transform of (2.27) with respect to time, and then using this transformed equation to eliminate Q_{1r} from the Laplace transform of (2.28). This leaves an expression for the Laplace transform of C_{1r} which is perfectly well defined in the limit $\sigma_r \to \infty$. Once C_r has been evaluated then the result may be substituted back into (2.27) and the resulting equation solved for Q_{1r}. The important point is that the final expression for Q_{1r} exhibits no degenerate behaviour when σ_r becomes infinite, suggesting strongly that Q_{1r} does not become unbounded, contrasting with the case where the resonant harmonic is directly forced at the wall. Again, this seems consistent with the physical picture that if a particular harmonic grew without bound, the continual entrainment of new q-contours by the growing cat's eyes would, for that harmonic, lead to a non-zero value of the integral appearing in (2.26).

2.3. *The time-integrated absorptivity and the effect of artificial dissipation*

For future reference it is helpful to apply the argument presented in Killworth & McIntyre (1985, §1.4) directly to the case being discussed here (in which the critical layer is very thin). In the small-amplitude limit appropriate in the outer flow, the flux $\boldsymbol{F}$ appearing in (1.1) has components $(\frac{1}{2}\overline{(v'^2-u'^2)}, -\overline{u'v'})$, where the prime denotes the difference from the x-average, denoted by an overbar. It follows that the mean flux of wave activity into the critical layer is just equal to the jump in the Reynolds stress across the critical layer, $[\overline{u'v'}]$, where the limits are taken in the outer flow. This quantity will be referred to as the absorptivity. Matching with the inner region, and substitution of the Fourier series expressions for C and Q_1 gives the result that

$$[\overline{u'v'}]_{0^-}^{0^+} = -\int_{-\infty}^{\infty} \overline{Q_1 \frac{\partial C}{\partial x}}\,\mathrm{d}Y = -\mathrm{Im}\left\{\sum_{n=0}^{\infty} nB_n C_n^*\right\}, \tag{2.30}$$

and gives an equation for the absorptivity in terms of quantities known in the inner region. Note from (2.24) that the only contribution to the right-hand side comes from harmonics which are directly forced at the wall (in this case $n = 1$).

We now derive an alternative expression for the absorptivity, but first allow the addition of a dissipative term D to the right-hand side of (2.19). This term might be a suitably scaled representation of a physically motivated dissipation, e.g. viscous diffusion, in which case it takes the form $\lambda\partial^2 Q/\partial Y^2$, where λ, a suitably scaled viscosity, is Haberman's parameter (Haberman 1972), or else a purely artificial dissipative term introduced for computational purposes. In the numerical integrations such a term is found to be essential in order to inhibit the cascade of vorticity to scales which are at the limit of the numerical resolution of the model.

Taking the x-average (denoted by an overbar) of (2.19) with the dissipative term added leads to

$$\frac{\partial \bar{Q}_1}{\partial t} = -\overline{\frac{\partial C}{\partial x}}\,\overline{\frac{\partial Q_1}{\partial Y}} + \bar{D}. \tag{2.31}$$

Provided that $\bar{D}$ decays sufficiently rapidly as $|Y| \to \infty$, this equation is multiplied by Y, and then integrated with respect to Y from $Y = -\infty$ to $Y = \infty$. The first term on the right-hand side may then be integrated by parts and the boundary terms estimated to be zero using the asymptotic form of Q_1 for large Y, to leave, after using (2.30),

$$[\overline{u'v'}]_{0^-}^{0^+} = \int_{-\infty}^{\infty} \overline{Q_1 \frac{\partial C}{\partial x}}\,\mathrm{d}Y = \frac{\partial}{\partial T}\int_{-\infty}^{\infty} Y\bar{Q}_1\,\mathrm{d}Y - \int_{-\infty}^{\infty} Y\bar{D}\,\mathrm{d}Y. \tag{2.32}$$

The time-integrated absorptivity, which we shall refer to as $\alpha(T)$, following Killworth & McIntyre (1985), may therefore be written as

$$\alpha(T) = \int_0^T \mathrm{d}T'[\overline{u'v'}]_{0^-}^{0^+} = \int_{-\infty}^{\infty} Y\,\mathrm{d}Y[\bar{Q}_1]_0^T - \int_0^T \mathrm{d}T' \int_{-\infty}^{\infty} \mathrm{d}Y Y\bar{D}. \tag{2.33}$$

It follows that if $D = 0$ then, provided that the integral in the first term on the right-hand side of (2.33) remains bounded, $\alpha(T)$ remains bounded for all T. This is an alternative derivation of Killworth & McIntyre's (1985) result, which is allowed in the small-amplitude case because the behaviour of the vorticity field as $|Y| \to \infty$ is rather tightly constrained by (2.19). As may be seen from (2.32), $\alpha(T)$ may still be bounded even if D is not zero, providing that the zonal-mean D is zero. In order to ensure that a direct relation between the distribution of $\bar{Q}$ and the time-integrated

absorptivity remained true in the numerical experiments, the artificial dissipation, when it was used, was chosen to be such that its zonal-mean part vanished. The other condition required of the dissipation was that it should act only on the smallest scales resolved in the model. The form of D was therefore chosen to be

$$D = -\delta(\partial/\partial x)^8 Q_1, \tag{2.34}$$

with δ a constant. Equation (2.33) may also be interpreted as a statement about the x-average momentum, relating the time-integrated force on the fluid in the critical layer to the rate of change of Kelvin's impulse (and therefore to the rearrangement of vorticity within the critical layer). The form of D chosen in (2.34) is such that no net force is exerted on the critical layer. The form of the dissipation means that it is inappropriate to define Haberman's parameter. However the value of δ was taken to be $7.5N^{-7}$, where N was the highest non-zero Fourier coefficient retained in the calculation. Thus, broadly speaking, for most of the x-scales in the model, the dissipation was weak, but it acted strongly at the shortest scales.

3. Numerical solution of the critical-layer equations

Systems of equations similar to (2.19) and (2.24) have been integrated by Warn & Warn (1978), Ritchie (1985) and Haynes (1985). Given the requirement of periodicity in x, it is clearly simplest to solve the corresponding equations for the Fourier coefficients, after truncating the Fourier series at a finite number of terms, N say. Use may be made of the technique, now standard, of evaluating nonlinear terms by transforming to a grid, multiplying, and then calculating the Fourier coefficients of the product. The equations for the coefficients must then be integrated in time, and, in addition, the Y-integral appearing in (2.24) must be evaluated. In principle the time integration may be based on evaluating the new Q_1 field at each time step using the old stream function, perhaps through the Adams–Bashforth or leapfrog methods, and then evaluating the integral in (2.24) to give the C_n and hence the new stream function. In fact a considerable improvement on this is to use an implicit time-stepping scheme for the terms representing advection by the basic shear, which is very strong near the edge of the computational domain. Such a semi-implicit scheme was used by Ritchie (1985) and Haynes (1985, see §3 and Appendix B). However the evaluation of terms representing the advection by the cross-stream velocity, which may vary in time, is then still based entirely on information from previous time-steps.

The primary difficulty that arose during this investigation was the requirement that the numerical scheme remain accurate for all values of σ_n, however large. For reasons discussed in §2, the equations are expected to have a perfectly well-behaved solution, even when $\sigma_n, \rho_n \to \infty$, for some harmonic n, providing that harmonic is not externally forced. However it turns out that, if σ_n is large for some n and the equations are integrated using the semi-implicit scheme outlined earlier, violent numerical instability results. It is not difficult to see why this is so; if $\int Q_{1n}\,\mathrm{d}Y$ is non-zero at some time then the calculated value of C_n used for the next time step will be very large. The adjustment process by which $\int Q_{1n}\,\mathrm{d}Y$ remains small enough for C_n to be moderate in size is a very delicate one and, for typical values of σ_n which arise, can be achieved in the numerical integration only by taking time steps which are prohibitively small. This problem was avoided by using a different numerical scheme in which the time stepping of Q_1 and the evaluation of the integral were more closely linked. Details of this new scheme are now given.

We write the Fourier transform of (2.19) in time-differenced form as

$$\frac{Q_{1n}^{T+\Delta T}-Q_{1n}^{T}}{\Delta T}+\tfrac{1}{2}\mathrm{i}nY(Q_{1n}^{T+\Delta T}+Q_{1n}^{T})+\tfrac{1}{2}\mathrm{i}n\left(1+\frac{\partial Q_{10}^{T+\Delta T}}{\partial Y}\right)C_n^{T+\Delta T}$$
$$+\tfrac{1}{2}\mathrm{i}n\left(1+\frac{\partial Q_{10}^{T}}{\partial Y}\right)C_n^{T}=\tfrac{1}{2}(3P_n^{T}-P_n^{T-\Delta T})\quad(1\leqslant n\leqslant N),\quad(3.1)$$

where the superscript T denotes the value of a quantity at time T and P_n is defined by

$$\mathrm{Re}\left\{\sum_{n=1}^{\infty}P_n\,\mathrm{e}^{\mathrm{i}nx}\right\}=\frac{\partial C}{\partial x}\left\{\frac{\partial Q_1}{\partial Y}-\frac{\partial \bar{Q}_1}{\partial Y}\right\}.$$

For $n=0$ we may simply integrate using an Adams–Bashforth scheme, to give $Q_{10}^{T+\Delta T}$. To solve (3.1) for $Q_{1n}^{T+\Delta T}$ we also need the relation (2.24), written in the form

$$C_n^{T+\Delta T}=\rho_n\,\delta_{n1}\,W(T+\Delta T)+\sigma_n\int_{-\infty}^{\infty}Q_{1n}^{T+\Delta T}\,\mathrm{d}Y.\quad(3.2)$$

Note that the terms on the left-hand side of (3.1) are evaluated as in the trapezoidal scheme and the nonlinear terms on the right-hand side are evaluated as in the Adams–Bashforth scheme. The difference between this scheme and that used in Haynes (1985) (and, apparently, that in Ritchie 1985) is that the last term on the left-hand side is evaluated as in the trapezoidal scheme rather than as in the Adams–Bashforth scheme. It is this difference that makes the scheme stable even when one or more of the constants σ_n is very large, or indeed, infinite.

The two equations (3.1) and (3.2) may be solved for $C_n^{T+\Delta T}$, by rearranging (3.1) to give an expression for $Q_{1n}^{T+\Delta T}$, and then integrating this expression with respect to Y from $-\infty$ to ∞. Substituting this expression into (3.2) gives a linear equation for $C_n^{T+\Delta T}$, and hence that

$$C_n^{T+\Delta T}=\left(\frac{1}{\sigma_n}+\int_{-\infty}^{\infty}\left\{1+\frac{\partial Q_{10}^{T+\Delta T}}{\partial Y}\right\}\frac{\mathrm{d}Y}{Y-2\mathrm{i}/(n\Delta T)}\right)^{-1}\left\{\frac{\rho_n\,\delta_{n1}\,W(T)}{\sigma_n}-\int_{-\infty}^{\infty}\{Y+2\mathrm{i}/(n\Delta T)\}\right.$$
$$\left.\times\left[Q_{1n}^{T}+C_n^{T}\left(1+\frac{\partial Q_{10}^{T}}{\partial Y}\right)+\{P_n^{T-\Delta T}-3P_n^{T}\}/(n\Delta T)\right]\frac{\mathrm{d}Y}{Y-2\mathrm{i}/(n\Delta T)}\right\}.\quad(3.3)$$

This expression is well defined in the limit $\sigma_n\to\infty$. (Note that the ratio ρ_n/σ_n remains finite in that limit.) An expression for $Q_{1n}^{T+\Delta T}$ may now be obtained by substituting (3.3) into (3.1).

