Nonlinear Balance and Potential Vorticity Thinking at Large Rossby Number

D. J. Raymond

Physics Department and Geophysical Research Center,
New Mexico Institute of Mining and Technology,
Socorro, NM 87801

SUMMARY

Two rational approximations are made to the divergence, potential vorticity, and potential temperature equations resulting in two different nonlinear balance models. *Semibalance* is similar (but not identical to) the nonlinear balance model of Lorenz. *Quasibalance* is a simpler model that is equivalent to quasigeostrophy at low Rossby number and the barotropic model at high Rossby number. Practical solution methods that work for all Rossby number are outlined for both models. A variety of simple initial value problems are then solved with the aim of fortifying our insight into the behaviour of large Rossby number flows. The flows associated with a potential vorticity anomaly at very large Rossby number differ in significant ways from the corresponding low Rossby number results. In particular, an isolated anomaly has zero vertical radius of influence at infinite Rossby number, while the induced tangential velocity in a horizontal plane containing the anomaly decreases as the inverse of radius rather

than the inverse square. The effects of heating and frictional forces are approached from a point of view that is somewhat different from that recently expressed by Haynes and McIntyre, though the physical content is the same.

1 Introduction

In a highly influential paper, Hoskins, McIntyre, and Robertson (1985) (hereafter HMR – see also Hoskins, McIntyre, and Robertson, 1987) presented a conceptual framework for thinking about large scale flow patterns that are nearly balanced. In this framework potential vorticity plays a central role. First, as a quantity that is conserved in adiabatic, frictionless flows, it makes the evolution of large scale flow patterns accessible to the intuition. Second, assuming that some balance condition is approximately satisfied by the flow, all interesting dynamic and thermodynamic fields may be recovered. This is called the "invertibility principle".

This framework arguably adds no physics that isn't already included in the traditional way of thinking about large scale dynamics in terms of the tendency and omega equations. However, many people find that the potential vorticity framework makes it much easier to obtain physical insight into the workings of large scale flows. This aid to physical insight may in fact be the primary benefit of such "potential vorticity thinking".

Though HMR touch upon the possibility of using other balance conditions (and considered gradient wind balance in an axisymmetric vortex), most of their qualitative insights are based on the use of geostrophic balance. This leads to a particularly nice approximation in which the potential vorticity becomes the source term of a Poisson equation for the pressure or geopotential height. However, geostrophic balance has obvious limitations on time scales of order or less than the inverse of the Coriolis parameter. It is thus not generally applicable to mesoscale dynamics, which may be defined as atmospheric circulations with Rossby number of order unity.

It should be noted that this does not impose a fundamental limitation on potential vorticity thinking, since the Ertel potential vorticity is exactly conserved for *any* adiabatic, frictionless flow. It is conceivable that flows on short time scales might be approximately balanced in some sense (i. e., that some diagnostic relationship exists between thermodynamic and dynamic fields), leading to invertibility on these time scales.

Nonlinear balance is a possible candidate for a balance condition on short time scales. It is approximately valid when circulations are nearly horizontal, and it is obtained (in most cases) by dropping all terms in the divergence equation containing the divergence, vertical velocity, and the divergent components of horizontal velocity. Bolin (1955, 1956) and Charney (1955, 1962) were the first to use nonlinear balance in numerical models. However, Lorenz (1960) was the first to develop a nonlinear balance approximation to the primitive equations in a manner that conserved an energy-like quantity. In the Lorenz model all first order terms in the vorticity equation involving the divergent part of the wind must be retained. The Lorenz model thus arises from an inconsistent truncation of the primitive equations, since only the leading order terms are retained in the divergence equation.

McWilliams (1985) pointed out that the condition of nonlinear balance is valid under certain circumstances for all Rossby number. Thus, replacing geostrophic balance with nonlinear balance should allow potential vorticity thinking to be extended to time scales much less than the inverse of the Coriolis parameter as long as the flow remains nearly horizontal.

Recently the Lorenz model has been used to study small scale oceanic circulations. Norton, McWilliams, and Gent (1986) developed a numerical scheme to integrate the Lorenz model, while McWilliams and Gent (1986) and McWilliams, Gent, and Norton (1986) have used this model to study the evolution of oceanic vortices on a beta plane.

One problem with the Lorenz model is that it does not exhibit exact parcel conservation of potential

vorticity. Allen (1991) developed a model in which more terms are retained in the divergence equation than in the Lorenz model. The Allen model not only conserves an energy, but also exhibits parcel conservation of potential vorticity. Under certain circumstances the Allen model produces spurious high frequency modes, but apparently these are easily suppressed. The Allen model has been used to produce high accuracy simulations of ocean flow over bottom topography.

In the Lorenz model only a single prognostic equation survives. Both the vorticity equation and the thermodynamic equation retain time derivatives, but the nonlinear balance equation relates these, making one into a diagnostic equation for the vertical velocity. The vorticity equation is generally taken to be the prognostic equation.

Charney (1962) used a different approach to integrate the nonlinear balance equations by taking the potential vorticity equation rather than the vorticity equation as the prognostic equation. The diagnostic equation relating the pressure or streamfunction to the potential vorticity is more complex than the equivalent relationship for the vorticity, which makes the mathematical solution more complex. However, the potential vorticity approach has the attractive feature of making the mathematical procedure congruent with the conceptual picture of the inversion process.

Davis and Emanuel (1991) have recently been able to invert the potential vorticity and nonlinear balance equations in a diagnostic study of midlatitude cyclonic systems, and found that the diagnosed winds closely matched the observed winds. McIntyre and Norton (1992) showed how to invert a hierarchy of balance conditions for the shallow water equations on a hemisphere, the lowest order equations being equivalent to the Charney model. The inversion was used in a prognostic calculation, and the results compared favourably with equivalent primitive equation results, even for cases with strong vortices in the tropics.

Eliassen (1952) developed what is basically a nonlinear balance model for an axisymmetric vortex,

and showed how the vortex would evolve under the application of heat and friction. Versions of this model have been invoked many times, particularly in the study of hurricanes, e. g., Sundqvist (1970), Rosenthal (1971), Willoughby (1979), Smith (1981), Shapiro and Willoughby (1982), Schubert and Hack (1982), Thorpe (1985), etc. Shutts and Thorpe (1978) explored the overturning of a vortex with unstable stratification using this model.

Craig (1991) has recently developed a theory for small deviations from an Eliassen balanced vortex. The essential assumption of the theory is that the ratio of the radial and tangential velocity scales is much less than unity. The model equations conserve both potential vorticity and energy.

Raymond and Jiang (1990) (hereafter RJ) proposed that the essential features of long-lived mesoscale convective systems could be represented by the nonlinear balance equations, and suggested a mechanism by which such systems could regenerate themselves after all convection has died. The dynamical heart of such a system was conjectured to be a midlevel potential vorticity anomaly produced by convective heating in the early stages of the mesoscale convective system. Subsequent lifting of low level air and the regeneration of convection was hypothesised to be due to the interaction of the potential vorticity anomaly and the low level environmental shear.

RJ developed a method for solving the nonlinear balance equations for small deviations from quasi-geostrophic theory. However, a limitation of RJ's model is the absence of advection by the unbalanced part of the wind. Comparison of the results of semigeostrophic and quasigeostrophic theory, particularly on small horizontal scales, suggests that this limitation is quite serious.

In this paper I develop two nonlinear balance models from two different scale analyses of the divergence, potential vorticity, and potential temperature equations. Each model is the leading order approximation to an expansion in small vertical velocity. Since the potential vorticity equation is used, parcel conservation of potential vorticity is automatic, and the existence of a potential enstrophy invariant

is assured.

Quasibalanced scaling leads to model equations that are equivalent to quasigeostrophic theory with geostrophic balance replaced by nonlinear balance. The small dimensionless parameter in this case turns out to be the Froude number. At large Rossby number the model reduces to the barotropic model.

Semibalanced scaling leads to the approximate nonlinear balance analogue of the semigeostrophic equations. The fundamental assumption behind semibalanced scaling is that the balanced advections of potential temperature and potential vorticity are approximately cancelled by the respective time derivatives in the equations for these two quantities. This happens exactly when the potential vorticity distribution exhibits slab or axial symmetry – e. g., in the case of an isolated front or vortex. Semibalanced scaling should therefore be applicable when strong vortices or fronts weakly interact with each other or with an environmental flow.

In semibalanced scaling the exact potential temperature and potential vorticity equations are retained with the exception that the definition of Ertel potential vorticity is modified to incorporate only the balanced part of the absolute vorticity. Balance models retain only one component of the vorticity equation as a prognostic equation. The semibalanced model uses the potential vorticity equation as its prognostic equation, which is equivalent to using the component of the vorticity equation normal to isentropic surfaces. This is the only way in which the semibalanced model differs from the Lorenz model, which uses the vertical component of the vorticity equation.

Neither the quasibalanced nor the semibalanced model exactly conserves energy, though of course both do so asymptotically in the limit of small expansion parameter. However, the semibalanced model satisfies all four of the criteria posed by Gent and McWilliams (1984) for selecting an appropriate intermediate model, i. e., one that falls between quasigeostrophy and the primitive equations in accuracy:

1. The semibalanced model uses the exact hydrostatic, continuity, and heat equations.

2. It uses the nonlinear balance condition, which is an improvement on geostrophic balance.

3. It has an integral invariant, namely the potential enstrophy.

4. It results from a minimal truncation of the divergence and potential vorticity equations.

Semibalance differs from the Lorenz equation in these criteria in that the latter has energy as the integral invariant. The Allen (1991) model has two integral invariants, but exhibits added complexity in the balance equation. Quasibalance fails the first criterion, because a truncated heat equation is employed. It nevertheless may be useful under certain circumstances.

Reasonably efficient numerical techniques for solving both sets of model equations are demonstrated, though the emphasis is on the solution of the semibalanced equations.

In order to establish a basis for potential vorticity thinking on small space and time scales, it is necessary to build up a catalogue of solutions to simple problems. Three types of solutions are useful in this regard:

1. Given a distribution of potential vorticity, what does the invertibility principle imply about the distributions of fluid velocity, and the perturbation potential temperature and pressure?

2. In the absence of external forcing, how does the potential vorticity distribution evolve?

3. How do externally applied forces and heat sources modify the potential vorticity distribution?

I attempt to answer these questions for a simple set of cases. In order to understand the nonlinear balance condition at high Rossby numbers in its purest form, most of the cases studied have environmental rotation suppressed.

Section 2 of this paper discusses the basic equations and their inversion, while section 3 presents the numerical methods used. Sections 4 and 5 respectively discuss unforced and forced flows, and conclusions are summarised in section 6.

2 Basic equations and scale analysis

I take as fundamental the divergence equation,

$$\frac{d\delta}{dt} + \delta^2 \; - \; 2\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} \; - \; \frac{\partial u}{\partial y}\frac{\partial v}{\partial x}\right) + \left(\frac{\partial w}{\partial x}\frac{\partial u}{\partial z} + \frac{\partial w}{\partial y}\frac{\partial v}{\partial z}\right) + \nabla_h^2(\theta_0\pi' \; - \; f\psi) \; - \; \nabla_h \cdot \mathbf{F} = 0 \; , \tag{1}$$

the potential temperature and potential vorticity conservation equations,

$$\frac{d\theta}{dt} = H \; , \tag{2}$$

$$\frac{dq}{dt} = \; - \; \rho_0^{-1}\nabla \cdot \mathbf{Y} \; , \tag{3}$$

where

$$\mathbf{Y} = \; - \; H\zeta + \nabla\theta \; \times \; \mathbf{F} \tag{4}$$

is the nonadvective flux of potential vorticity substance; the definition of potential vorticity,

$$q = \frac{1}{\rho_0}\left[(f + \nabla_h^2\psi)\left(\Gamma + \frac{\partial\theta'}{\partial z}\right) \; - \; \left(\frac{\partial v}{\partial z}\frac{\partial\theta'}{\partial x} \; - \; \frac{\partial u}{\partial z}\frac{\partial\theta'}{\partial y}\right)\right] \; , \tag{5}$$

the mass continuity equation,

$$\rho_0\delta + \frac{\partial\rho_0 w}{\partial z} = 0 \; , \tag{6}$$

and the hydrostatic equations for the mean and perturbation fields:

$$\theta_0\frac{\partial\pi_0}{\partial z} = \; - \; g \; , \qquad \theta_0\frac{\partial\pi'}{\partial z} = \frac{g\theta'}{\theta_0} \; . \tag{7}$$

The meanings of symbols are given in Table 1.