In the computation the integrals and Y-derivatives appearing in (3.3) were approximated by finite-difference expressions. The integrals were evaluated between the limits of the computational domain, which were at $Y=-Y_{\max}$ and $Y=Y_{\max}$, where $Y_{\max}$ was fixed at various values greater than or equal to 5. It was therefore necessary to make some estimate of the contribution to the integral from outside the domain, which is particularly important when a disturbance is caused by a forcing that is switched on rapidly. A large part of the response then consists of sheared disturbances spread over a wide region of the flow, and these give a contribution to the Y-integral from outside $Y_{\max}$ which is initially substantial, although it subsequently decays as $1/T$. The contribution to the integral from disturbances which result from a smoothly growing forcing, on the other hand, decay with $Y_{\max}$ at all times. Haynes (1985) described a scheme by which this contribution could be estimated correct to $O(Y_{\max}^{-3})$. This scheme can be modified to deal with a flow made up of some

disturbances which are forced suddenly (as was expected to be the case in the flow simulations to be described later). However the $O(Y_{\max}^{-3})$ scheme could not be combined with the semi-implicit time-stepping scheme just described, nor could one which was accurate to $O(Y_{\max}^{-2})$ not least because the sequence of approximations described in Haynes (1985, Appendix) does not remain valid when σ_n is very large. It was therefore necessary to estimate the infinite integrals by ignoring the contribution from outside the computational domain, except in the case of harmonics directly forced at the wall (i.e. $n = 1$ in this case), where the contribution may be estimated in terms of a sine integral and the initial forcing amplitude (Warn & Warn 1978). This approximation is relatively inaccurate, with errors $O(Y_{\max}^{-1})$. However, tests with different domain sizes suggest that choosing $Y_{\max} = 7.5$ led to small quantitative inaccuracy (less than 5% in $B_1(T)$ at $T = 10$, for example), and no qualitative change, in the results of the numerical experiments. Typically, 151 grid points were used to cover this domain. Again, this number was chosen on the basis of the results of a series of experiments in which the resolution in the y-direction was varied.

4. Numerical experiments: I

In this section the results of some experiments in flow configurations rather close to that for which the SWW solution is valid will be presented. The configurations chosen are close in the sense that the wall position y_b is taken to be 3.67, the smallest value such that the SWW solution is valid in the limit $\mu \to 0$, i.e. such that $g(0, y_b) = 0$. In practice this ensures that σ_1 is small for all $\mu < 0.5$, so that changes in the primary harmonic component of the velocity jump do not greatly affect the streamline pattern in the critical layer. However, if μ is taken to be greater than zero the coefficients σ_n are not necessarily small for all $n > 1$. This allows the possibility of higher harmonics changing the streamline pattern either through unstable growth or simply through their being forced by nonlinear interactions originating in the primary wave. In each of the simulations reported the truncation in the zonal direction was taken at $N = 31$ (for $\mu > 0.2$) or else at $N = 63$ (giving 63 or 127 degrees of freedom with which to represent the x-variation).

We shall first consider a case where μ is small and which therefore lies within the parameter range for which the Killworth & McIntyre instability analysis should be valid, at least qualitatively. The first experiment (experiment A) is in every way identical to the SWW case, except that μ is taken to be 0.1 rather than 0. Figure 1 shows a series of pictures of the absolute vorticity distribution within the critical layer at times $T = 3.2$, 3.4, and 3.5. For comparison, the SWW absolute vorticity distribution is shown in the same manner in figure 2 for $T = 3.5$. Note that the absolute vorticity field is shown only in the region $-5 < Y < 5$ although the computational domain was taken to be $-7.5 < Y < 7.5$. The Fourier series were truncated at $N = 31$. There is a spectacular difference between figures 1 (*c*) and 2. The nature of the flow in the former case, from which it may be deduced that the absolute vorticity field is being advected around secondary cat's eyes of much shorter wavelength than that of the forced wave, is consistent with that expected if the instability described by Killworth & McIntyre (1985) and Haynes (1985) occurs. Figure 2 of Haynes (1985) shows the effect of the instability at one location along the primary cat's eye. To be certain that this instability is indeed important in this case it is necessary to make quantitative comparison between the results of the numerical experiments and the theory.

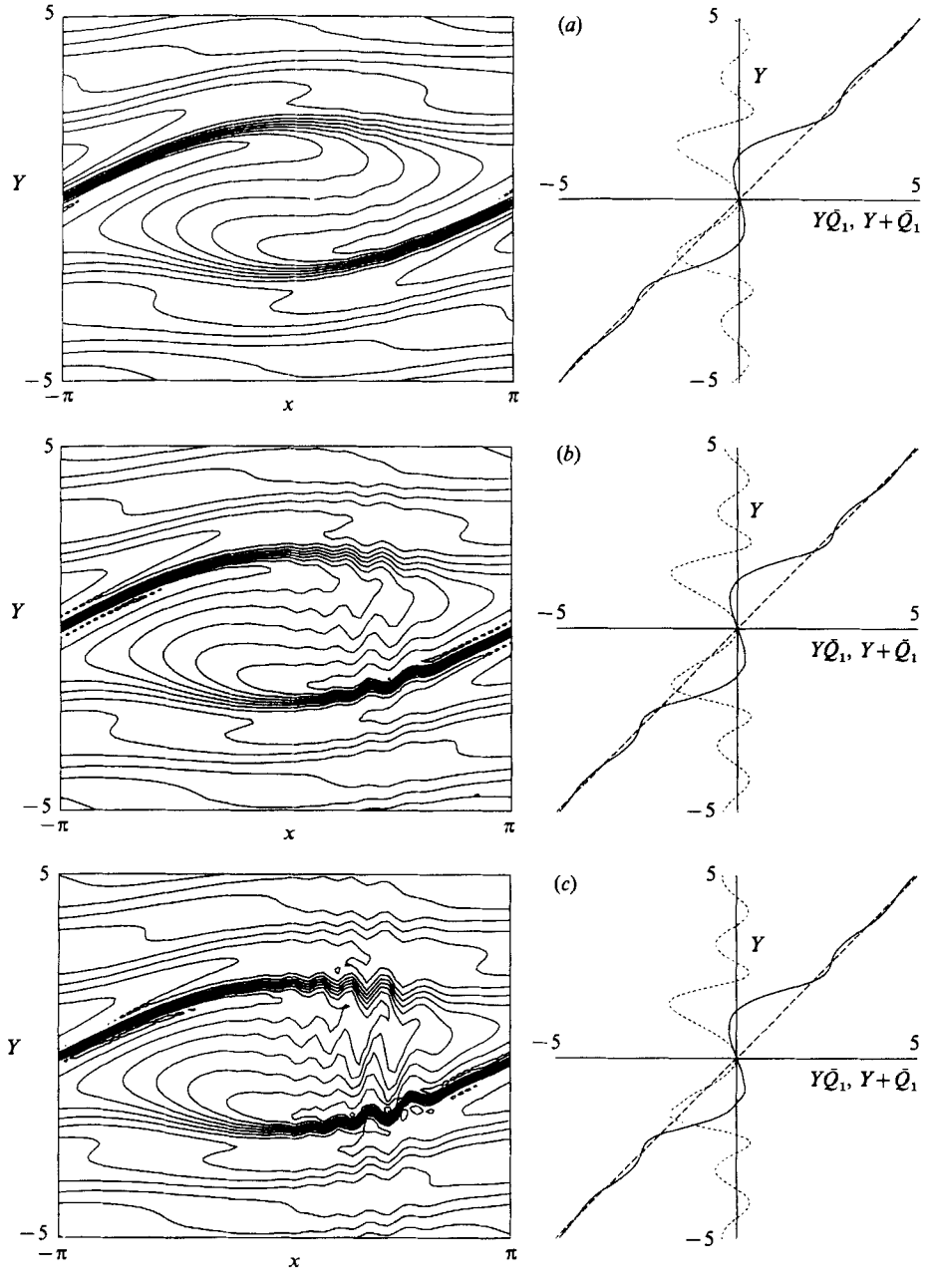

FIGURE 1. Contours of absolute vorticity, Q, in the critical layer at (*a*) $T = 3.2$, (*b*) $T = 3.4$ and (*c*) $T = 3.5$ during experiment A, for which $\mu = 0.1$. The contour interval is 0.5. On the right are graphs of the x-average of Q (solid line), the x-average of the quantity $Y(Q-Y)$ (dotted line), the time derivative of which appears in (2.32), and the basic-state absolute-vorticity profile (dashed line) which was present for $T < 0$.

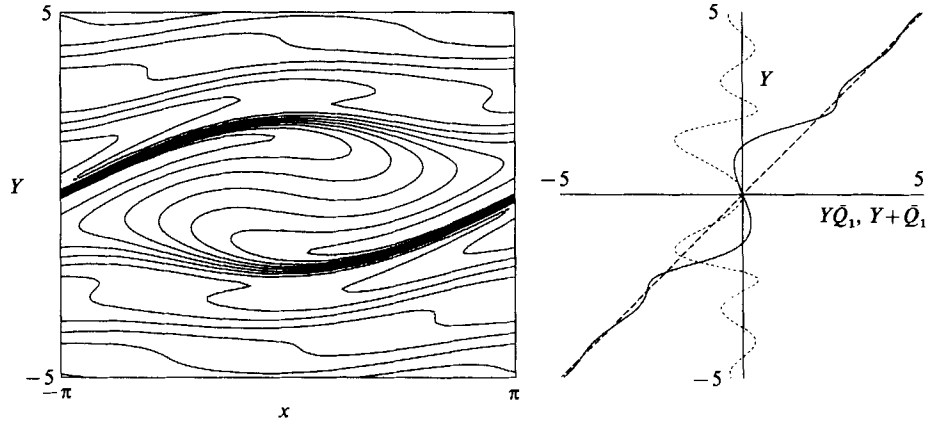

FIGURE 2. As figure 1, for the absolute vorticity field given by the SWW solution at $T = 3.5$.

First we may estimate the growth rate of the disturbances. This is not completely straightforward – the disturbance does not comprise a single harmonic in x and has a strongly variable structure along the cat's eye. However, at the times in question, the disturbance has a much larger amplitude at one location along the cat's eye than at any other. We might therefore expect the dominant scale of the disturbance at that location, providing that it is sufficiently well-separated from the wavelength of the forced wave, and from the lengthscale on which the amplitude 'envelope' varies, to correspond to a distinct peak in the spectrum of the vorticity or the stream function (even though this spectrum is calculated from the functional form over the entire length of the cat's eye). Furthermore, the amplitude of that peak should also be roughly proportional to the amplitude of the disturbance at the location where it is largest.

In fact the peak in the spectrum seems to be shifted some way towards lower wavenumbers, as may be seen from figure 3(*a*), which shows the spectrum of enstrophy (squared relative vorticity) at various times from $T = 1.6$ to 3.5, with alternate times (at an interval of 0.1) being shown as solid and dashed curves. For comparison the graphs of the enstrophy spectra at $T = 3.0$ and 3.5 are repeated in figure 3(*b*) as solid lines with the corresponding results from the $\mu = 0$ case superimposed as dashed lines. The strongest peak in the spectrum is at around wave 10, whereas from figure 1(*c*) it appears that the local wavenumber of the disturbance is about 14 times that of the primary wave. Both the peak in the spectrum, and the amplitude of the short-wavelength disturbances as estimated graphically from figure 1, for example, grow from $T = 3.3$ to $T = 3.5$ at a rate corresponding roughly to doubling in a time unit of 0.1 (an e-folding rate of 0.69). The local instability analysis for this time gives the fastest growing wavenumber at $x = \frac{1}{2}\pi$ as 1.12 and the corresponding growth rate as 0.59. (For $x = \frac{1}{2}\pi$ and $x = \frac{3}{4}\pi$ the relevant wavenumbers are 1.47 and 0.97 and the growth rates 0.91 and 0.31 respectively.) The disturbances seen here therefore appear to have characteristics which are well within the bounds of those predicted by the local instability theory.

As discussed in §1, it is also possible that multiple cat's eyes might arise as a predictable and deterministic part of the evolution of the system, as seemed to be the case in Ritchie's (1985) experiments. A clear indication that for $\mu = 0.1$ the basic cat's eye flow is *unstable* is given by the results of a second experiment (B) in which 'noise' was added to the system at $T = 1.5$, in harmonics 7–16. The amplitude of the

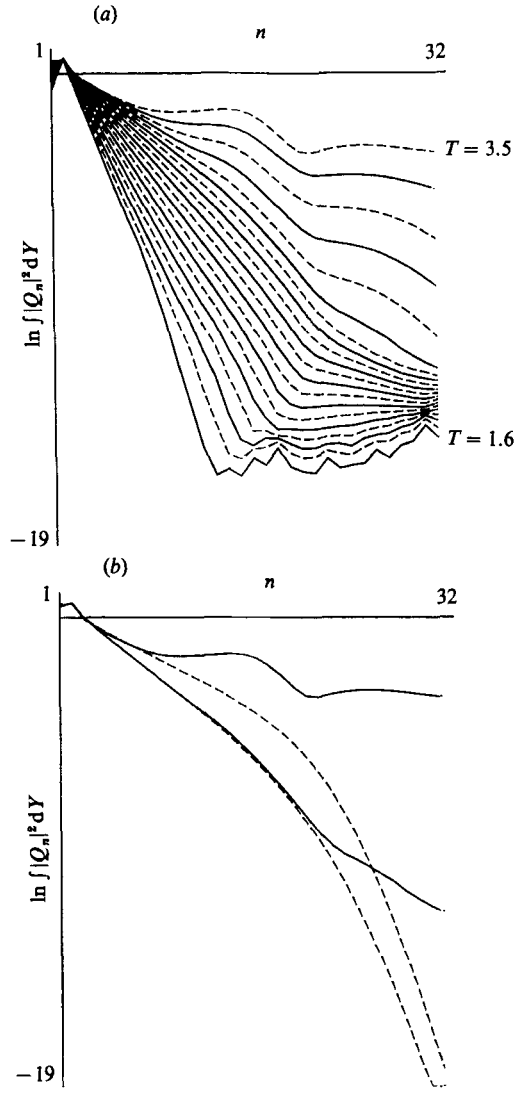

FIGURE 3. Spectra of enstrophy (squared relative vorticity) integrated across the critical layer, plotted against zonal wavenumber. (*a*) Experiment A, $T = 1.6$ to $T = 3.5$; (*b*) experiment A, $T = 3.0$ and $T = 3.5$ (solid lines), experiment with $\mu = 0$, $T = 3.0$ and $T = 3.5$ (dashed lines).

noise is such that there are perceptible, but small, differences between the vorticity distribution for this experiment at $T = 2$ shown in figure 4(*a*), and that in the case with no noise. The maximum value of the vorticity associated with the noise was less than 0.1. Well before $T = 3$, these differences grow and the associated disturbances to the flow saturate, so that at $T = 3$ the vorticity patterns in the two cases differ drastically, as may be seen by comparing figure 4(*b*), for the case with noise added, and figure 4(*c*), for the case without noise.

It is also interesting to note from figure 1(*c*), for example, that the flow in the regions $-\pi < x < 0$ and $\frac{1}{2}\pi < x < \pi$ is almost completely undisturbed from that in figure 2. The reason is the relatively long time that it takes for information to be communicated along the length of the primary cat's eye as compared with that for

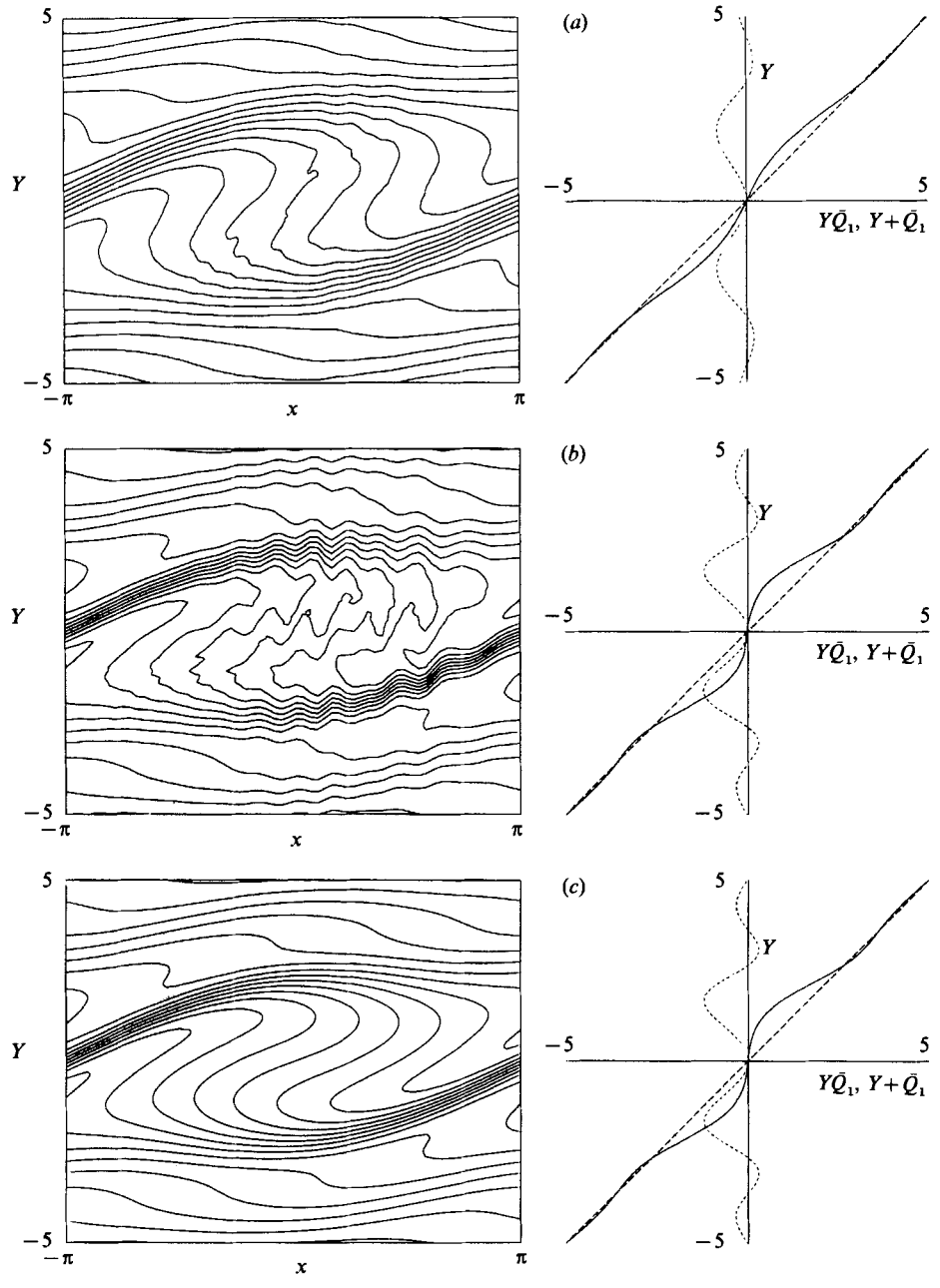

FIGURE 4. As figure 1, for experiment B in which, at $T = 1.5$, the vorticity field was slightly perturbed from that in experiment A. The absolute vorticity field is here shown for (*a*) $T = 2.0$ and (*b*) $T = 2.5$. The unperturbed flow in experiment A at $T = 2.5$ is shown in (*c*).

the growth of the secondary cat's eyes. Of course, the idea that, on the e-folding timescale for unstable disturbances, different locations along the cat's eye are relatively independent was the whole basis of the local analysis performed by Killworth & McIntyre (1985) and Haynes (1985). The local growth rates are actually largest in the interval $-\frac{1}{4}\pi < x < 0$. However the group velocities of the growing

modes, i.e. the rate of translation of a packet of disturbances, equal to $\mathrm{d}(kc_r)/\mathrm{d}k$, where k is the wavenumber in the x-direction and c_r is the real part of the complex phase speed, are positive (verified directly during the eigenvalue calculation). A packet of disturbances has therefore been growing for a longer time when it is to the right of a region in which the unstable growth rates are large and we might expect the largest amplitudes to occur there, provided that the instability is a purely convective one. A search was made for modes which were absolutely unstable (i.e. which grow in place, see e.g. Drazin & Reid 1981, pp. 151–153), but none were found.

It is clear from figures 1, 2 and 4 that the instability makes a considerable difference to the vorticity distribution within the critical layer. This in turn raises the question of the effect of the instability on the absorption/reflection properties which are felt well away from the critical layer. In the examples just discussed, where the evolution was followed until the instability had reached finite amplitude, there was little detectable difference in the graph of B_1 against T, for example. This is not unexpected; the effect of the instability in the case of small μ is expected to be rearrangement of the vorticity over a region of small width in the x-direction. The integral (2.22) should therefore be largely unchanged immediately after the first growth of the disturbances to finite amplitude. It is expected to change on a slower timescale as advection communicates the change in the vorticity field to other locations along the cat's eye.

In order to follow the evolution further it was necessary to introduce artificial dissipation, chosen to be of the form (2.34). The experiments A and B reported earlier, with $\mu = 0.1$, were repeated with this dissipation included and with higher resolution in the x-direction, and run out to $T = 10$. Figure 5(*a*) shows a typical absolute vorticity field during the evolution, in this case in experiment A at $T = 5.5$.

The graphs in figure 5(*b*) show the variation of $B_1(T)$ and $\alpha(T)$ in experiments A and B. For comparison the variation of $B_1(T)$ according to the SWW solution is also shown. As may be seen the effect of the instability is apparent in $B_1(T)$ at about $T = 4.5$ in experiment A and at about $T = 3$ in experiment B. (Recall that the vorticity field was considerably disrupted at about $T = 3.5$, as shown in figure 1*c*.) The effect of the instability in experiment A is therefore to reduce the strength of the first over-reflection considerably, but not inhibit it completely. After a period of weak over-reflection the critical layer returns to a configuration in which it is acting as an absorber. There appears to be a tendency for the strength of the absorptivity to decrease with time. In experiment B the unstable disturbances grow to finite amplitude earlier and their effect on $B_1(T)$ is felt before an over-reflecting state has been reached. It appears that in this case the critical layer never acts as an over-reflector. Also shown in figure 5(*b*) is the time-integrated absorptivity $\alpha(T)$. For the SWW case, when $\alpha(T)$ is simply equal to $-\frac{1}{2}$ multiplied by the time integral of $B_1(T)$, the ultimate value of $\alpha(T)$ is 3.0858 (Killworth & McIntyre 1985). As may be seen from figure 5(*b*), the effect of the instability is to increase this by a factor of two or three.

A similar experiment (C) was performed for $\mu = 0.05$ (and with increased resolution in the zonal direction). No initial noise was added, but as in experiment B the instability of the flow allowed disturbances to grow out of the errors associated with the computation. The vorticity pictures at $T = 2.9$ and $T = 3.5$ are shown in figures 6(*a*) and 6(*b*). The disturbances appear to have smaller wavelength than in A and B; in fact their dimensional wavelength is comparable. Because the non-dimensional wavelength is $O(\mu^{-1})$ the number of Fourier components required is at least $3\mu^{-1}$, assuming a bare minimum of three waves to represent the disturbances.