The following relations also hold:

$$u = \; - \; \frac{\partial\psi}{\partial y} \; - \; \frac{\partial\phi}{\partial x} \; , \qquad v = \frac{\partial\psi}{\partial x} \; - \; \frac{\partial\phi}{\partial y} \; ; \tag{8}$$

Symbol	Meaning
x, y, z	spatial coordinates
t	time
u, v, w	velocity components
$u_0(z)$	ambient wind in x direction
$\zeta = (\zeta_x, \zeta_y, \zeta_z)$	absolute vorticity
δ	horizontal divergence
ψ	horizontal streamfunction
ϕ	horizontal velocity potential
χ	integrated velocity potential
$\theta = \theta_0(z) + \theta'$	potential temperature
$\Gamma = d\theta_0/dz$	ambient potential temperature gradient
$N^2 = g\Gamma/\theta_0$	Brunt frequency squared
$U,\ S,\ C$	constants defining ambient wind in troposphere
T	absolute temperature
$\pi = \pi_0(z) + \pi' = C_p T/\theta$	Exner function
Σ	geostrophic deviation
$\rho_0(z)$	ambient air density
q	potential vorticity
$q_0 = f\Gamma/\rho_0$	planetary potential vorticity
H	potential temperature source term
$\mathbf{F}$	horizontal force (e. g., friction)
$\mathbf{Y}$	potential vorticity flux
ϵ	expansion parameter for scaling
F	Froude number
R	Rossby number
S	additional dimensionless parameter
∇_h, ∇_h^2	horizontal gradient and Laplacian
∇, ∇^2	3-D gradient and Laplacian
f	Coriolis parameter (constant)
g	acceleration of gravity
C_p	specific heat of air at constant pressure
$\Delta x,\ \Delta y,\ \Delta z$	sizes of grid cells
Δt	time step
$N_x,\ N_y,\ N_z$	numbers of grid cells in each dimension
M	linear density (per grid cell) of MAC particles
a	fraction of old ψ estimate mixed in

Table 1: Definitions of symbols, excluding scaling parameters.

$$\zeta_x = -\frac{\partial^2\psi}{\partial x\partial z} + \frac{\partial^2\phi}{\partial y\partial z}\ , \qquad \zeta_y = -\frac{\partial^2\psi}{\partial y\partial z} - \frac{\partial^2\phi}{\partial x\partial z}\ , \qquad \zeta_z = \nabla_h^2\psi + f\ ; \tag{9}$$

$$\delta = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = -\nabla_h^2\phi\ . \tag{10}$$

Several minor approximations have been made in deriving (1) - (7) from the primitive equations. Aside from the hydrostatic and f-plane approximations, the term $\theta'\nabla\pi'$ is ignored in the momentum equation, and the density is replaced by its mean value ρ_0 in several places. These approximations are justified on the mesoscale, where fractional variations in θ and ρ at a given level are small.

I now make a partial scale analysis of (1) - (10), with the following scaling assumptions:

Horizontal scale: L

Vertical scale: Z

Horizontal velocity: V

Streamfunction: VL

Time: L/V

Vertical velocity: $\epsilon ZV/L$

Divergence: $\epsilon V/L$

Velocity potential: ϵVL

Vertical gradient of ambient potential temperature: G

Note that the vertical velocity w, the horizontal divergence δ, and the velocity potential ϕ are all assumed proportional to a dimensionless parameter ϵ which is taken to be much less than unity. As indicated by Lilly (1983), the key assumption with this scaling is that the time scale is that of advection rather than

that of gravity waves. The assumption that $\epsilon \ll 1$ means that circulations are nearly horizontal, i. e., the vertical velocity is much less than that naively expected from a scale analysis of the continuity equation.

Ignoring terms containing ϵ in the divergence equation yields the nonlinear balance equation,

$$\nabla_h^2 \Sigma + 2 \left[\left(\frac{\partial^2 \psi}{\partial x \partial y} \right)^2 - \frac{\partial^2 \psi}{\partial x^2} \frac{\partial^2 \psi}{\partial y^2} \right] = \nabla_h \cdot \mathbf{F} \ , \tag{11}$$

where a new variable, the *geostrophic deviation*, has been defined:

$$\Sigma = \theta_0 \pi' - f\psi \ . \tag{12}$$

Two alternate ways of completing a scale analysis valid for all Rossby number lead to two different sets of equations which I call the quasibalanced and semibalanced equations, as discussed in the introduction. Both of these are leading order approximations in small ϵ.

If the potential temperature perturbation θ' is assumed to scale as $\mathrm{S}\epsilon GZ$ where $\mathrm{S} = \max(1, \mathrm{R}^{-1})$ and $\mathrm{R} = V/(Lf)$ is the Rossby number, then leading term approximations to (2), (3), and (5) valid for all Rossby number, small ϵ, and $\mathrm{S}\epsilon$ order unity or less are

$$\frac{\partial \theta'}{\partial t} - \frac{\partial \psi}{\partial y} \frac{\partial \theta'}{\partial x} + \frac{\partial \psi}{\partial x} \frac{\partial \theta'}{\partial y} + w\Gamma = H \ , \tag{13}$$

$$\frac{\partial q}{\partial t} - \frac{\partial \psi}{\partial y} \frac{\partial q}{\partial x} + \frac{\partial \psi}{\partial x} \frac{\partial q}{\partial y} + w \frac{dq_0}{dz} = - \rho_0^{-1} \nabla \cdot \mathbf{Y} \ , \tag{14}$$

$$q = \frac{1}{\rho_0} \left[(f + \nabla_h^2 \psi)\Gamma + f \frac{\partial \theta'}{\partial z} \right] \ , \tag{15}$$

where $q_0 = f\Gamma/\rho_0$. The last term on the left side of (14) must be included even though it contains w, which is of order ϵ, since it also scales with S, and thus becomes important at small Rossby number. The last term on the right side of (15) is retained for the same reason.

In order for the pressure perturbation to be consistently scaled between (7), (11), and (12), ϵ must equal the Froude number squared,

$$\epsilon = \mathrm{F}^2 \equiv \left(\frac{V}{NZ} \right)^2 \ , \tag{16}$$

where N is the Brunt frequency. I call this scaling *quasibalanced* since dropping the nonlinear terms from (11) results in the quasigeostrophic equations (Charney and Stern, 1962). (The quasigeostrophic potential vorticity differs from (15), but may be obtained by eliminating w between (13) and (14).)

An alternate form of scaling results when θ' is potentially of the same order of magnitude as GZ for all Rossby number, but the *balanced parcel tendency* of θ',

$$\frac{D\theta'}{Dt} \equiv \frac{\partial\theta'}{\partial t} - \frac{\partial\psi}{\partial y}\frac{\partial\theta'}{\partial x} + \frac{\partial\psi}{\partial x}\frac{\partial\theta'}{\partial y} , \tag{17}$$

is first rather than leading order in ϵ, i. e., scales as $\epsilon\theta' V/L$. This occurs when the flow is a small deviation from an axisymmetric circulation (as discussed by Craig, 1991) or from a slab symmetric flow, e. g., a front, since these flows have $D\theta'/Dt$ precisely equal to zero. For such flows $Dq/Dt = 0$ as well, which means that dq/dt scales as $\epsilon qV/L$ rather than qV/L when deviations from symmetry are weak.

For small ϵ, the leading order approximations to equations (2), (3), and (5) are

$$\frac{\partial\theta'}{\partial t} + u\frac{\partial\theta'}{\partial x} + v\frac{\partial\theta'}{\partial y} + w\left(\Gamma + \frac{\partial\theta'}{\partial z}\right) = H , \tag{18}$$

$$\frac{\partial q}{\partial t} + u\frac{\partial q}{\partial x} + v\frac{\partial q}{\partial y} + w\frac{\partial q}{\partial z} = -\rho_0^{-1}\nabla\cdot\mathbf{Y} , \tag{19}$$

$$q = \frac{1}{\rho_0}\left[(f+\nabla_h^2\psi)\left(\Gamma + \frac{\partial\theta'}{\partial z}\right) - \left(\frac{\partial^2\psi}{\partial x\partial z}\frac{\partial\theta'}{\partial x} + \frac{\partial^2\psi}{\partial y\partial z}\frac{\partial\theta'}{\partial y}\right)\right] . \tag{20}$$

I call this *semibalanced* scaling in analogy with semigeostrophy, since unbalanced advections of potential temperature and potential vorticity are included.

For semibalanced scaling, ϵ is not related to the Froude number as in quasibalanced scaling. Thus, flows with Froude number of order unity can be treated as long as the flow remains quasihorizontal. It is clear, e. g., in the case of an instability where small deviations from symmetry grow, that ϵ may increase with time for semibalanced scaling. Thus, the assumption that $\epsilon \ll 1$ must be continually checked when computing the evolution of a flow with the semibalanced equations.

The scalings of the forcing terms, H and $\mathbf{F}$, have yet to be considered. However, since $d\theta/dt$ scales as $\epsilon\Gamma ZV/L$ in both models, H must scale this way also. This is also consistent with the assumed scaling of dq/dt. Examination of (3) and (4) further indicates that $\mathbf{F}$ scales as $\epsilon \mathrm{S} V^2/L$. For the balanced models to remain valid, H and $\mathbf{F}$ technically should remain within the magnitudes specified by the above scaling. Adherence to the scaling on $\mathbf{F}$ obtained from the potential vorticity equation insures that $\mathbf{F}$ will not overwhelm the nonlinear balance equation, (11), as well.

It is interesting to speculate on the effects of forcing stronger than allowed by the above scale analysis. Heating in particular can easily induce strong vertical circulations which violate the model assumptions. However, in the semibalanced model, the instantaneous effects of these forcing terms on the potential vorticity distribution are exactly represented. As long as advection by the unbalanced part of the velocity, which may be poorly represented during the forcing, doesn't seriously distort the potential vorticity distribution, and as long as the unbalanced part of the flow doesn't itself modify the forcing, the flow after the forcing ceases should be accurate.

3 Numerical solution

The numerical method used to solve the equations is a descendant of that used by RJ. The multigrid method (Briggs, 1987) is invoked to solve the various elliptic equations. Three main differences exist between the present method and RJ's. First, the potential vorticity equation is inverted for the streamfunction rather than the Exner function perturbation, and the latter is obtained from the nonlinear balance equation. This has the advantage of working uniformly in Rossby number, i. e., it doesn't fail when the Coriolis parameter goes to zero. It represents a departure from the conceptual picture of the invertibility principle on large scales, wherein the pressure perturbation is obtained directly from the potential vorticity distribution, and the streamfunction from the pressure perturbation.