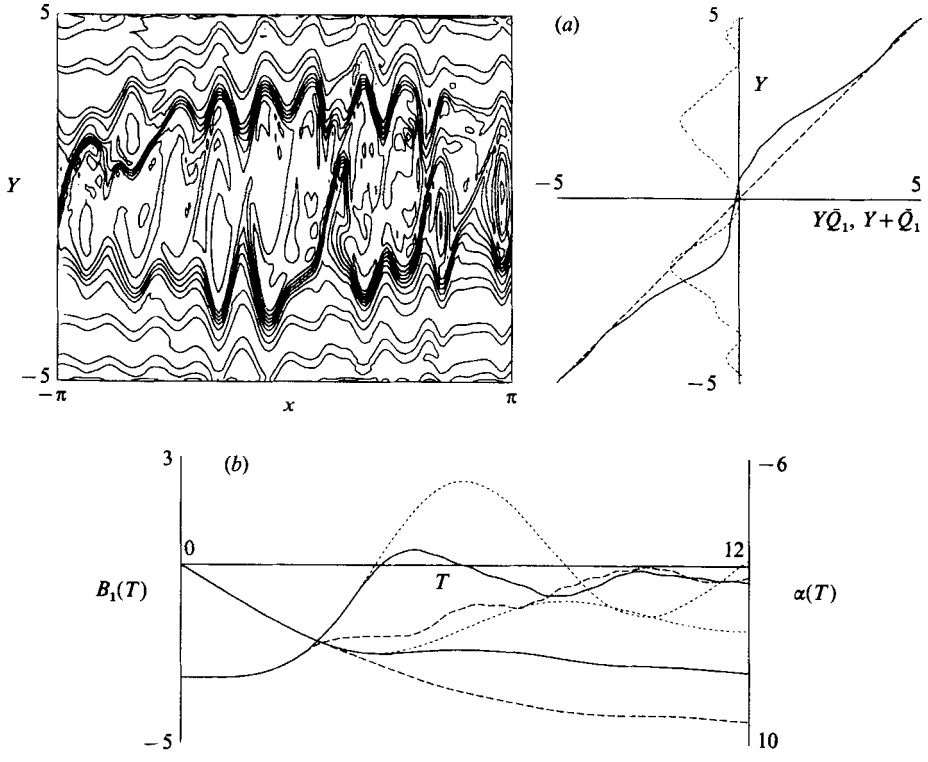

FIGURE 5. (*a*) As figure 1, for experiment A ($\mu = 0.1$) at $T = 5.5$. (*b*) Graphs of absorptivity, $B_1(T)$ against time T for experiment A (solid curve), experiment B (dashed curve) and SWW solution (dotted curve). In each case $B_1(0) = -\pi$. Also shown is the time-integrated absorptivity $\alpha(T)$. Each of these latter curves begins at the origin.

Calculations for smaller and smaller values of μ therefore become expensive very rapidly.

The graph of $B_1(T)$ for experiment C is shown in figure 7, together with that for another experiment, D, in which noise was added at $T = 1.5$. Note that the over-reflection has been inhibited in both cases. Note also that the difference between the cases C and D is much less than that between A and B (figure 5). We might reasonably suppose that this is because the instability acts relatively more rapidly in the former case and because the gross effect of the instability on the critical layer is then less sensitive to the details of the evolution of the unstable disturbances. Some further ideas on this are presented in §6. It appears in this case that the effect of the instability is to increase the ultimate value of the time-integrated absorptivity $\alpha(\infty)$ to two and a half times that predicted by the SWW solution.

Also of interest is the behaviour for larger values of μ, outside the formal range of validity of the local instability theory. The difference in the coefficients ρ_n and σ_n from the $\mu = 0$ case make it almost certain that the evolution will be different from that predicted by the SWW solution. What is in question is whether it is appropriate to interpret this different behaviour as instability. Here we take the working definition that the flow may be regarded as being unstable if the evolution is sensitive to the introduction of noise and that this instability is closely related to the local shear instability if the sensitivity is exhibited within one turnover time for the basic cat's eye flow (i.e. $\frac{1}{2}\pi$ multiplied by the rotation period of a material line element at

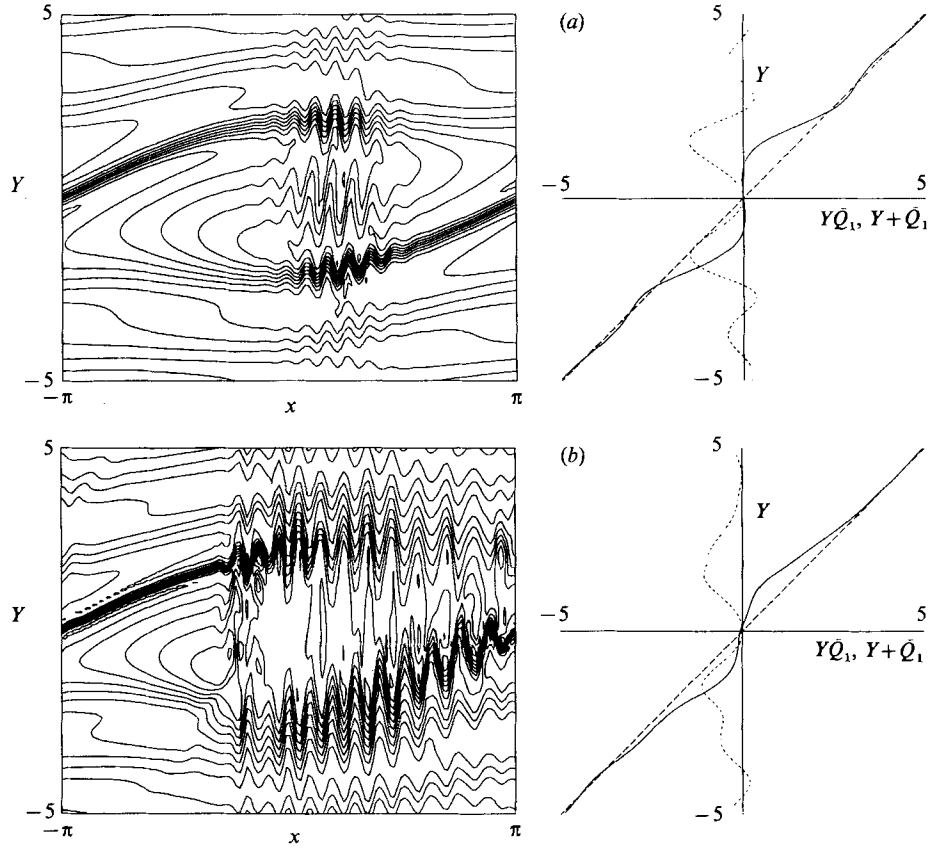

FIGURE 6. As figure 1, for experiment C with $\mu = 0.05$. (*a*) $T = 2.9$, (*b*) $T = 3.5$.

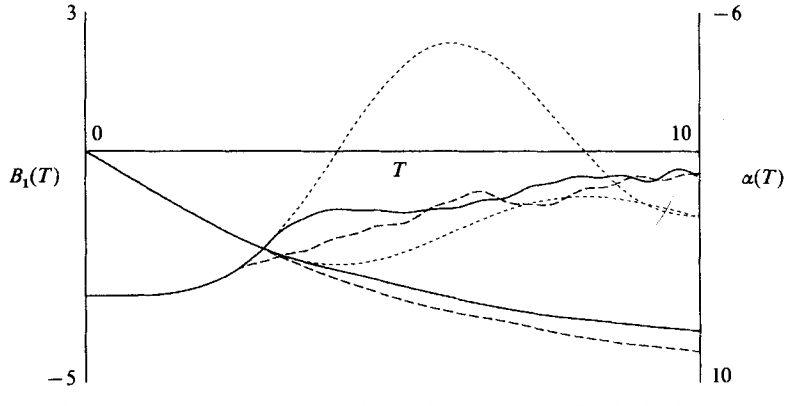

FIGURE 7. As figure 5(*b*), for experiments C (solid curves) and D (dashed curves), both with $\mu = 0.05$, and for the SWW solution (dotted curve).

the centre of the cat's eye). There might well be other instabilities of two-dimensional vortex flow which are important in the time evolution over longer periods.

Numerical experiments were performed for various values of μ between 0.1 and 1 (keeping the same wall position y_0). For each value of μ the tendency towards instability was tested by running two experiments, one with a small-amplitude

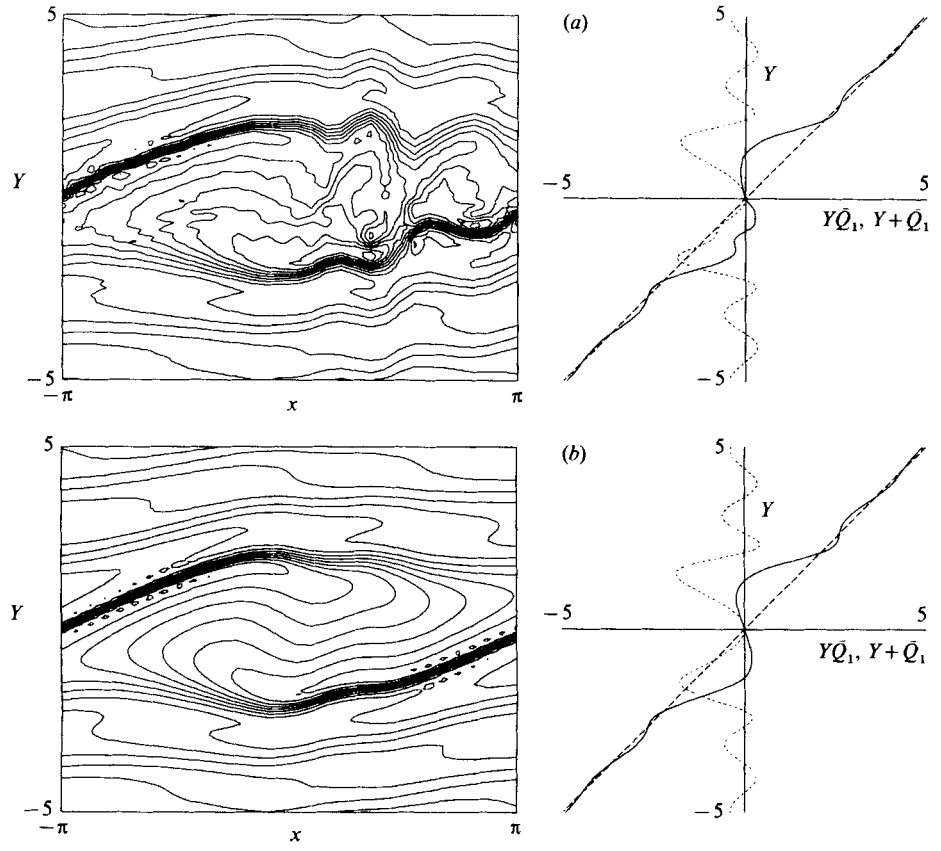

FIGURE 8. As figure 1, for (*a*) experiment E and (*b*) experiment F, with initial noise, both at $T = 3.5$ and both with $\mu = 0.25$.

disturbance superimposed, associated with a maximum value of the vorticity of less than 0.1. For $\mu = 0.25$, experiments E and F, the absolute vorticity fields at $T = 3.5$ are shown in figures 8(*a*) and 8(*b*) respectively. Differences between the evolution in the two experiments may also be seen in the absorptivity curves, shown in figure 9. In particular the ultimate value of the integrated absorptivity appears likely to be about 30% larger in F than in E. When μ was increased to 0.4, experiments G and H, the sensitivity to the introduction of noise was markedly less, as may be seen from figures 10(*a*) and 10(*b*), which show the absolute vorticity fields at $T = 4.0$.

These results might lead us to conclude that the instability plays an important role in the evolution for $\mu < 0.25$, but not for larger values. Does resonance also play a role? For this choice of y_0, the fourth harmonic is in fact very close to resonance when $\mu = \frac{1}{4}$; the third is close when $\mu = \frac{1}{3}$, the fifth when $\mu = \frac{1}{5}$, and so on. There is therefore a sense in which the system is far from resonance when $\mu = 2/(2l+1), l = 1, 2, 3, 4, \ldots$. Experiments were therefore run for $\mu = \frac{2}{7}$ (I) and $\mu = \frac{2}{9}$ (J), to investigate whether any markedly different behaviour could be detected from the case $\mu = 0.25$. Experiments G and H, figure 10, with $\mu = \frac{2}{5}$, are also relevant. Figures 11(*a*) and 11(*b*) show the absolute vorticity fields in experiments E and J respectively at $T = 4.0$. Figure 12 shows the evolution of $B_1(T)$ in experiments E, I and J. There does not appear to be any qualitative difference between the resonant and non-resonant

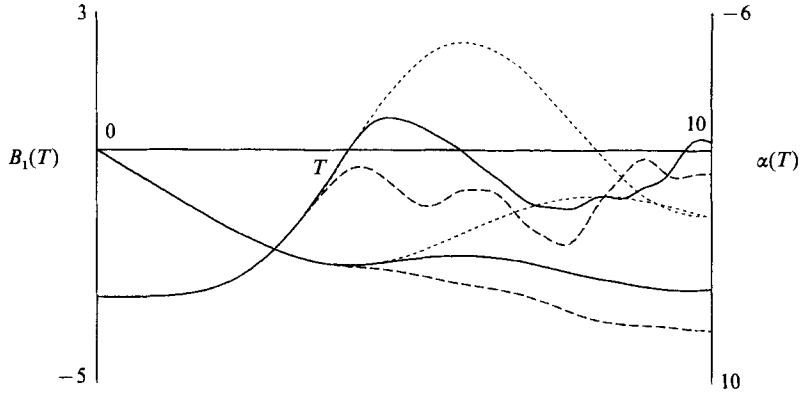

FIGURE 9. As figure 5(*b*), for experiment E (solid line), F (dashed line), both with $\mu = 0.25$, and the SWW solution (dotted line), with $\mu = 0$.

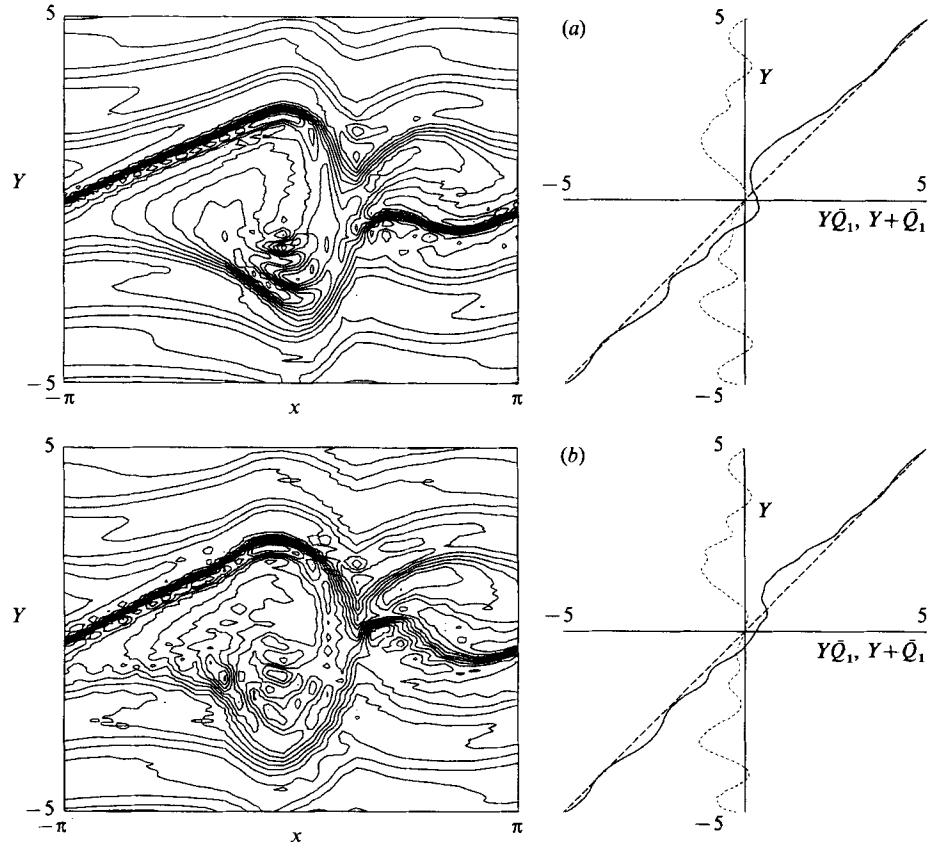

FIGURE 10. As figure 1, for (*a*) experiment G, and (*b*) experiment H, with initial noise, both at $T = 4$ and both with $\mu = 0.4$.

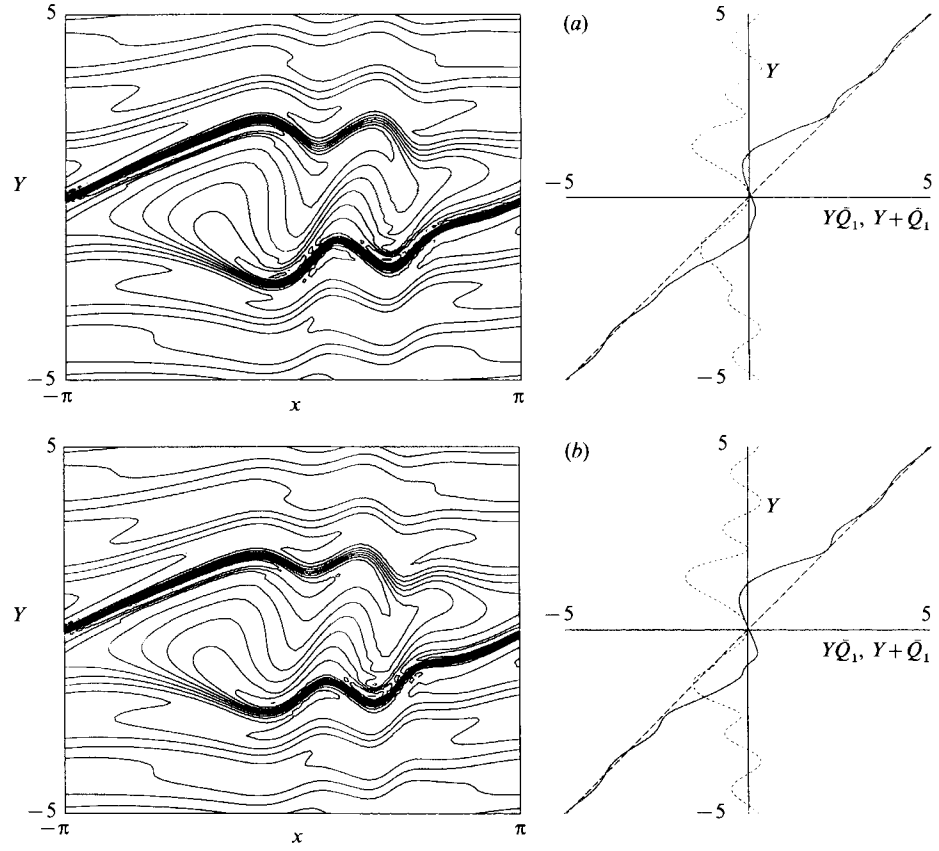

FIGURE 11. As figure 1, for (*a*) experiment E with $\mu = 0.25$ (fourth harmonic near resonance) and (*b*) experiment J with $\mu = \frac{2}{9}$ (no harmonic near resonance), both at $T = 4.0$.

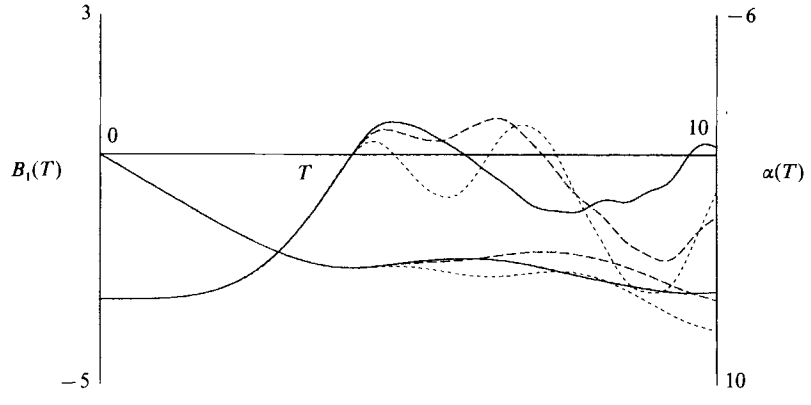

FIGURE 12. As figure 5(*b*) for experiment I (solid curves), experiment J (dashed curve) and experiment E (dotted curve). I and J both have no harmonics near resonance, E has the fourth harmonic near resonance.

cases, indicating again that the resonance of higher harmonics is not important in the time evolution of the flow in the critical layer, at least during this early, essentially inviscid, stage.

5. Re-examination of Béland's numerical experiments

It has already been remarked that Béland's (1976) experiments showed up interesting small-scale structure in the critical layer, e.g. in his figure 4(*c*) at $T = 59$, 86. (Béland's figures are reproduced in Killworth & McIntyre 1985, figure 5.) It has been suggested that this is either a manifestation of the resonance of higher harmonics (Ritchie 1985) or of local instability (Killworth & McIntyre 1985). It should also be noted that the short-wave features in Béland's figure are at the limits of his numerical resolution, since only six harmonics were included in the x-direction. There is also evidence of strong grid-scale structure in the meridional vorticity gradient, seen in his figure 7(*b*). It was decided to attempt to reproduce Béland's results in order to confirm that the behaviour seen was not sensitive to resolution and also to allow more detailed investigation which might throw further light on the question of instability versus resonance.

Béland's experiments were repeated in two ways. First the matching constants ρ_n and σ_n were calculated for the relevant flow profile and external parameter values. Details of this calculation, following Warn & Warn (1978), are given in Appendix A and the results are displayed in table 1. Note that none of these numbers seems particularly large, indicating that resonance (of the forced harmonic or of higher harmonics) is unlikely to be important, whether or not one accepts the stronger conclusions of §§2.2 and 4 concerning resonant effects. The critical-layer model was run with these values of ρ_n and σ_n. Secondly a channel model (developed from a model originally supplied by Dr P. D. Killworth) was run with the parameter values, flow configuration and numerical resolution used by Béland, except that the number of zonal harmonics was increased from 6 to 31. Some details of this calculation are given in Appendix B. It is, of course, entirely independent of matched-asymptotic concepts. Whilst this method was relatively inefficient and expensive, because high resolution was needed over the whole flow domain not just the critical layer, it was well within computing resources typical of those widely available now (as opposed to ten years ago).