The second difference is that unbalanced horizontal velocities are included (at least in the semibal-

anced model) in the advection of potential vorticity and surface potential temperature. The unbalanced parts of the horizontal velocity and the vertical velocity are computed diagnostically from the potential temperature equation for $z > 0$ as follows: As shown by Lorenz (1960), an integrated velocity potential

$$\chi = \int_0^z \rho_0 \phi \, dz' \tag{21}$$

may be defined. Both unbalanced horizontal velocities and the vertical velocity may be obtained from χ:

$$\phi = \frac{1}{\rho_0} \frac{\partial \chi}{\partial z} \, , \tag{22}$$

and

$$w = \rho_0^{-1} \nabla_h^2 \chi \, . \tag{23}$$

The potential temperature equation may then be written for the semibalanced scaling as

$$\left(\Gamma + \frac{\partial \theta'}{\partial z} \right) \nabla_h^2 \chi \; - \; \frac{\partial \theta'}{\partial x} \frac{\partial^2 \chi}{\partial x \partial z} \; - \; \frac{\partial \theta'}{\partial y} \frac{\partial^2 \chi}{\partial y \partial z} = B \, , \tag{24}$$

where

$$B = \; - \; \rho_0 \left(\frac{\partial \theta'}{\partial t} \; - \; \frac{\partial \psi}{\partial y} \frac{\partial \theta'}{\partial x} + \frac{\partial \psi}{\partial x} \frac{\partial \theta'}{\partial y} \; - \; H \; \right) \tag{25}$$

is computable once the streamfunction and geostrophic deviation are known. Backward differencing is used to approximate the time derivative of potential temperature in (25). For the quasibalanced equations, (24) simplifies to

$$\Gamma \nabla_h^2 \chi = B \, . \tag{26}$$

Three elliptic equations must be inverted to solve the balance equations, the potential vorticity equation (15) or (20), the nonlinear balance equation (11), and the potential temperature equation (24) or (26). Multigrid coarsening is done on all three dimensions for the potential vorticity and nonlinear balance equations, which are solved simultaneously. The nonlinear terms in these equations are allowed to vary during relaxation. Two dimensional coarsening is done for the potential temperature equation, which is

solved separately after the potential vorticity and nonlinear balance equations are solved. Because of the nonlinearity of the problem, underrelaxation is sometimes needed for convergence while solving for the streamfunction and geostrophic deviation. In addition, after each V cycle, it is sometimes necessary to apply the low pass filter

$$\psi_{i+1} = a\psi_i + (1 - a)\psi_{i+1} \ , \tag{27}$$

where ψ_i is the estimate of the streamfunction after the ith V cycle and $0 \leq a \leq 0.5$, to keep the solution from oscillating. Convergence was deemed to be sufficient when the volume-integrated value of $(\psi_{i+1} - \psi_i)^2$ reached 10^{-4} times its original value.

The final major difference between RJ and the present model is that the marker and cell (MAC) method (Harlow and Welch, 1965; Daly, 1967) is used to solve the potential vorticity and surface potential temperature advection equations. This is done because of its nondiffusive properties. In the MAC method, the trajectories of a swarm of imaginary particles are obtained by time integrating the fluid velocity interpolated to the positions of the particles. Each particle carries the potential vorticity (and potential temperature at the surface) existing at the initial position of the particle at the initial time. The potential vorticity is allowed to vary according to the integrated effect of the right side of (14) or (19) interpolated to the particle trajectory. Changes in the particle's potential temperature are computed in a similar fashion at the surface. A second order half step-full step scheme is used in the time integration. The only restriction on the time step is that the Courant condition be satisfied, i. e., $\Delta x/\Delta t > \max(|v_x|)$ and $\Delta y/\Delta t > \max(|v_y|)$.

Only the horizontal part of the advection is computed using the MAC method. The vertical advection of potential vorticity is weak in models with small ϵ, and is done with upstream differencing. This is, of course, diffusive, but the diffusion is negligible in the present case, and it has the advantage of not introducing any advective instabilities.

M^2 imaginary particles are put initially in each $x-y$ grid box at each level, where generally $M=3$. These particles are not centered in the $\Delta x/M$ by $\Delta y/M$ sub-box allocated to each particle, but are initially located randomly in each sub-box. This keeps nondivergent deformations from developing regular patterns of particle distribution that empty some boxes while overfilling others. Empty boxes can still be generated at large time by persistent horizontal divergence. This causes problems with the interpolation of potential vorticity and potential temperature back to the Eulerian grid.

Eulerian interpolation at each grid point is accomplished by fitting the potential vorticity of all particles at the level of interest and with x and y positions within 1.5 grid dimensions of the grid point of interest to a linear function of x and y. The value of this function at the grid point is taken as the interpolated value. Potential temperature is treated similarly at the surface. This procedure causes some initial smoothing in the Eulerian potential temperature distribution. However, the potential vorticity is the source term of a Poisson-like equation, and as pointed out by HMR, small scale perturbations in the potential vorticity are of little consequence because of this. Furthermore, the smoothing doesn't increase with time, because reinterpolation from the grid to the particles is never done.

Simulations were done in a domain with a flat bottom surface, a rigid upper lid, and solid north and south walls. As suggested by Gent and McWilliams (1983), free slip boundary conditions were imposed individually on the solenoidal and irrotational parts of the velocity at these walls. Periodic boundary conditions were applied in the east-west direction. The domain was divided vertically into a troposphere and a stratosphere at $z=z_t$, with separate lapse rates of ambient temperature. The ambient wind and temperature were assumed to be in thermal wind balance, which means that ambient tropospheric and stratospheric temperature profiles generally vary with y. The ambient wind profile in the troposphere was taken to be in the east-west plane, with the form

$$u_0(z) = U + Sz + Cz^2/2 \ . \tag{28}$$

(Note that the ambient shear, S is distinct from the dimensionless parameter S defined in section 2 and set in Roman type.) A similar quadratic polynomial for the stratosphere was adjusted so that ambient wind and shear were continuous at the tropopause and caused no ambient horizontal potential temperature gradient at domain top. This prevented spurious surface Rossby waves from occurring there.

In all simulations, either an initial potential vorticity anomaly is imposed, or distributions of heating, H, and force, $\mathbf{F}$, are allowed to create potential vorticity anomalies. All such distributions take the form

$$A = A_{max} P[(x - x_l)^2/x_0^2 + (y - y_l)^2/y_0^2 + (z - z_l)^2/z_0^2] \tag{29}$$

for any variable A with maximum value A_{max}, where

$$P(\mu) = \begin{cases} 1 - \mu , & \mu \leq 1 \\ 0 , & \mu > 1 , \end{cases} \tag{30}$$

(x_l, y_l, z_l) are the Cartesian coordinates of the center of the distribution, and (x_0, y_0, z_0) are the half widths of the distribution in the three cardinal directions.

4 Unforced flows

In this section I present the results of semibalanced simulations with forcing turned off. Potential vorticity is conserved, and is simply advected around the domain. Considered first are the flow and thermodynamic fields resulting from an isolated, axially symmetric potential vorticity anomaly in an unsheared, nonrotating environment. Next, the effect of shear on vortices is investigated. This is the situation envisioned by RJ in their model of mesoscale convective systems. In the final subsection the interaction of a pair of potential vorticity anomalies of opposite sign is studied. This leads to a solution that is very close to Drazin's (1961) solution for flow past a three dimensional obstacle at low Froude number.

The arena in which the simulations in this section occur is an 800 km by 800 km by 16 km domain with $\Delta x = \Delta y = 25$ km, and $\Delta z = 1$ km. The tropopause is at 10 km. In the center of the domain the ambient temperature lapse rate is -7 K km^{-1} in the troposphere, while the stratosphere is isothermal.

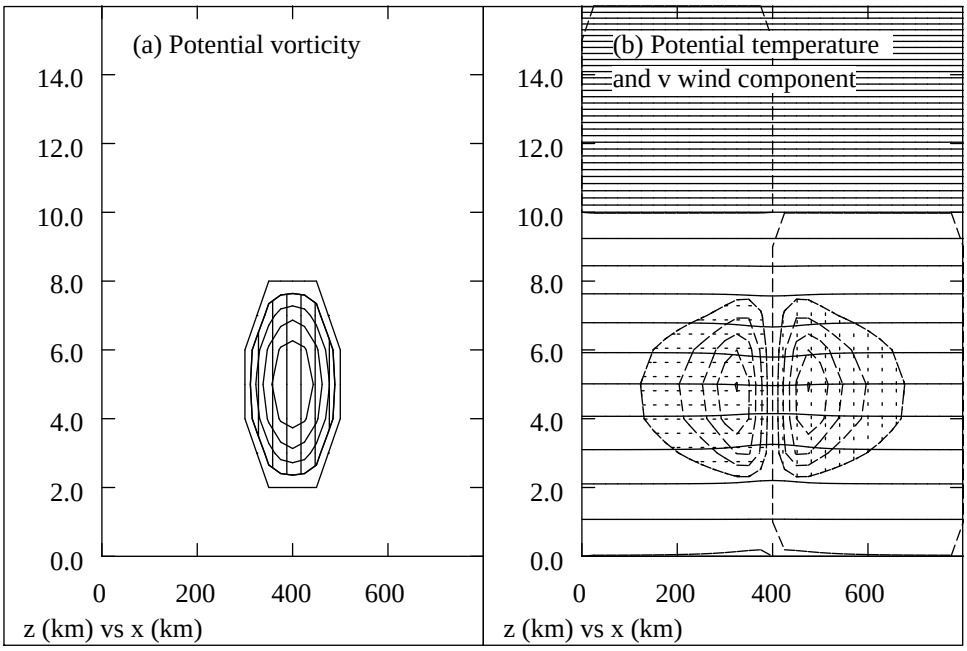

Figure 1: East-west cut through the center of an isolated potential vorticity anomaly of magnitude 1 pvu in an unsheared environment with no rotation. (a) Contours of potential vorticity at 0.2 pvu intervals, with hatching where the potential vorticity exceeds 0.2 pvu. (b) Solid contours are of potential temperature at 3 K intervals, while dashed contours show the wind component normal to the page at 1 m s^{-1} intervals. Vertical and horizontal hatching indicate flow in or out of the page less than 1 m s^{-1}.

a. Isolated vortices

In the absence of Coriolis force, the balance in a vortex is purely between pressure gradient force and the centripetal acceleration. Figure 1 shows a semibalanced simulation of an isolated vortex in an unsheared, nonrotating environment. In this simulation, $x_0 = y_0 = 100$ km, $z_0 = 3$ km, $z_l = 5$ km, and the maximum value of the potential vorticity is $q_{max} = 1$ pvu (1 pvu = 10^{-6} m^2 kg^{-1} s^{-1} K). The Froude number in this case is F = 0.17, where the vertical scale is taken to be z_0, and the velocity is the maximum rotational velocity of 5 m s^{-1}. The Rossby number is, of course, infinite. (The Froude number is a useful dimensionless parameter for semibalanced simulations even though it is not equal to ϵ.)

As figure 1 indicates, the horizontal extent of the vortex is greater than that of the potential vorticity anomaly. This, of course, is also true in the low Rossby number limit. However, at large horizontal

distances from the anomaly center, the tangential velocity varies as the inverse of radius in the infinite Rossby number case, whereas the dependence is inverse square at low Rossby number. This is easily seen by noting that the quasibalanced theory is valid at large radius even for strong vortices, resulting in $\nabla_h^2 \psi = \rho_0 q/\Gamma$ for $f = 0$. This yields a logarithmic variation of streamfunction with horizontal radius, and therefore inverse radius behaviour for the tangential velocity. Thus, at infinite Rossby number the specific angular momentum associated with the tangential circulation is constant with radius.