Contour plots of absolute vorticity from the two models, at corresponding times, are shown in figures 13 and 14. Note that the former corresponds to a thin slice taken from the central region of the latter. In both models the vorticity pattern in the critical layer appears to break up into features with zonal lengthscale smaller than that of the forced wave. In both cases there are three such features: the central one is longest and has largest extent in the y-direction, the other two do not extend so far in the y-direction and the leftmost is the longer of the two. Agreement between the two models is therefore about as good as could be expected. These features are also qualitatively similar to those seen in Béland's experiment. Close agreement would seem unlikely given the low zonal resolution of Béland's model. Sensitivity to noise inserted in the vorticity field was seen in the evolution of both models, as we might have expected from the fact that the value of μ for these experiments was 0.25. Taking these results together with those of §§2.2 and 4, we conclude, in all likelihood that local instability, and not resonance, played a significant role in the evolution of the disturbances seen in Béland's experiments.

n	ρ_n	σ_n
0	−1.539	−0.931
1	−1.086	−0.650
2	−0.559	−0.290
3	−0.304	−0.295
4	−0.212	0.314
5	−0.478	3.128
6	0.048	−1.038
7	0.003	−0.533
8	0.002	−0.379

TABLE 1. Matching constants for the first eight harmonics for the flow configuration used in Béland's (1976) numerical experiments

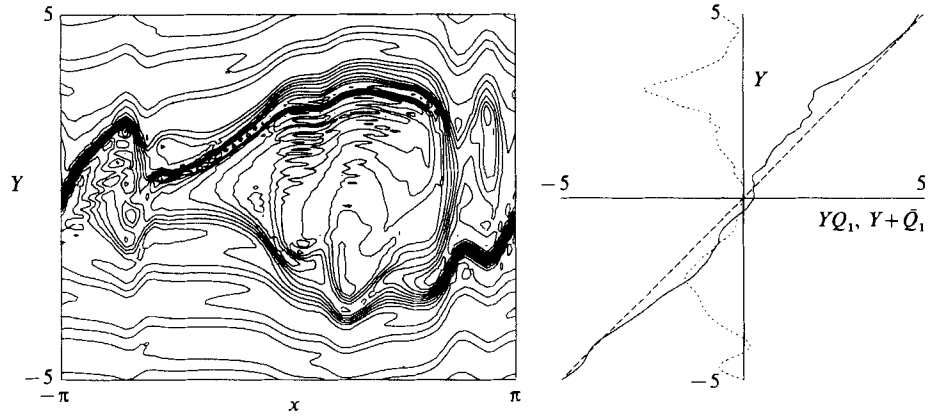

FIGURE 13. As figure 1, for experiment in Béland's configuration at $T = 5.0$.

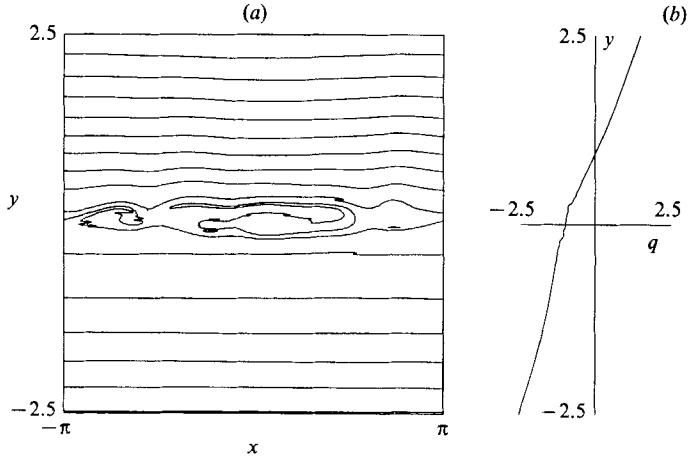

FIGURE 14. (*a*) Contours of absolute vorticity, at $T = 65$, in an experiment designed to be as close as possible to Béland's, except that the x-resolution is much finer. The model had 401 grid points across the domain and represented x-variation using a 31 harmonic spectral expansion. Note that the entire flow is shown, not just the critical layer. (*b*) The x-average of the absolute vorticity profile.

6. A simplified model of the effect of the instability on absorption and reflection

The numerical experiments described in the previous two sections have used direct numerical simulation of the flow in the critical layer to assess the effects of the shear instability. The results confirm what we might have expected from the stability analysis in Killworth & McIntyre (1985), that the instability is most effective in the limit $\mu \to 0$. The minimum value of μ for which a simulation was run was 0.05. For reasons already mentioned, the computational resources required for still smaller values of μ are very large. Nonetheless, it is interesting to consider the limit $\mu \to 0$. In this limit the turnover time for the eddies is much less than for the flow as a whole. We might imagine that the large-scale flow then depends only on some statistical measure of the eddies, rather than details of individual eddies. It seems likely that in this limit the family of curves which form graphs of B_1 against time (as shown in figures 5(*b*) and 7, for example) tend to some limiting shape.

Although direct numerical simulation is out of the question for very small values of μ, there remains the possibility of investigating the limiting shape by somehow modelling the effects of the instability. The model described and applied in this section is based on the idea that, for small μ, the instability is expected to grow very rapidly in comparison with the rate at which the basic flow evolves (Killworth & McIntyre 1985; Haynes 1985). The basic model assumption is that, as soon as the gradient of absolute vorticity reverses sign at any point along the critical layer, the effect of the instability is to adjust the local vorticity profile until it becomes stable. The cat's eye flow then acts upon it in such a way that it inevitably becomes unstable, and the process repeats.

To carry this idea further, and make it precise enough to define the simplified model, let us recall the implications of the zonal mean vorticity equation in the critical layer, i.e. (2.33) with $D = 0$, being

$$\frac{\partial \bar{Q}}{\partial T} = -\overline{\frac{\partial C}{\partial x}\frac{\partial Q}{\partial Y}}. \tag{6.1}$$

Just as in §2, this equation may be multiplied by Y and integrated with respect to Y across the critical layer to give

$$\begin{aligned}\frac{\partial}{\partial T}\int_{-\infty}^{\infty} YQ\,\mathrm{d}Y &= -\int_{-\infty}^{\infty} Y\overline{\frac{\partial C}{\partial x}\frac{\partial Q}{\partial Y}}\,\mathrm{d}Y \\ &= -\int_{-\infty}^{\infty} Q\frac{\overline{\partial C}}{\partial x}\,\mathrm{d}Y = [\overline{u'v'}]_{0^-}^{0^+}.\end{aligned} \tag{6.2}$$

The second equality may be demonstrated by integrating by parts and using the asymptotic form of Q as $|Y| \to \infty$, the third by using the matching conditions with the outer flow. As remarked in §2.3, it may be shown that the only terms which contribute are those which are directly forced at the wall, in this case only the first harmonic $n = 1$. This may be regarded as a consequence of the fact that, in this flow configuration, the convergence of the momentum flux into the critical layer is equal to the force exerted on the flow at the wall. The presence of the wall, together with the quasi-steady property of the wave field, means that there can be no momentum flux associated with the systematic propagation of wave activity away from the critical layer in higher harmonics, whether they arise through the instability, or through any other cause.

This result may now be used to constrain the possible parametrizations of the effect of the short-wavelength rapidly growing instability on the large-scale flow. The vorticity equation may be written in the form

$$\frac{\partial Q}{\partial T}+Y\frac{\partial Q}{\partial x}+\frac{\partial C}{\partial x}\frac{\partial Q}{\partial Y}=F, \tag{6.3}$$

where the quantities Q and C are taken to represent only the slowly varying parts of the absolute vorticity and stream-function fields and F represents the averaged effects on the large scale of the nonlinear terms associated with the rapidly varying part of the flow. There is no contribution to F from those harmonics which are directly forced and, from the arguments given earlier, it is therefore required that F does not result in an artificial source or sink of momentum. A repeat of the analysis leading to (2.31), applied to (6.3), shows that an artificial momentum source will be avoided if F satisfies the constraint

$$\int_{-\infty}^{\infty} YF\,\mathrm{d}Y=0 \tag{6.4}$$

at each point along the critical layer.

Within this restriction there are still an enormous number of possible forms of F which may be chosen. Given that the parametrization is to be implemented within a numerical model it seems easiest to specify F within a numerical scheme. The chosen method will now be described.

At each time step the Q-field is incremented in two stages. First the Q-field is changed as if it was advected solely by the large-scale flow over a time interval ΔT, say. Because the special case $g(0, y_b)$ is being considered, where the velocity field in the critical layer is independent of the vorticity field in the critical layer ($C = \cos x$), it is convenient to use a rather crude semi-Lagrangian method for this part of the time step. A grid of equally spaced x-values (ranging from $-\pi$ to π) and of equally spaced Y-values (ranging from $-Y_{\max}$ to $Y_{\max}$) is set up. An array of ordered pairs (x_0, Y_0) is then calculated, each member of the array corresponding to a point on the grid, and being the x- and Y-coordinates of the starting position of a particle that would arrive at the point on the grid a time ΔT later. The updated value of Q at a given point (x, Y) was therefore calculated by evaluating the old Q-value at the corresponding 'departure' point (x_0, Y_0) using linear interpolation on the grid.

The next part of the time step depends on whether the Y-derivative of the new Q-field was everywhere positive, or changed sign somewhere in the domain. At those values of x for which $\partial Q/\partial Y$ was positive for all Y, no further adjustment is made to the Q-field. On the other hand, at those values of x for which $\partial Q/\partial Y$ is negative for some range of Y, the effect of the instability is modelled by redistributing the Q-field in the Y-direction in the manner shown in figure 15. The redistribution is taken to be such that, in the simplest case, the adjusted $\partial Q/\partial Y$ is non-negative everywhere and the adjusted Q-field at the value of x in question takes the constant value, Q_m say, between $Y = Y_-$ and $Y = Y_+$. In more complicated cases, where the Q-field before adjustment includes more than one contiguous region of reversed gradients, it is necessary to have a number, I, of regions with, again at given x, $Q = Q_m^{(i)}$, $Y_-^{(i)} < Y < Y_+^{(i)}$, $i = 1, \ldots, I$.

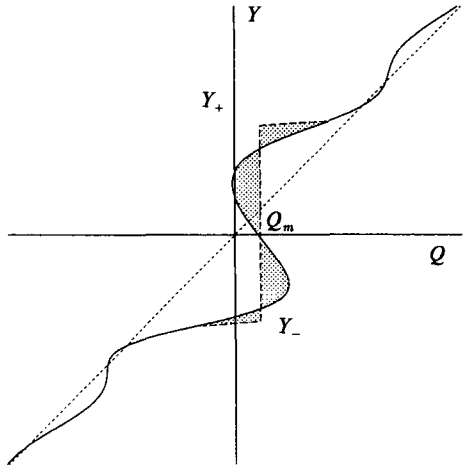

FIGURE 15. Adjustment scheme for the absolute vorticity field. The solid curve shows the absolute vorticity profile at a given value of x plotted against Y. The dashed curve shows the adjusted profile, which takes a constant value Q_m between $Y = Y_+$ and $Y = Y_-$. In order that total absolute vorticity is conserved, the net positive and negative contributions from the shaded areas cancel. The contributions from the Y-weighted areas must also cancel, to ensure no artificial momentum source (see text).

Each adjustment depends on the three constants Y_-, Y_+ and Q_m and therefore requires three constraints in order to be specified uniquely. The first is that

$$\int_{Y_-}^{Y_+} (Q-Q_m)\,\mathrm{d}Y = 0 \tag{6.5}$$

so that total vorticity is conserved. The second is that

$$\int_{Y_-}^{Y_+} Y(Q-Q_m)\,\mathrm{d}Y = 0 \tag{6.6}$$

as implied by (6.4) and ensuring that there is no artificial source of momentum in the critical layer. The third condition has to be chosen more arbitrarily. For the vorticity gradients to be one-signed after the adjustment has been applied, Q_m must be such that $Q(Y_-) \leqslant Q_m \leqslant Q(Y_+)$ and, bearing this in mind, we specify the third condition as

$$Q_m = \tfrac{1}{2}(Q(Y_-)+Q(Y_+)). \tag{6.7}$$

It is shown in Appendix C that this condition implies that the width $|Y_+ - Y_-|$ of the mixed region is maximized, subject to the constraints (6.5) and (6.6). The computational problem therefore reduces to solving (6.5), (6.6) and (6.7), given an unadjusted profile $Q(Y)$. Details of the algorithm used, which required some delicacy, are presented in Appendix D.