In figure 1 the vertical extent of the vortex and the associated potential temperature anomaly are almost completely confined to levels where the potential vorticity anomaly is nonzero. This is because the circulation around any loop contained in an isentropic surface with uniformly zero potential vorticity must be zero (excluding those surfaces penetrating a boundary), causing the flow on that surface to be irrotational. Thus, only isentropic surfaces cutting through the potential vorticity anomaly can exhibit vortex flow at infinite Rossby number. (Ronald Smith, personal communication.) Hence, vertical influence is precisely zero in isentropic coordinates. This is in contrast to a rotating environment in which the potential vorticity anomaly has a nonzero vertical as well as horizontal radius of influence.

HMR showed that a surface temperature anomaly mimics the effects of a very thin potential vorticity anomaly located just above the surface. In the infinite Rossby number case, such anomalies have a more limited dynamical significance than at low Rossby number, since the lack of background vorticity eliminates the vertical penetration of the induced circulation in isentropic coordinates. They are thus unable to interact with potential vorticity anomalies aloft that are located on isentropic surfaces which don't intersect the surface anomaly.

Figure 2 shows the results of a simulation that is identical to that in figure 1, except that the potential vorticity anomaly is 4 times as strong. Notice that the maximum rotational velocity is now 14 m s^{-1}, resulting in a Froude number F = 0.47. The circulation in the first simulation is sufficiently weak that

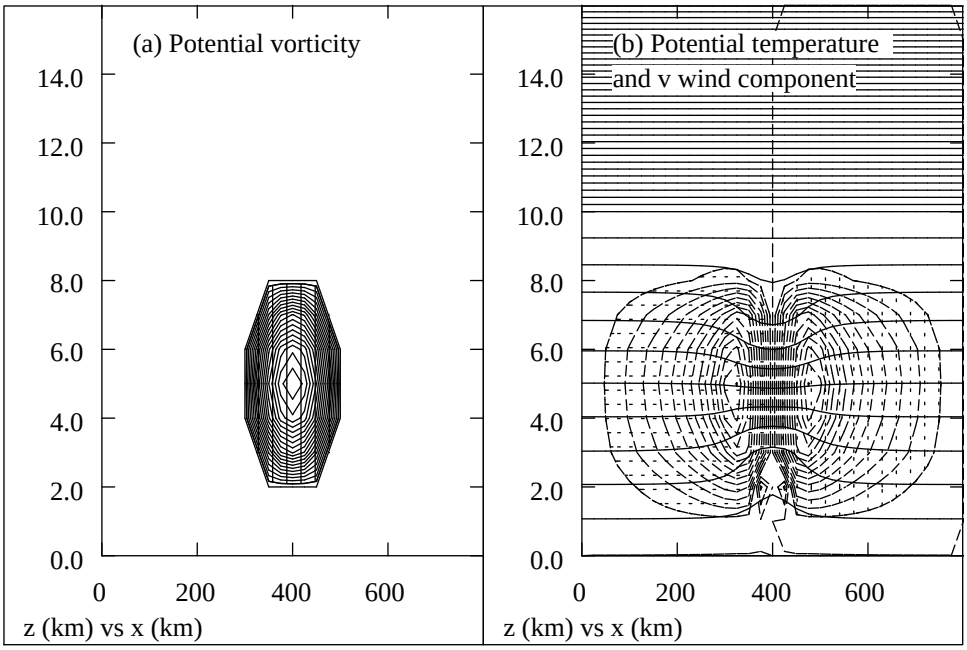

Figure 2: East-west cut through the center of an isolated potential vorticity anomaly of magnitude 4 pvu in an unsheared environment with no rotation. (a) Contours of potential vorticity at 0.2 pvu intervals, with hatching where the potential vorticity exceeds 0.2 pvu. (b) Solid contours are of potential temperature at 3 K intervals, while dashed contours show the wind component normal to the page at 1 m s^{-1} intervals. Vertical and horizontal hatching indicate flow in or out of the page exceeding 1 m s^{-1}.

the quasibalanced model would have produced an accurate result, whereas this is not true of the second case. This is reflected in the values of the Froude number, which is almost 3 times greater in the second case than in the first.

Since the Froude number depends on the velocity scale, it can only be computed a posteriori in this case. However, F $\approx$ 1 implies that $V^2 \approx N^2 Z^2$. In the case of infinite Rossby number the potential vorticity anomaly scales as $q \approx \Gamma V/(\rho_0 L)$. Eliminating V between these two equations results in a scaling parameter for the potential vorticity anomaly, Q, of the form

$$Q = \frac{\Gamma N Z}{\rho_0 L} \ . \tag{31}$$

$Q \approx$ 1 pvu in these two cases. Semibalanced results differ significantly from quasibalanced results only when $q_{max} \gg Q$ for an isolated potential vorticity anomaly in a stationary, nonrotating environment.

b. Vortices in shear

The above simulations exhibit no vertical velocity of any significance. (A slight vertical velocity results from the interaction of the vortex with its images induced by the lateral boundary conditions.) As is well known, a potential vorticity anomaly in a sheared, rotating environment has upward motion on its downshear side. RJ suggested that this results from (a) vortex-induced motion up or down the tilted ambient isentropic surfaces, and (b) ambient motion up or down isentropic surfaces distorted by the potential vorticity anomaly. Third and fourth possibilities not mentioned by RJ are (c) the time-dependent lifting or descent of isentropic surfaces resulting from shear-induced changes in the potential vorticity distribution, and (d) vortex-induced motion up and down isentropic surfaces distorted by the potential vorticity anomaly.

In a nonrotating environment, (a) cannot act. However, (b), (c), and (d) are all operating in the simulation shown in figures 3 and 4. In this simulation an initial potential vorticity anomaly identical to that shown in figure 2 is subjected to a sheared environment with wind profile parameters $U = -10$ m s^{-1},

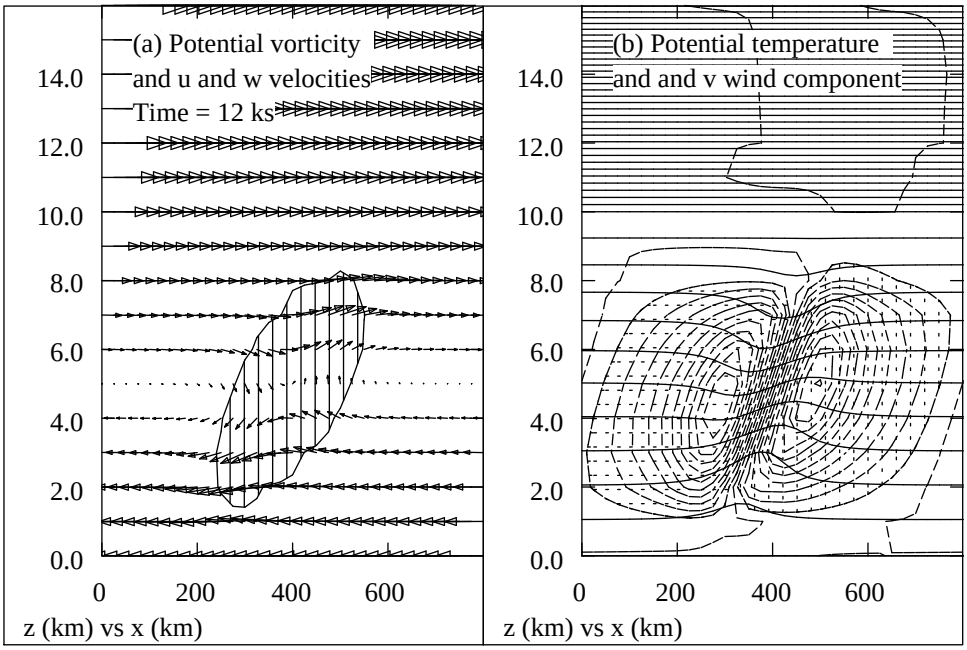

Figure 3: East-west cut through the center of a simulation like that in figure 2, except that the potential vorticity anomaly is allowed to evolve for 12 ks in an ambient shear flow $u_0(z) = -10 + 2z$, where z is in kilometers and u_0 in meters per second. (a) Regions with potential vorticity exceeding 0.2 pvu are hatched. The scale on the velocity vectors is 10 m s^{-1} per 100 km. (b) Solid contours are of potential temperature at 3 K intervals, while dashed contours show the wind component normal to the page at 1 m s^{-1} intervals. Vertical and horizontal hatching indicate flow in or out of the page exceeding 1 m s^{-1}.

$S = 2$ ks^{-1}, and $C = 0$. With these values, the center of the potential vorticity anomaly at 5 km experiences zero ambient wind, while material below this level is advected to the left, and material above to the right.

Figure 3 shows the wind and potential vorticity in a vertical plane parallel to the shear, passing through the center of the potential vorticity anomaly. Note that upward motion exists on the downshear side of the anomaly, as in the rotating case. In the upper and lower parts of the anomaly, this is mainly caused by ambient advection up and down distorted isentropic surfaces, as envisioned by RJ. However, there is still vertical motion at $z = 5$ km, even though the ambient flow is zero at this level. Figures 4a, 4b, and 4c reveal the origin of this vertical motion.

Figures 4a and 4b show the potential temperature perturbation induced by the potential vorticity anomaly in a horizontal plane at 5 km for $t = 0$ and $t = 12$ ks. At the initial time a weak, nearly symmetric warm anomaly is seen, while at 12 ks the region upshear of the vortex core is anomalously warm, and the region downshear is cool. This change is caused by the shearing of the potential vorticity anomaly by ambient flow, and is clearly seen in figure 3a. (The initial warm anomaly wouldn't have occurred if the Boussinesq approximation had been used, since the elevation of the display plane is vertically centred on the potential vorticity anomaly. Non-Boussinesq effects make the potential temperature distribution nonsymmetric in the vertical.)

The vertical velocity at 12 ks is shown in figure 4c. Comparison with figure 4b shows that the strongest vertical motion at 5 km is collocated with the strongest warm advection, and hence is caused by self-induced advection up and down the distorted isentropic surface, i. e., mechanism (d) above.

For purposes of comparison, a simulation identical to that discussed above is shown in figure 5, except that the Coriolis force is turned on. The vertical velocity in this case has the same general pattern, but is about twice as strong. Thus, the vortex-induced motion up and down ambient isentropic surfaces

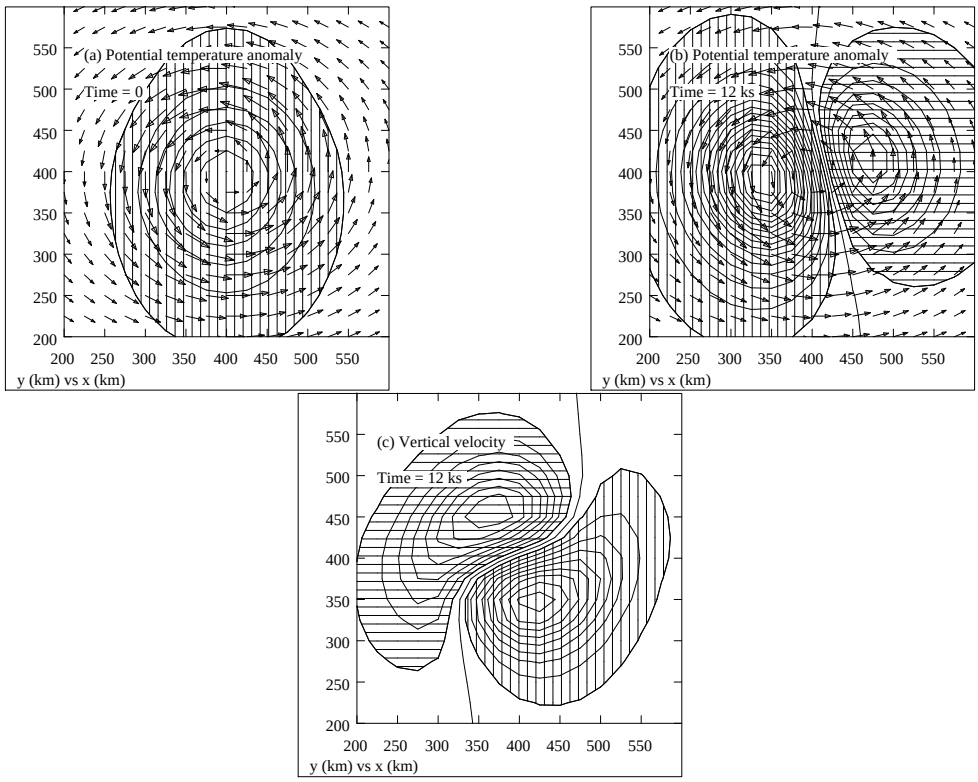

Figure 4: Horizontal cut through the simulation of figure 3 at $z = 5$ km. Only the central region of the simulation is shown. (a) Potential temperature anomaly and horizontal winds at $t = 0$. The contours of potential temperature anomaly are 0.1 K, and regions with the anomaly exceeding 0.1 K are hatched. The velocity vector scale is 10 m s^{-1} per 25 km. (b) Potential temperature anomaly and horizontal winds at $t = 12$ ks. Contours and hatching are the same as in (a), except that horizontal hatching indicates a potential temperature anomaly less than -0.1 K. (c) Vertical velocity. Contour interval is 0.005 m s^{-1}. Regions with upward motion greater than 0.005 m s^{-1} are vertically hatched, while those with downward motion exceeding -0.005 m s^{-1} are horizontally hatched.

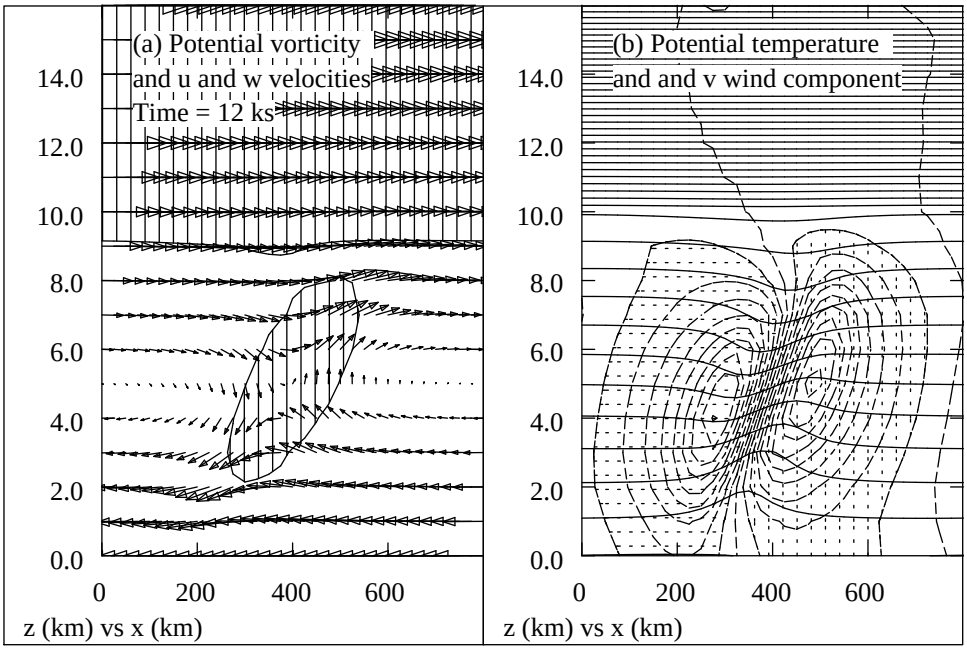

Figure 5: East-west cut through the center of a simulation like that in figure 3, except that the Coriolis parameter $f = 0.1$ ks. (a) Regions with potential vorticity exceeding 1 pvu are hatched. The scale on the velocity vectors is 10 m s^{-1} per 100 km. (b) Solid contours are of potential temperature at 3 K intervals, while dashed contours show the wind component normal to the page at 1 m s^{-1} intervals. Vertical and horizontal hatching indicate flow in or out of the page exceeding 1 m s^{-1}.

(mechanism a) must be about equal to all the other mechanisms in strength in this case.

The maximum vertical velocity that could be expected from the continuity equation here is of order VZ/L. Taking $L = x_0 = 100$ km, $Z = z_0 = 3$ km, and $V \approx 10$ m s^{-1} (estimated from figure 3), then $VZ/L \approx 0.3$ m s^{-1}. The actual maximum computed vertical velocity is about 0.05 m s^{-1}, so $\epsilon \approx 0.15$ at most, and semibalanced scaling is appropriate to this situation.

c. Vortex interactions and idealised flow around mountains

It is easy to demonstrate the usual types of interactions between barotropic vortices with this model. For instance, two horizontally adjacent potential vorticity anomalies of the same sign rotate around each other, whereas two of opposite signs move in a direction normal to the line between the anomalies. Distortion in the vertical structure of the potential vorticity anomalies occurs in these cases, as the induced velocities are not uniform with height.

An interesting potential use of the two vortex configuration is to explore low Froude number flow around an isolated mountain. However, as will be shown, there is a significant problem in this application that illustrates the limits of the semibalanced approximation.

If two potential vorticity anomalies of opposite sign are set on a line normal to the ambient wind, they will (if the signs are correct) tend to move upstream against the flow. For the proper anomaly strength, this motion can be made to just counteract the downstream advection, leaving the vortex pair stationary. The flow then resembles idealised flow around terrain similar to that demonstrated by Drazin (1961) using a different method of analysis.

Figure 6 shows the flow computed at the surface induced by two potential vorticity anomalies, one 50 km north of the center of the domain, with $q_{max} = 5$ pvu, the other 50 km south, with $q_{max} = -5$ pvu. The anomalies are centered at the surface, i. e., $z_l = 0$, and the dimensions are $x_0 = y_0 = 100$ km and $z_0 = 4$ km. In addition to the induced flows from the vortices, an ambient flow from right to left of

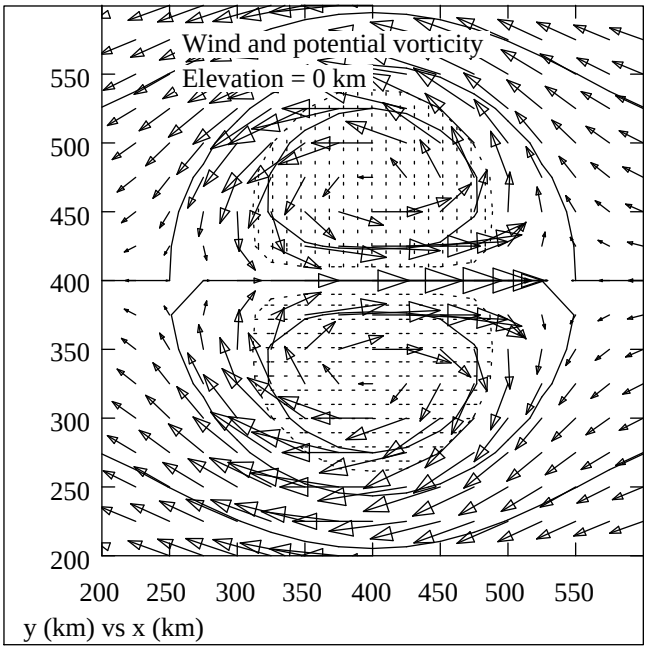

Figure 6: Flow at the surface due to two potential vorticity anomalies of opposite sign centered at the surface combined with a 10 m s^{-1} wind from the east. Only the central region of the simulation is shown. The velocity scale is 10 m s^{-1} per 25 km. Solid lines are contours of constant streamfunction. Vertical and horizontal hatching respectively show regions with potential vorticity greater than 1 pvu and less than -1 pvu. Environmental rotation is turned off.

10 m s^{-1} is superimposed. There is no environmental rotation in this case, and the semibalanced model is used. The Froude number is F = 0.25.

The flow is divided into three parts consisting of two counter-rotating vortices collocated with the potential vorticity anomalies and an external flow that passes around the vortices. The dividing line between the vortices and the external flow is nearly circular, and all regions of nonzero potential vorticity are confined within this circle.

Figure 6 represents a snapshot of the flow at the initial time. The parameters chosen for vortex strength and spacing cause the two vortices to remain approximately stationary relative to the surface, thus making the flow pattern roughly steady. However, an alternate interpretation of the results can be made that doesn't depend on the stationarity of the vortex pair if all potential vorticity is confined within the internal vortex flow. The boundary between the vortex flow and the external flow can then be interpreted as the surface of a three-dimensional mountain. The evolution of the internal flow then becomes irrelevant to the problem of flow around the mountain, as the initial external flow is just the flow that satisfies free slip boundary conditions at the mountain surface. This flow is steady if the potential vorticity outside the mountain is zero. We have already seen that potential vorticity is confined to the interior at $z = 0$. Closer inspection of the results shows that it is nearly confined at all levels, with only very minor penetration through the mountain surface at $z = 3$ km.

Figure 7 shows the elevation of the 303 K isentropic surface for this simulation. The ambient elevation of this surface is approximately 1.1 km, but its level is depressed to a minimum of 0.6 km and elevated to a maximum of 1.3 km near the potential vorticity anomalies. The higher elevations upstream and downstream of the mountain are correlated with lower wind speeds, while the lower elevations near the lateral flanks of the mountain are associated with higher wind speeds. These results are consistent with Drazin's (1961).

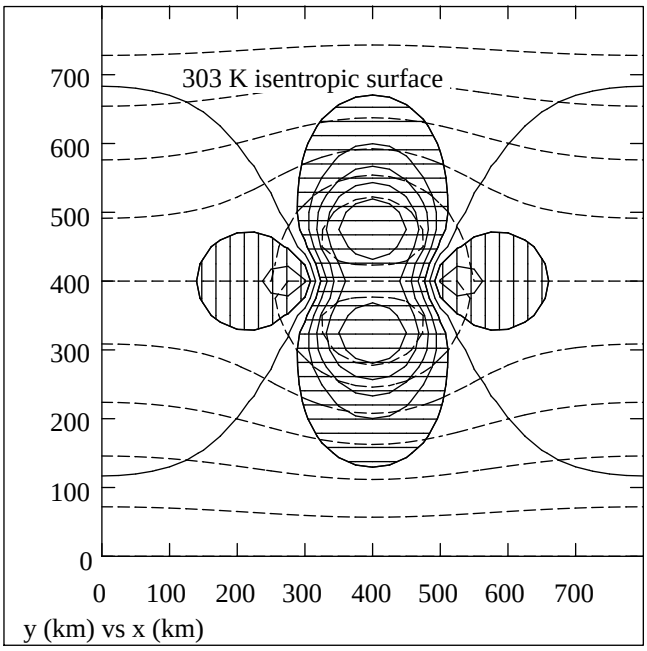

Figure 7: Plot of elevation of 303 K isentropic surface for simulation of figure 6. The ambient elevation of this surface is 1.1 km. Solid contours show elevation at 0.1 km intervals. Vertical hatching shows elevations greater than 1.2 km, while horizontal hatching indicates regions less than 1.0 km. Dashed contours indicate streamfunction.