The model was run with a number of different resolutions in the x- and Y-directions and with a number of different values of ΔT. The results presented are for resolutions at which further increase gave no qualitative change in the behaviour. Figure 16 shows a typical configuration of the vorticity field during the evolution, at $T = 3$. This might therefore be compared with figure 6, for example, showing the vorticity field at a similar stage during the direct simulation, with $\mu = 0.05$. The variation of B_1 and the time-integrated absorptivity with time are shown in figure 17. The value

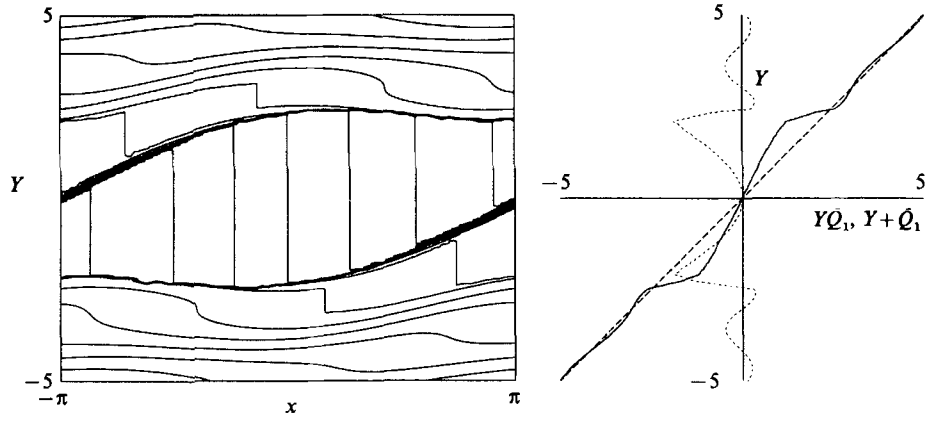

FIGURE 16. As figure 1, for an experiment in which the absolute vorticity field was adjusted under the scheme depicted in figure 15. The absolute vorticity field is shown at $T = 3.0$.

of $\alpha(10)$ predicted by this model was 8.15. (Compare the SWW result of $\alpha(\infty) = 3.0858$.) For comparison the corresponding curves from the direct simulation are superimposed. Considering the crudity of the model, it gives surprisingly close quantitative agreement with the direct simulation. This agreement lends support to our earlier picture of the evolution for small μ, i.e. rearrangement of the vorticity field in the critical layer by the flow associated with the unstable disturbances, inhibition of over-reflection and much more gradual attainment of a reflecting state.

A second mixing algorithm was also tried, in which the requirements were (6.5), together with $Q_m = Q(Y_-) = Q(Y_+)$. A typical adjustment under this scheme is shown in figure 18. It may be seen that

$$\int_{Y_-}^{Y_+} Y(Q-Q_m)\,\mathrm{d}Y < 0.$$

Considerations similar to those in §2.3 suggest that such an adjustment requires a momentum-flux *convergence* into the critical layer and must therefore be associated with a flux of wave activity away from the critical layer. As discussed earlier, this is not possible within the configuration considered here, where the waves are forced at a wall, but would be possible if some alternative boundary condition, perhaps a radiation condition, were applied. Of course, altering this boundary condition would violate the conditions necessary for the SWW-like case in which the stream function pattern in the critical layer is independent of time and, strictly speaking, the numerical method used in this section for following the evolution of the cat's eye flow as a whole would not be applicable. Nonetheless, as a heuristic exercise, the second mixing algorithm was applied to the SWW case, overlooking the inconsistency.

The graphs of B_1 and related quantities against time for this experiment are shown in figure 19. B_1 may now remain positive (or negative) for all time, since the associated contribution to the absorptivity may be cancelled by contributions from higher harmonics. Also of interest is the behaviour of the total time-integrated absorptivity, including all harmonics, as diagnosed from the time integral of the first term on the right-hand side of (2.33). This is shown as the dotted line in figure 19. Taken at face value, and when compared with figure 15, this suggests that there might be a tendency for the critical layer to be less absorptive (in the time-integrated

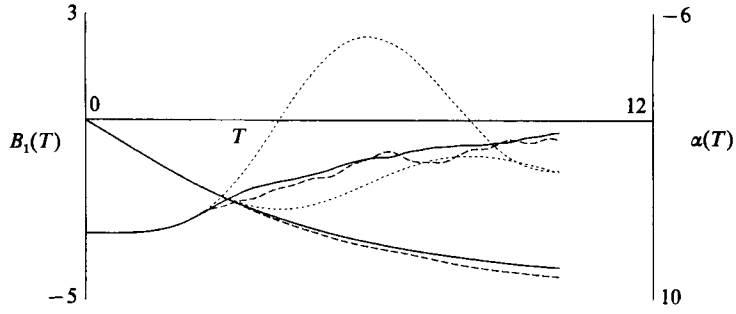

FIGURE 17. As figure 5(*b*), for an experiment in which the absolute vorticity field was adjusted (solid curves), for experiment C with $\mu = 0.05$ (dashed curves) and for the SWW analytical solution (dotted curves).

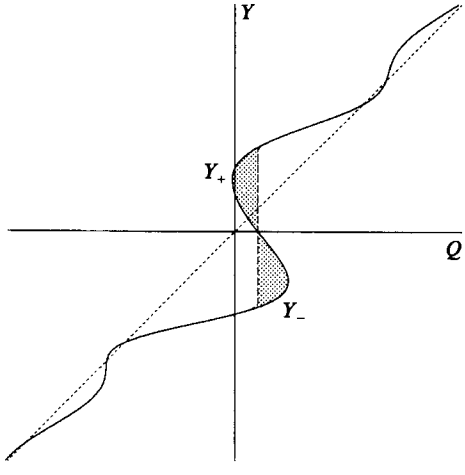

FIGURE 18. Adjustment scheme for the absolute vorticity field which requires a net source of momentum in the critical layer. As for figure 15, except that the Y-weighted areas do not cancel.

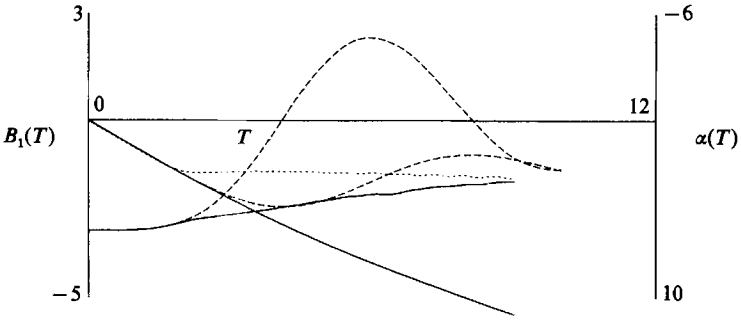

FIGURE 19. As figure 5(*b*), for an experiment in which the absolute vorticity field was adjusted under the scheme depicted in figure 18 (solid curves) and for the SWW analytical solution (dashed curve). The time integral of $\frac{1}{2}B_1(T)$ is shown in place of the absorptivity $\alpha(T)$. The dotted curve shows the quantity $\int YQ_1\,\mathrm{d}Y$ plotted on the same scale.

sense) when there is the possibility of net re-radiation into the outer flow. Of course, more careful study would be required before any definite conclusion could be reached on this point.

7. Discussion

The new work in this paper has been almost entirely based on numerical simulations, and it is therefore important to establish that the simulated time evolution of the nonlinear critical layer represents fluid-dynamical reality, and is not the result of artificial numerical instability, for example. As has been stressed earlier, the majority of the numerical experiments reported here have been repeated at different resolution, both with different numbers of harmonics in the x-direction, and with different grid and domain sizes in the y-direction. Good quantitative agreement has been found between different versions, provided, of course, that the resolution was fine enough. Furthermore, the behaviour seen in the simulations is entirely consistent with the local instability theory of Killworth & McIntyre (1985) and Haynes (1985). It seems almost inconceivable that the superposition of x-harmonics necessary to produce the field shown in figure 6(*b*), for example, which is physically highly plausible, could have arisen through numerical artifact.

The most important conclusion to be drawn from this work concerns the temporary absorptivity of the critical layer. Whilst, under appropriate conditions, the time-integrated absorptivity remains bounded whatever the details of the flow in the critical layer (Killworth & McIntyre 1985), what is now clear is that these details can have a drastic effect on instantaneous values of the absorptivity, and thus on the numerical value of the time-integrated absorptivity. For instance, it appears that the effect of the local barotropic instability in the critical layer can be to increase the time-integrated absorptivity to three or four times the value predicted by the SWW solution (see figures 5 and 6). The growth of unstable disturbances and the subsequent rearrangement of the vorticity field over a region wider than the width of the cat's eyes has the effect of inhibiting any over-reflection almost completely, as well as considerably increasing the time-integrated absorptivity. Indeed, Killworth & McIntyre's bound is sensitive to the effective width b of the region subject to vorticity rearrangement, going like b^3. Presumably this change in the matching condition across the critical layer has implications, not only in the forced-wave problem, but also where the waves are growing on an unstable shear flow, as studied recently by Churilov & Shukhman (1987), where it would presumably lead to an increased prediction of the saturation amplitude in the inviscid case.

For small values of μ (i.e. <0.2), the instability is spectacular in its effect on the critical layer, as was expected on the basis of the theory presented by Killworth & McIntyre (1985). As μ is increased the effect of the instability becomes less remarkable, although there are still grounds for regarding the flow as unstable, even up to values of μ of about 0.3, on the basis that the flow exhibits sensitivity to initial conditions. In the flow configurations investigated there does not seem to be any evidence for changes in the character of the flow when higher harmonics are resonant. However, there does not seem to be any reason to doubt the conclusion of Ritchie (1985) that on long timescales, when gross features of the flow are determined by weak dissipative processes, there may be more sensitivity to resonance. The author has independently reproduced Ritchie's results in these parameter regimes. It may also be that resonant effects would show up more strongly in flow configurations for which σ_1 was not small. The reason why Ritchie (1985) did not find unstable

behaviour was that the viscosity he chose was so large that the reversals achieved in the absolute vorticity gradient were rather weak. There is evidence of this from his figures 6 and 7, for example.

Under geophysical parameter regimes (where there seems little *a priori* reason for introducing dissipation in the form of a constant diffusivity) it seems that reversed gradients are commonplace and Juckes & McIntyre (1987) have recently reported high-resolution numerical simulations in which the accompanying shear instability is clear. In such simulations the nonlinear critical layers are no longer thin, but the essential dynamics are very closely related. However, the effect on the absorption/reflection properties in the simulations has not yet been quantified and the small-amplitude results in figures 5(*b*) and 7, for example, are therefore our first indication of that effect.

What has also become clear from the work reported here, and from previous work already cited, is that the critical-layer equations (2.19) and (2.24) admit solutions with a large range of different behaviour, despite the fact that they are considerably simplified from the equations for two-dimensional flow in general. Any numerical work, such as that presented in this paper and the earlier literature can only investigate a relatively small set of external parameters. It is interesting to speculate on whether, for example, the system does always eventually settle down to a steady state when a steady forcing is applied. The work of Ritchie (1985) and experiments done by this author, suggest, but do not prove, that it does not. And if this is not the case, is the eventual time-dependent motion periodic, or irregular? The results presented in §4, for example, suggest that with small but non-zero dissipation a disordered motion, which one might call 'critical-layer turbulence', is a strong possibility.