Figure 8 shows the vertical structure of the flow along the east-west centerline. Note how the isentropic surfaces are elevated at all levels below the mountain top outside the mountain along this centerline. This is a nonlinear effect, since each potential vorticity anomaly by itself depresses the isentropic surfaces everywhere in the domain.

Comparison with Smolarkiewicz and Rotunno (1990) shows that the motions in the present model are much more like those of the Drazin model than those in the primitive equation results, and agreement with the primitive equation simulation of the above authors is quite poor. The primary difference between the balance and Drazin results on one hand and the primitive equation results on the other hand is the existence of upstream-downstream symmetry in the former. Not only are the lee vortices seen in the primitive equation results not present in the balanced results, but vertical motions upstream of the mountain are poorly represented as well.

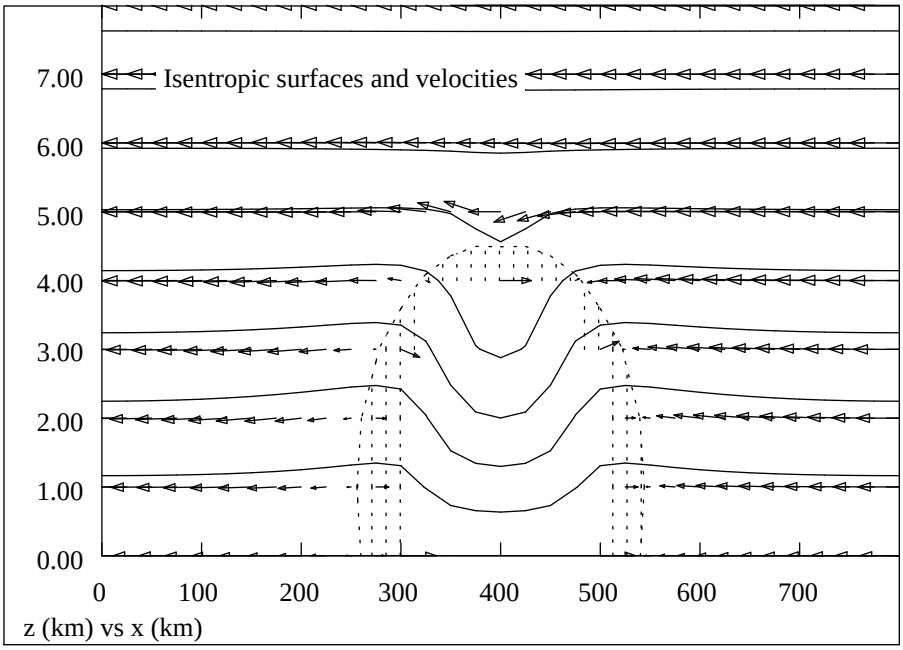

Figure 8: Vertical, east-west section through the center of the simulation shown in figure 6. The solid lines are contours of potential temperature at 3 K intervals, while the vectors show the $u - w$ velocity components. The outline of the "mountain" (actually, the region with reversed east-west velocities) is shown by the dashed lines. Very large velocity vectors in the interior of the mountain have been suppressed for clarity. The velocity scale is 5 m s^{-1} per 25 km.

Viewing this problem in the original context of two interacting potential vorticity anomalies explains why the semibalanced model results are so different from the primitive equation results. Though vertical velocities are very weak outside the boundaries of the "mountain", they reach 0.4 m s^{-1} in the interior. Taking $L = x_0 = 100$ km, $Z = z_0 = 4$ km, and $V = 10$ m s^{-1}, $ZV/L = 0.4$ m s^{-1}, and $\epsilon \approx 1$ there. Thus, the semibalanced model is outside of its range of validity, and the disagreement with primitive equation calculations is not surprising.

The above result suggests that $\epsilon \ll 1$ must be globally rather than locally satisfied for the semibalanced model to be valid.

5 Forced flows

Flows caused by diabatic heating and cooling, and by forces other than the Coriolis, pressure gradient, and gravitational forces are considered in this section. These forcing factors act through their effects on the potential vorticity distribution and the nonlinear balance equation.

As Haynes and McIntyre (1987) and others showed, the density-weighted volume integral of potential vorticity is conserved in the fluid interior, and can be changed only by processes at the fluid boundary. The role of forcing is therefore to nonadvectively transport potential vorticity substance around the fluid. The right sides of (14) and (19) are in the form of minus the divergence of a flux, $\mathbf{Y}$, divided by the density. Multiplying both sides of (14) and (19) by the density and integrating over volume verifies the above stated conservation law for both quasibalanced and semibalanced theory.

Haynes and McIntyre (1987) show in addition that the amount of potential vorticity substance between two isentropic surfaces remains constant. In their formula for the potential vorticity flux, this is obvious since all nonadvective fluxes are parallel to isentropic surfaces. However, since the nonadvective flux is only defined up to an additive nondivergent vector field, alternate forms exist that cause local transport across

isentropic surfaces while maintaining zero global transport, i. e., flux of potential vorticity substance across an isentropic surface is compensated by flux in the opposite direction elsewhere on the surface.

As defined here, the nonadvective transport of potential vorticity by diabatic heating is in the direction opposite the absolute vorticity vector, ζ. Thus, in an irrotational fluid, diabatic heating causes no potential vorticity transport, while in a fluid with vorticity but no initial potential vorticity, the transport is parallel to isentropic surfaces.

The form of the external force contribution to the potential vorticity flux presented here is different from that in RJ. The two forms are related by a vector identity, but the form used here is easier to interpret. Note that with the form used here, the potential vorticity flux is normal both to the applied force and the potential temperature gradient. Thus, external forces can never transport potential vorticity across isentropes.

Viscous forces are among those that can contribute to potential vorticity transport. We tend to think of such forces as being negligible in the atmosphere except very near boundaries, but the accumulated effect of viscous dissipation can have nontrivial effects, as I demonstrate in the last subsection.

a. Applied forces

As the first example of forced flow, I show the response of the semibalanced model to a force in the x-direction confined to a restricted region of fluid. The force has the form given by (29) with $x_0 = y_0 = 100$ km, $z_0 = 2$ km, $z_l = 5$ km, and $F_{max} = 1$ km ks^{-2}. For the sake of simplicity, the initial wind is zero, and there is no environmental rotation.

Figure 9 shows the resulting flow in a horizontal plane at $z = 5$ km and $t = 30$ ks. As (4) suggests, the eastward force causes a northward flux of potential vorticity, thus forming a positive anomaly on the north side of the forced region, and a negative anomaly on the south side. The induced flow from the two vortices advects potential vorticity to the east. Since potential vorticity is being continuously replenished

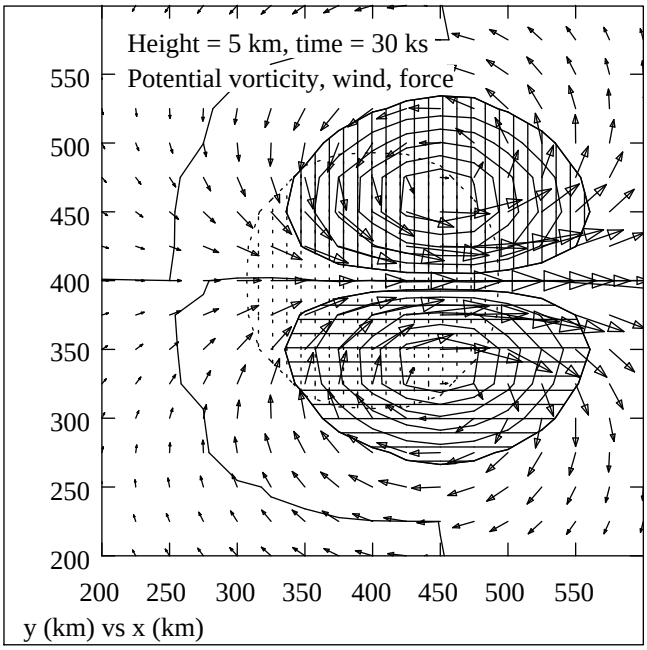

Figure 9: Horizontal section at 5 km through a flow created by an applied force 30 ks after it is turned on. Only the central region of the simulation is shown. The scale for the velocity vectors is 5 m s^{-1} per 25 km. Dashed hatching shows where the applied force per unit mass, F_x, exceeds 0.1 km ks^{-2}. The solid contours are of potential vorticity with a contour interval of 0.3 pvu. Solid vertical hatching shows regions with potential vorticity exceeding 0.3 pvu, while solid horizontal hatching indicates values less than -0.3 pvu.

by the forcing, this results in elongated patterns of potential vorticity and a westerly jet extending to the east of the forced region. Notice also that there is induced recirculation, with return flow on the north and south sides of the jet. This is partially the result of the lateral boundary conditions on the north and south edges of the computational domain, which keep the difference in the streamfunction between the two faces constant at each level. Thus, the average velocity at each level remains constant, requiring a return flow in response to the central jet.

The scaling parameter for force given at the end of section 2 is $\epsilon \mathrm{S} V^2/L$. Since there is no initial velocity in this simulation, the scaling parameter doesn't make any sense at the beginning of the calculation. However, at the end, a horizontal wind of about 8 m s^{-1} has been generated. Using the horizontal radius of the forcing, 100 km, as L, and realising that S = 1 in this case, the force should be much less than 0.64 km ks^2. Since the maximum applied force is 1 km ks^{-2}, the scaling assumptions have been formally violated even at the end of the calculation. However, at that point the maximum vertical velocity is approximately 0.016 m s^{-1}, while the maximum possible vertical velocity from mass continuity is $VZ/L \approx 0.16$ m s^{-1}. Thus, $\epsilon \approx 0.1$, and the conditions for the semibalanced approximation are well satisfied.

b. Diabatic heating

Equation (4) indicates that potential vorticity transport associated with diabatic heating is along the absolute vorticity vector. Figure 10 illustrates the result of diabatic heating in a sheared environment at middle latitudes. Assuming that thermal wind balance is valid in the environmental flow, the absolute vorticity vector is normal to the shear and slants upward toward the north (i. e., toward colder environmental air). The slope of the vorticity vector is just f/S, where the shear $S = du_0/dz$. If diabatic heating is confined to a region of width d (in a direction normal to the shear) and depth h, then potential vorticity is removed from the upper right corner of the heated region, creating a negative potential

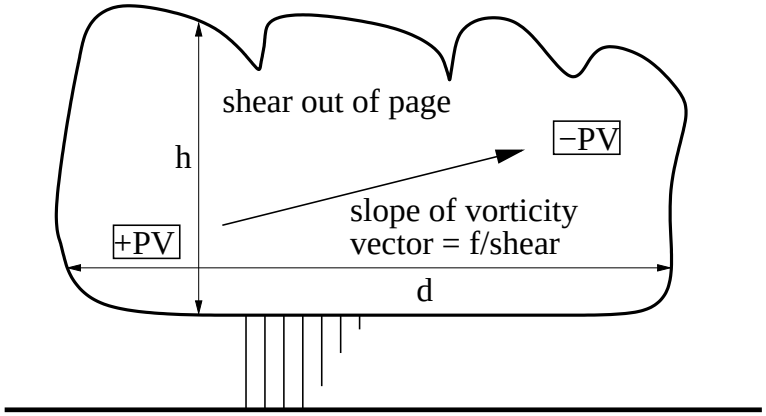

Figure 10: Sketch of potential vorticity transport in a heated region (a convective system is indicated for definiteness). The shear is out of the page, so the right side of the domain is cooler by the thermal wind relationship. The absolute vorticity vector slopes up to the right, and potential vorticity is transported from the upper right corner of the heated region to the lower left corner.

vorticity anomaly there, and deposited in the lower left corner, creating a positive anomaly.

The aspect ratio of the heated region, h/d, determines whether the vertical or horizontal component of the potential vorticity transport is more important. If $h/d \gg f/S$, then the transport is effectively horizontal, from the cold to the warm side of the heated region. On the other hand, the reverse condition makes the vertical transport dominate, resulting in a negative anomaly overlying a positive anomaly. It is this latter situation that was assumed by RJ in their analysis of mesoscale convective systems. If $S = 2 \times 10^{-3}$ s^{-1} and $f = 10^{-4}$ s^{-1}, then $d/h \gg 20$ for RJ's assumption to be valid. As typical values for d and h are 300 km and 10 km for mesoscale convective systems, this assumption is clearly marginal.

For small scale thunderstorms in shear, $d/h \ll 20$, and the transport of potential vorticity is essentially horizontal. The counter-rotating vortices that are commonly seen in simulations of such systems (e. g., Schlesinger, 1978) are manifestations of the potential vorticity anomalies created by such transport, though they are more commonly explained as the result of tilting of ambient vortex lines into the vertical.

Figure 11 shows the winds and potential vorticity resulting from heating a sheared, nonrotating environment with $u_0(z) = -6 + 2z$ (m s^{-1}) for 40 ks. The diabatic heating is horizontally centered

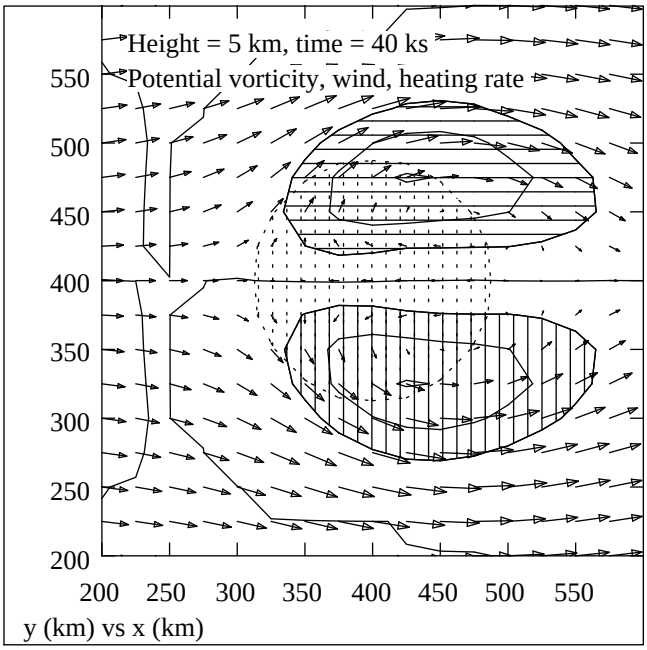

Figure 11: Horizontal section at 5 km through a flow created by diabatic heating of a sheared environment. Displayed are results 40 ks after the onset of the heating. Only the central region of the simulation is shown. The scale for the velocity vectors is 5 m s^{-1} per 25 km. Dashed hatching shows where the heating rate H exceeds 0.1 K ks^{-1}. The solid contours are of potential vorticity with a contour interval of 0.2 pvu. Solid vertical hatching shows regions with potential vorticity exceeding 0.2 pvu, while solid horizontal hatching indicates values less than -0.2 pvu.

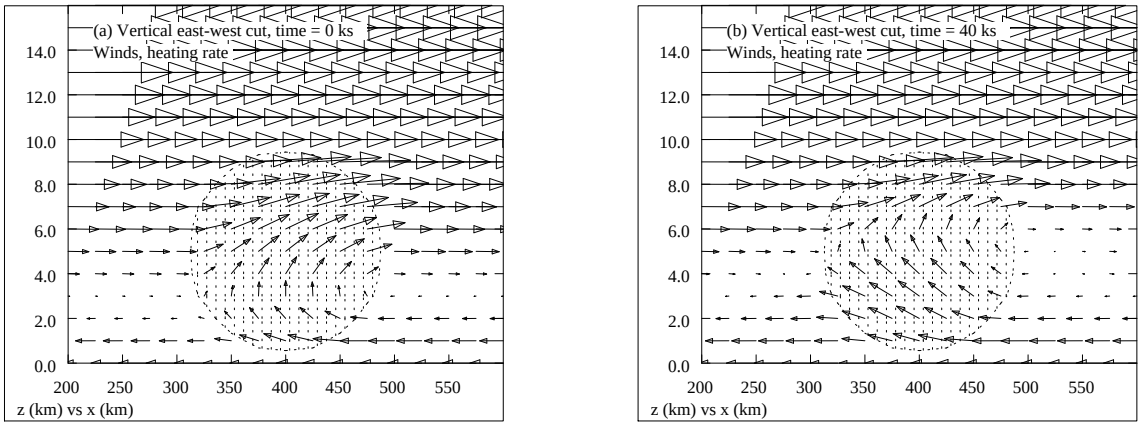

Figure 12: Vertical, east-west section through the simulation shown in figure 11. The scale for the velocity vectors is 5 m s^{-1} per 25 km, and the region where the heating exceeds 0.1 K ks^{-1} is hatched. (a) Time = 0. (b) Time = 40 ks.

in the environment, has $x_0 = y_0 = 100$ km, $z_0 = 5$ km, $z_l = 5$ km, and a maximum heating rate of 0.5 K ks^{-1}. Since the Coriolis force is turned off in this simulation, potential vorticity transport is completely horizontal, at least initially. The result is as expected, with a positive potential vorticity anomaly building up on the south side of the heated region, and a negative anomaly to the north. The ambient flow at 5 km is 4 m s^{-1} to the right, which explains why the potential vorticity anomalies are extended in that direction.

The upper limit on the imposed heating rate according to the scale analysis is $\epsilon\Gamma ZV/L$. Taking $\Gamma =$ 3 K km^{-1}, $Z = z_0$, $L = x_0$, and V equal to the vertical shear times Z, this upper limit is $\epsilon \times 0.75$ K ks^{-1}. The heating rate of 0.5 K ks^{-1} used in this example thus stretches that allowed by the scale analysis, and the results can only be considered qualitatively correct.

The most interesting aspect of figure 11 is that the flow along the central east-west axis is actually reversed, with a maximum flow toward the west of about 2 m s^{-1}. This flow reversal results from the circulations around the two induced potential vorticity anomalies.

The structure of this feature is clarified in vertical, east-west sections through the center of the heated

region. Figures 12a and 12b show the flow in this plane at the instant that the heating is turned on, and 40 ks later. At the initial time the heating-induced updraft's horizontal wind component is essentially the same as in the undisturbed environment, leading to a down-shear tilt to the updraft above 3 km. At the later time, the induced flow from the potential vorticity anomalies causes the tilt of the updraft to be up-shear below 6 km.

Strong front-to-rear and rear-to-front jets are often seen in squall lines and mesoscale convective systems, e. g., Ogura and Liou (1980), Smull and Houze (1987). The former are generally associated with mesoscale upward motion and the latter with evaporatively induced sinking. These two jets could be caused respectively by the nonadvective transport of potential vorticity associated with latent heat release in the updraft and evaporative cooling and melting in the underlying downdraft. The environmental shear associated with squall lines almost always has a strong rear-to-front component (Bluestein and Jain, 1985; Wyss and Emanuel, 1988), which would result in the development of jets of the proper sense.

c. Potential vorticity flux and molecular and turbulent processes

Mountain lee waves being absorbed at a critical level would generate an applied force similar to that simulated in the first subsection, though of course the ambient wind pattern would necessarily be sheared in that case. It is interesting to imagine how the momentum carried by the gravity waves is transferred to the jet. Since the jet and the potential vorticity doublet are inexorably coupled, the deposition of momentum by the gravity waves must create the potential vorticity anomalies as well. However, since potential vorticity is conserved in adiabatic, inviscid flows, including those involving gravity waves, and since there is initially no potential vorticity, the creation of the potential vorticity anomalies *must* be associated with the final dissipation of the gravity wave motions by viscosity as they become unstable and cascade to small scales at the critical level. Thus, the viscous force is a necessary intermediary in the process of adjustment whereby the gravity waves transfer their momentum to the mean flow.

An alternate way of looking at the above problem is to start the analysis from the Reynolds-averaged primitive equations rather than the equations in their raw form. In this case the applied force no longer refers directly to viscous forces, since these are truly negligible for the averaged flow. Instead, the force is minus the divergence of the turbulent momentum flux, $\mathbf{F} = -\rho_0^{-1}\nabla\cdot(\rho_0\mathbf{v}'\mathbf{v}')$. The advantage of this approach is that models are available for turbulent momentum transport, thus opening the way to the computation of potential vorticity fluxes in real situations.

There would seem to be a contradiction between these two approaches, in that the potential vorticity flux is driven by viscous forces in one case and by turbulent mixing in the other. However, as Keyser and Rotunno (1990) point out, the potential vorticity defined in terms of Reynolds-averaged quantities is a different quantity than the potential vorticity defined in terms of the raw velocity and potential temperature fields. This difference is important precisely when turbulence exists. Thus, it is not surprising that the two quantities have different source terms. It is also worth remembering that the existence of a turbulent cascade implies viscous dissipation at small scales. Thus, the two source terms, though not necessarily identical, must be intimately related.

The nature of this relationship may be made clear in another way. Imagine a situation in which an episode of turbulence is confined to a finite time interval. While the turbulence is active, the two above-defined potential vorticities will differ. However, before and after the turbulent event the two should be substantially the same, since the difference between raw and Reynolds-averaged fields would be negligible in the absence of turbulence. Thus, the net, time-integrated effect of the two source terms considered together with the different advective transports in the two cases must be identical.

All of the considerations expressed above for molecular and turbulent fluxes of momentum also hold for fluxes of heat. The main lesson is that potential vorticity transport has no relationship to the potential vorticity gradient in either molecular or turbulent processes. Instead, it is given by (4), with $\mathbf{F}$ and H

appropriately defined so as to either exclude or include the divergence of turbulent fluxes of momentum and heat, depending on whether potential vorticity is defined in terms of raw or Reynolds-averaged fields. Thus, K-theory models of the turbulent transport of potential vorticity, as well as conventional higher order closure schemes, are inappropriate when applied to the potential vorticity field. Misunderstanding of this point has recently generated a certain amount of controversy, e. g., Danielsen (1990) and Haynes and McIntyre (1990).

6 Discussion

In this paper two versions of the nonlinear balance equations are consistently derived from the divergence, potential vorticity, and potential temperature equations. These are called the quasibalanced and semibalanced equations in analogy with the quasigeostrophic and semigeostrophic equations of low Rossby number dynamics. The quasibalanced equations are valid for small Froude number uniformly in Rossby number, while the semibalanced equations simply require that the vertical velocity be uniformly small compared to naive continuity equation scaling.

Truncated versions of the potential vorticity definition are inverted for the streamfunction in both cases, and the nonlinear balance equation is solved for the pressure (actually, Exner function) perturbation. The potential temperature perturbation is obtained using the hydrostatic equation, and the vertical velocity and unbalanced flow are obtained diagnostically from the potential temperature equation. The model is time stepped by advecting the potential vorticity and surface potential temperature.

The rest of the paper is devoted to presenting a variety of idealised semibalanced simulations in three dimensions. The purpose of these simulations is to provide a basis for "potential vorticity thinking" on mesoscale flows, i. e., those flows with Rossby number of order unity or greater. In order to understand large Rossby number flows in their purest form, most simulations are done at infinite Rossby number, i. e., with the Coriolis force suppressed. Flows with order one Rossby number should in principle meld

the effects of low Rossby number flows, which are already well understood, with the effects demonstrated here.