Much of this work was performed in the Department of Atmospheric Sciences, University of Washington, financial support being provided by the Joint Institute for the Study of the Atmosphere and the Ocean and numerical simulations run on the PR1ME computer. The remainder of the computation was carried out on the IBM3084 in Cambridge, courtesy of the University of Cambridge Computing Service, and on the CRAY-1 in London, courtesy of the University of London Computing Centre, while the author held a Research Fellowship at Queens' College, Cambridge, and subsequently a Royal Society Meteorological Office Research Fellowship in Dynamical Meteorology. Additional support during the preparation of the manuscript was provided by the US Office of Naval Research. The work benefited from conversations or correspondence with S. J. Cowley, P. D. Killworth and M. E. McIntyre.

Appendix A. Calculation of ρ_n and σ_n for a general flow profile

For a general profile the functions $f(\mu n, y), g(\mu n, y)$ and $h(\mu n, y)$ do not take the forms (2.13*a*, *b*, *c*) but are solutions of the differential equation

$$U(y)\left(\frac{\mathrm{d}^2\chi}{\mathrm{d}y^2}-\mu^2n^2\chi\right)+G(y)\,\chi=0, \tag{A 1}$$

where U and G are the velocity and vorticity-gradient distributions, non-dimensionalized using the *local* values of shear Λ and vorticity gradient β at the critical line $y=y_c$. We thus have $U(y_c)=0$, $U'(y_c)=1$ and $G(y_c)=1$. Two

independent solutions of (A 1) may be specified by their asymptotic forms near the critical line,

$$\chi_a(\mu n, y) = 1-(y-y_c)\ln|y-y_c|+O\{(y-y_c)^2\ln(y-y_c)\} \quad (A\ 2a)$$

and

$$\chi_b(\mu n, y) = y-y_c+O((y-y_c)^2). \quad (A\ 2b)$$

By analogy with (2.13) and (2.14) we may write

$$f(\mu n, y) = \chi_a(\mu n, y)+a_1\chi_b(\mu n, y) \quad (y > y_c), \quad (A\ 3a)$$

$$g(\mu n, y) = \chi_b(\mu n, y) \quad (y > y_c) \quad (A\ 3b)$$

and

$$h(\mu n, y) = \chi_a(\mu n, y)+a_1\chi_b(\mu n, y) \quad (y < y_c), \quad (A\ 3c)$$

where the constant a_1 is chosen such that $h(\mu n, y)$ satisfies the required boundary condition as $y \to -\infty$ or at some other boundary in $y < y_c$. These forms may then be substituted in (2.25a) and (2.25b) to give ρ_n and σ_n.

The problem therefore reduces to that of calculating the constant a_1 and the functions χ_a and χ_b, which may be done through numerical integration of (A 1) with appropriate care near the singular point at $y = y_c$.

Appendix B. Repeat of Béland's numerical experiments

Equations (2.1) and (2.3) were solved in a channel geometry, with boundaries at $y = \pm 2.5$. The basic vorticity gradient β was taken to be 1.6 and the flow specified to be $(u, v) = (\tanh y, 0)$ at $t = 0$. The boundary conditions were taken as $\psi = 0$ at $y = -2.5$ and $\psi = 0.009\cos(0.4x)[1-\cos \pi t/3.5]$ (for $t < 3.5$) and $\psi = 0.018\cos(0.4)x$ (for $t >= 3.5$) at $y = 2.5$. The matching constants for a critical-layer problem in which these were the specified boundary conditions were calculated using the method described in Appendix A (see table 1). It was found that there was little change in the matching constants when the value of y was decreased from -2.5 and it was concluded that there were no serious differences between using these boundary conditions and those used by Béland (1976, 1978), where a radiation condition was applied for negative y. Béland used a different function of time in the boundary condition at $y = -2.5$, in which the value of ψ increased linearly to its steady state value at $t = 3.5$. The function used here has continuous first and second derivatives at $t = 0$ and at $t = 3.5$.

The above configuration may be represented in terms of the non-dimensional quantities introduced in §2 by taking $\mu = 0.25$, $\epsilon = 0.06$ and $y_b = 1.5$. A time unit on the dimensionless 'inner' timescale represents a dimensionless time interval in the channel model of about 11. The critical-layer thickness is about 0.11. Inspection of the results from the two models suggested that the times $t = 65$ and $T = 5.5$, for example, were roughly equivalent, and these are in a ratio which seems entirely consistent with the scalings, bearing in mind the level of precision of the comparison.

Appendix C

It is required to show that one root of (6.7) maximizes $Y_+ - Y_-$ subject to the constraints (6.5) and (6.6). Note that Q_m may be eliminated from (6.6) using (6.5) to give the single constraint

$$\int_{Y_-}^{Y_+} Q(Y-\tfrac{1}{2}(Y_+ + Y_-))\,dY = 0. \quad (C\ 1)$$

If Y_+ and Y_- vary by the amounts δY_+ and δY_- respectively, then the first variation in (C 1) implies that

$$\delta Y_+\left\{\tfrac{1}{2}(Y_+-Y_-)\,Q(Y_+)-\tfrac{1}{2}\int_{Y_-}^{Y_+}Q\,\mathrm{d}Y\right\}+\delta Y_-\left\{\tfrac{1}{2}(Y_+-Y_-)\,Q(Y_-)-\tfrac{1}{2}\int_{Y_-}^{Y_+}Q\,\mathrm{d}Y\right\}=0. \quad \text{(C 2)}$$

This may be written in more succinct form as

$$\frac{\delta Y_+}{\delta Y_-}=-\frac{Q(Y_+)-Q_m}{Q(Y_-)-Q_m}. \quad \text{(C 3)}$$

If $|Y_+-Y_-|$ is made stationary as Y_- and Y_+ vary subject to (C 1), then $\delta Y_+-\delta Y_-=0$, and (C 3) leads to (6.7), i.e. that

$$Q_m=\tfrac{1}{2}(Q(Y_+)+Q(Y_-)). \quad \text{(C 4)}$$

It remains to identify which roots of this equation are of interest. The easiest case to consider is that where (for given x) there is a single region in which $\partial Q/\partial Y<0$. Suppose that this region is the interval $(Y_-^{(r)}, Y_+^{(r)})$. (C 3) represents a differential equation which describes the locus of the points Y_-, Y_+ which satisfy (C 1). After some analysis it may be shown that this locus begins at $(Y_-, Y_+)=(Y_-^{(r)}, Y_+^{(r)})$ and ends at $(Y_-, Y_+)=(Y_+^{(r)}, Y_+^{(r)})$. The quantity $\mathrm{d}Y_+/\mathrm{d}Y_-$ is negative at $(Y_-^{(r)}, Y_-^{(r)})$, positive on some intermediate part of the locus, becomes unbounded when Y_- stops increasing and begins to decrease, and then remains negative until $(Y_+^{(r)}, Y_+^{(r)})$. Thus Y_+-Y_- has one maximum on the intermediate part of the locus where $\mathrm{d}Y_+/\mathrm{d}Y_->0$, $Q(Y_-)<Q_m<Q(Y_+)$ and therefore, as may be seen by considering figure 15 in the light of (6.5), $Y_-<Y_-^{(r)}$ and $Y_+^{(r)}<Y_+$. When there is more than one region where $\partial Q/\partial Y<0$ it is not straightforward to show that there is only one local maximum in Y_+-Y_- although numerical evidence suggests that this is the case.

Appendix D

The algorithm used to find the roots of (6.5), (6.6) and (6.7) (or equivalently (C 1) and (C 4) was based on consideration of the locus in the (Y_-, Y_+) plane of points satisfying (6.5) and (6.6) (or equivalently (C 1)) only. The Q-field at the various grid points was carried along in the calculation and so for the purposes of the adjustment the profile was specified at a finite number of values of Y. These values were interpolated linearly to give a continuously varying (but not differentiable) profile on which the problem to be solved was well-posed.

The adjustment algorithm has two stages.

(i) The value of Y_+ was fixed at $Y_+^{(r)}$ and the equation (C 1), with the requirement $Y_-<Y_-^{(r)}$, was solved for Y_- using interval halving. This gave one solution with $Q_m>\frac{1}{2}(Q(T_+)+Q(Y_-))$.

(ii) Y_+ was increased by increments, and for each Y_+, the Y_- satisfying (C 1) was calculated using interval halving. The first guesses for Y_- were generated using the expression for the derivative (C 3). When Q_m was found to be less than $\frac{1}{2}(Q(Y_+)+Q(Y_-))$ the procedure was modified to interval halving on Y_+, followed at each stage by interval halving on Y_- to solve (C 1), until $Q_m-\frac{1}{2}(Q(Y_+)+Q(Y_-))$ was sufficiently small.

This method might appear cumbersome, but was sufficiently well conditioned that it gave solutions for all the profiles encountered, including those for which there was more than one Y-interval on which $\partial Q/\partial Y<0$.

REFERENCES

ABRAMOWITZ, M. & STEGUN, I. A. 1965 *A Handbook of Mathematical Functions.* Dover, pp. 1046.

BÉLAND, M. 1976 Numerical study of the nonlinear Rossby-wave critical layer 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.

CHURILOV, S. M. & SHUKHMAN, I. G. 1987 The nonlinear development of disturbances in a zonal shear flow. *Geophys. Astrophys. Fluid Dyn.* **38**, 145–175.

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

DRAZIN, P. G. & REID, W. H. 1981 *Hydrodynamic Stability.* Cambridge University Press, pp. 525.

GEISLER, J. E. & DICKINSON, R. E. 1974 Numerical study of an interacting Rossby-wave and barotropic zonal flow near a critical level. *J. Atmos. Sci.* **31**, 946–955.

HABERMAN, R. 1972 Critical layers in parallel flows. *Stud. Appl. Maths.* **51**, 139–161.

HAYNES, P. H. 1985 Nonlinear instability of a Rossby-wave critical layer. *J. Fluid Mech.* **161**, 493–511.

HAYNES, P. H. 1987 On the instability of sheared disturbances. *J. Fluid Mech.* **175**, 463–478.

HAYNES, P. H. 1988 Forced, dissipative generalizations of finite-amplitude wave-activity conservation relations for zonal and non-zonal basic flows. *J. Atmos Sci.* **45**, 2352–2362.

HAYNES, P. H. & MCINTYRE, M. E. 1987 On the representation of Rossby-wave critical layers and wave breaking in zonally truncated models. *J. Atmos. Sci.* **44**, 2359–2382.

HELD, I. M. & PHILIPS, P. J. 1987 Linear and nonlinear barotropic decay on the sphere. *J. Atmos. Sci.* **44**, 200–207.

JUCKES, M. N. & MCINTYRE, M. E. 1987 A high-resolution one-layer model of breaking planetary waves in the stratosphere. *Nature* **328**, 590–596.

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

MASLOWE, S. A. 1986 Critical layers in shear flows. *Ann. Rev. Fluid Mech.* **18**, 405–432.

MCINTYRE, M. E. & PALMER, T. N. 1985 A note on the general concept of wave breaking. *Pure Appl. Geophys.* **123**, 964–975.

RITCHIE, H. 1984 Amplification of forced Rossby waves in the presence of a nonlinear critical layer. *J. Atmos. Sci.* **41**, 2012–2019.

RITCHIE, H. 1985 Rossby-wave resonance in the presence of a nonlinear critical layer. *Geophys. Astrophys. Fluid Dyn.* **31**, 49–92.

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. Maths* **27**, 133–175.

TEMME, N. M. 1983 The numerical computation of the confluent hypergeometric function $U(a, b, z)$. *Numer. Maths* **41**, 63–82.

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 critical level. *Stud. Appl. Maths* **59**, 37–71.