The main differences between the behaviour of flows at high and low Rossby numbers are as follows:

1. At infinite Rossby number, an isolated, unsheared potential vorticity anomaly of either sign produces a warm region in the upper part of the anomaly and a cool region in the lower part. For a weak anomaly (i. e., $q_{max} < Q$), the potential temperature perturbations scale as the square of the strength of the potential vorticity anomaly, and are weak enough that the potential vorticity becomes essentially the vertical component of absolute vorticity. Wind speed scales linearly with anomaly strength. This is the regime of validity of quasibalanced theory. When $q_{max} > Q$, quasibalanced theory gives way to semibalanced theory, and the relationships between potential temperature perturbations, wind speeds, and the strength of the potential vorticity anomaly become more complicated.

2. At infinite Rossby number the vertical radius of influence of an isolated potential vorticity anomaly is zero in isentropic coordinates. In pressure coordinates the vertical influence is therefore limited to the distortion of isentropic surfaces.

3. The tangential velocity around an isolated potential vorticity anomaly varies as the inverse of the horizontal radius at infinite Rossby number. Thus the specific angular momentum associated with rotation around the vortex is asymptotically constant as a function of radius. Laterally separated potential vorticity anomalies therefore interact much as do barotropic vortices in two dimensions.

4. Vertical motion is generated when a potential vorticity anomaly is embedded in shear. It arises from advection of parcels up and down deformed isentropic surfaces, and from

time evolution of these surfaces caused by deformation of potential vorticity fields. The only difference between the high and low Rossby number situation is that the ambient tilt of isentropic surfaces associated with shear doesn't exist in the former case.

5. Diabatic heating and frictional effects are manifested in the creation of a nonadvective potential vorticity flux. This flux has no relationship to the gradient of potential vorticity, rendering conventional K-theory and similar techniques useless in directly determining the transport of potential vorticity. Instead, (4) must be used, with appropriate source terms for diabatic heating and viscous forces inserted. If the primitive equations from which the analysis is begun are Reynolds-averaged, then a potential vorticity constructed from the Reynolds-averaged potential temperature and absolute vorticity must be used. The source terms must then include the divergences of the turbulent fluxes of heat and momentum. This procedure is generally more useful than working with the unaveraged primitive equations.

6. Externally applied forces result in transport of potential vorticity in a direction normal to both the force and the potential temperature gradient. This is generally close to the horizontal, but, for instance, could be up or down tilted frontal surfaces. The transport due to diabatic heating is antiparallel to the product of the heating rate and the absolute vorticity vector.

Considerable work needs to be done to compare the results of the semibalanced approximation to equivalent primitive equation solutions at large Rossby number. Currently, the only useful solutions known to the author are the model results of Smolarkiewicz and Rotunno (1989; 1990) for flow around a three dimensional mountain. Comparison with these results shows that significant errors exist in both the semibalanced model results for this problem and those of Drazin (1961). However, the approximation

criterion used for deriving the model equations is strongly violated in this case, so the lack of agreement is understandable.

It would be particularly interesting to compare semibalanced model results for a vortex in shear and heating in shear to the output of a primitive equation model, as both of these cases are pertinent to the dynamics of mesoscale convective systems.

One point remains to be resolved. The Lorenz theory and the semibalanced model are shown here to be very similar approximations. Yet the semibalanced theory results from a consistent leading order truncation of the primitive equations, while the Lorenz model truncates the divergence equation to leading order and the vorticity equation to first order. The resolution of this difference cannot be traced to the use of potential vorticity in one model and vorticity in the other. Instead, it lies in the additional assumption made in the derivation of the semibalanced model that the balanced advections of potential vorticity and potential temperature are nearly cancelled by the corresponding time tendencies. This assumption is unfortunate, in that it can only be verified *a posteriori* except in the simplest cases. However, its application to the derivation of the Lorenz equations shows that this model can also be obtained as a consistent leading order truncation of the primitive equations. In particular, the balanced advection and time tendency of the vertical component of vorticity should nearly cancel if this condition also holds for the potential vorticity and potential temperature, making the entire vertical component of the vorticity equation first order or higher. The virtue of this additional assumption is that it extends the range of validity of nonlinear balance models to at least some situations in which potential temperature and potential vorticity perturbations are "large", i. e., leading rather than first order at all Rossby number.

Acknowledgements. The author greatly benefited from discussions with Michael McIntyre and Richard Rotunno on this subject. Peter Gent and Ronald Smith also made helpful comments. Steve Schaffer introduced me to the mysteries of the multigrid method. The anonymous reviewers helped make this a

significantly better paper. This work was supported by the Atmospheric Sciences Section of the National Science Foundation under Grant No. ATM-8914116.

REFERENCES

Allen, J. S., 1991: Balance equations based on momentum equations with global invariants of potential enstrophy and energy. *J. Phys. Ocean.,* **21**, 265-276.

Bluestein, H. B. and M. H. Jain, 1985: Formation of mesoscale lines of precipitation: Severe squall lines in Oklahoma during the spring. *J. Atmos. Sci.,* **42,** 1711-1732.

Bolin, B., 1955: Numerical forecasting with the barotropic model. *Tellus,* **7**, 27-49.

Bolin, B., 1956: An improved barotropic model and some aspects of using the balance equations for three-dimensional flow. *Tellus,* **8**, 61-75.

Briggs, W. L., 1987: A multigrid tutorial. SIAM, Philadelphia, 88 pp.

Charney, J. G., 1955: The use of primitive equations of motion in numerical prediction. *Tellus,* **7**, 22-26.

Charney, J. G., 1962: Integration of the primitive and balance equations. Proc. Symp. on Num. Weather Prediction, Tokyo, 131-150.

Charney, J. G., and M. E. Stern, 1962: On the stability of internal baroclinic jets in a rotating atmosphere. *J. Atmos. Sci.,* **19**, 159-172.

Craig, G. C., 1991: A three-dimensional generalization of Eliassen's balanced vortex equations derived from Hamilton's principle. *Quart. J. Roy. Met. Soc.,* **117**, 435-448.

Daly, B. J., 1967: Numerical study of two fluid Rayleigh-Taylor instability. *Phys. Fluids,* **10,** 297-307.

Danielsen, E. F., 1990: In defense of Ertel's potential vorticity and its general applicability as a meteorological tracer. *J. Atmos. Sci.,* **47**, 2013-2020.

Davis, C. A., and K. A. Emanuel, 1991: Potential vorticity diagnosis of cyclogenesis. *Mon. Wea. Rev.,* **119**, 1287-1300.

Drazin, P. G., 1961: On the steady flow of a fluid of variable density past an obstacle. *Tellus,* **13,** 239-251.

Eliassen, A., 1952: Slow thermally or frictionally controlled meridional circulation in a circular vortex. *Astrophys. Norvegica,* **5, no. 2**, 19-60.

Gent, P. R., and J. C. McWilliams, 1983: Consistent balanced models in bounded and periodic domains. *Dyn. Atmos. Oceans,* **7,** 67-93.

Gent, P. R., and J. C. McWilliams, 1984: Balanced models in isentropic coordinates and the shallow water equations. *Tellus,* **36A**, 166-171.

Harlow, F. H., and J. E. Welch, 1965: Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *Phys. Fluids,* **8,** 2182-2189.

Haynes, P. H., and M. E. McIntyre, 1987: On the evolution of vorticity and potential vorticity in the presence of diabatic heating and frictional or other forces. *J. Atmos. Sci.,* **44,** 828-841.

Haynes, P. H., and M. E. McIntyre, 1990: On the conservation and impermeability theorems for potential vorticity. *J. Atmos. Sci.,* **47,** 2021-2031.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.,* **111,** 877-946.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1987: Reply (to a comment by J. S. A. Green). *Quart. J. Roy. Meteor. Soc.,* **113,** 402-404.

Keyser, D., and R. Rotunno, 1990: On the formation of potential-vorticity anomalies in upper-level jet-front systems. *Mon. Wea. Rev.,* **118,** 1914-1921.

Lilly, D. K., 1983: Stratified turbulence and the mesoscale variability of the atmosphere. *J. Atmos. Sci.,* **40,** 749-761.

Lorenz, E. N., 1960: Energy and numerical weather prediction. *Tellus,* **12,** 364-373.

McIntyre, M. E., and W. A. Norton, 1991: Potential-vorticity inversion on a hemisphere. *J. Atmos. Sci., (In press).*

McWilliams, J. C., 1985: A uniformly valid model spanning the regimes of geostrophic and isotropic, stratified turbulence: Balanced turbulence. *J. Atmos. Sci.,* **42,** 1773-1774.

McWilliams, J. C., and P. R. Gent, 1986: The evolution of sub-mesoscale coherent vortices on the β-plane. *Geophys. Astrophys. Fluid Dyn.,* **35**, 235-255.

McWilliams, J. C., P. R. Gent, and N. J. Norton, 1986: The evolution of balanced, low-mode vortices on the β-plane. *J. Phys. Ocean.,* **16**, 838-855.

Norton, N. J., J. C. McWilliams, and P. R. Gent, 1986: A numerical model of the balance equations in a periodic domain and an example of balanced turbulence. *J. Comput. Phys.,* **67**, 439-471.

Ogura, Y. and M. T. Liou, 1980: The structure of a mid-latitude squall-line: A case study. *J. Atmos. Sci.,* **37,** 553-567.

Raymond, D. J., and H. Jiang, 1990: A theory for long-lived mesoscale convective systems. *J. Atmos. Sci.,* **47**, 3067-3077.

Rosenthal, S., 1971: A circularly symmetric primitive-equation model of tropical cyclones and its response to artificial enhancement of the convective heating functions. *Mon. Wea. Rev.,* **99,** 414-426.

Schlesinger, R. E., 1978: A three-dimensional numerical model of an isolated thunderstorm: Part I. Comparative experiments for variable ambient wind shear. *J. Atmos. Sci.,* **35,** 690-713.

Schubert, W. H., and J. J. Hack, 1982: Inertial stability and tropical cyclone development. *J. Atmos. Sci.,* **39,** 1687-1697.

Shapiro, L. J., and H. E. Willoughby, 1982: The response of balanced hurricanes to local sources of heat and momentum. *J. Atmos. Sci.,* **39,** 378-394.

Smith, R. K., 1981: The cyclostrophic adjustment of vortices with application to tropical cyclone modification. *J. Atmos. Sci.,* **38,** 2021-2030.

Smolarkiewicz, P. K., and R. Rotunno, 1989: Low Froude number flow past three dimensional obstacles. Part I: Baroclinically generated lee vortices. *J. Atmos. Sci.,* **46,** 1154-1164.

Smolarkiewicz, P. K., and R. Rotunno, 1990: Low Froude number flow past three-dimensional obstacles. Part II: Upwind flow reversal zone. *J. Atmos. Sci.,* **47,** 1498-1511.

Smull, B. F., and R. A. Houze, 1987: Rear inflow in squall lines with trailing stratiform precipitation. *Mon. Wea. Rev.,* **115,** 2869-2889.

Sundqvist, H, 1970: Numerical simulation of the development of tropical cyclones with a ten-level model. *Tellus,* **22,** 359-390.

Thorpe, A. J., 1985: Diagnosis of balanced vortex structure using potential vorticity. *J. Atmos. Sci.,* **42,** 397-406.

Willoughby, H. E., 1979: Forced secondary circulations in hurricanes. *J. Geophys. Res.,* **84,** 3173-3183.

Wyss, J., and K. A. Emanuel, 1988: The pre-storm environment of midlatitude squall lines. *Mon. Wea. Rev.,* **116,** 790-794.