

Boundary Effects in Potential Vorticity Dynamics

TAPIO SCHNEIDER*

Courant Institute of Mathematical Sciences and Center for Atmosphere–Ocean Science, New York University, New York, New York

ISAAC M. HELD AND STEPHEN T. GARNER

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

(Manuscript received 19 December 2001, in final form 30 August 2002)

ABSTRACT

Many aspects of geophysical flows can be described compactly in terms of potential vorticity dynamics. Since potential temperature can fluctuate at boundaries, however, the boundary conditions for potential vorticity dynamics are inhomogeneous, which complicates considerations of potential vorticity dynamics when boundary effects are dynamically significant.

A formulation of potential vorticity dynamics is presented that encompasses boundary effects. It is shown that, for arbitrary flows, the generalization of the potential vorticity concept to a sum of the conventional interior potential vorticity and a singular surface potential vorticity allows one to replace the inhomogeneous boundary conditions for potential vorticity dynamics by simpler homogeneous boundary conditions (of constant potential temperature). Functional forms of the surface potential vorticity are derived from field equations in which the potential vorticity and a potential vorticity flux appear as sources of flow quantities in the same way in which an electric charge and an electric current appear as sources of fields in electrodynamics. For the generalized potential vorticity of flows that need be neither balanced nor hydrostatic and that can be influenced by diabatic processes and friction, a conservation law holds that is similar to the conservation law for the conventional interior potential vorticity. The conservation law for generalized potential vorticity contains, in the quasigeostrophic limit, the well-known dual relationship between fluctuations of potential temperature at boundaries and fluctuations of potential vorticity in the interior of quasigeostrophic flows. A nongeostrophic effect described by the conservation law is the induction of generalized potential vorticity by baroclinicity at boundaries, an effect that plays a role, for example, in mesoscale flows past topographic obstacles. Based on the generalized potential vorticity concept, a theory is outlined of how a wake with lee vortices can form in weakly dissipative flows past a mountain. Theoretical considerations and an analysis of a simulation show that a wake with lee vortices can form by separation of a generalized potential vorticity sheet from the mountain surface, similar to the separation of a friction-induced vorticity sheet from an obstacle, except that the generalized potential vorticity sheet can be induced by baroclinicity at the surface.

1. Introduction

Since potential vorticity is materially conserved in adiabatic and frictionless flows, and since it contains all relevant information about balanced flows in a single scalar field, many aspects of geophysical flows can be described compactly in terms of potential vorticity dynamics. For example, the propagation of Rossby waves and the development of baroclinic instability have traditionally been described in terms of quasigeostrophic

potential vorticity dynamics. And in situations in which the quasigeostrophic approximation is inadequate, such as for planetary-scale flows, considerations of potential vorticity dynamics on isentropes (or on isopycnals in the ocean) have proven fruitful (see, e.g., Tung 1986; Rhines and Young 1982). Since potential temperature can fluctuate at boundaries, however, the boundary conditions for potential vorticity dynamics are inhomogeneous, which complicates considerations of potential vorticity dynamics when boundary effects are dynamically significant. Within quasigeostrophic theory, Bretherton (1966) has shown that the inhomogeneous boundary condition implied by a fluctuating potential temperature at a boundary can be replaced by a homogeneous boundary condition of constant potential temperature if a singular surface potential vorticity proportional to the surface potential temperature fluctuations is included in the quasigeostrophic potential vorticity. Extending Bretherton's argumentation, Rhines

* Current affiliation: Division of Geological and Planetary Sciences and Division of Engineering and Applied Sciences, California Institute of Technology, Pasadena, California.

Corresponding author address: Tapio Schneider, California Institute of Technology, Mail Code 100-23, 1200 E. California Blvd., Pasadena, CA 91125.
E-mail: tapio@gps.caltech.edu

(1979) has shown that not only surface potential temperature fluctuations, but also the topography of a boundary can be taken into account in a quasigeostrophic surface potential vorticity. Bretherton's and Rhines's generalization of the quasigeostrophic potential vorticity concept has been used to describe the interaction between quasigeostrophic potential vorticity fluctuations in the interior of a flow on the one hand and surface potential temperature fluctuations and/or topographic slopes on the other hand, for example, in unstable baroclinic waves [see Hoskins et al. (1985) and Hallberg and Rhines (2000) for reviews]. Here we present a similar generalization of the potential vorticity concept that allows for the inclusion of boundary effects in the potential vorticity dynamics of arbitrary nongeostrophic flows.

In arbitrary flows as in quasigeostrophic flows, the inhomogeneous boundary conditions for potential vorticity dynamics can be replaced by homogeneous boundary conditions if the simplification of the boundary conditions is compensated by a generalization of the potential vorticity concept to a sum of the conventional interior potential vorticity and a singular surface potential vorticity. In addition to the contributions from surface potential temperature fluctuations that are contained in the quasigeostrophic surface potential vorticity, the surface potential vorticity generally also contains contributions from surface vorticity fluctuations, which are neglected in the quasigeostrophic approximation. We derive functional forms of the generalized potential vorticity and of its conservation law and discuss nongeostrophic effects described by the generalized potential vorticity conservation law. As an illustration of how the generalized potential vorticity concept can be used to describe flows for which the quasigeostrophic approximation is inadequate, we demonstrate that this concept can form a basis of theories of lee vortex formation in mesoscale flows past a mountain.

Sections 2 through 4 set up the formal framework of generalized potential vorticity dynamics. Section 2 casts the momentum equation with the help of the thermodynamic equation in the form of field equations in which the potential vorticity and the potential vorticity flux appear as sources of flow quantities in the same way in which an electric charge and an electric current appear as sources of fields in electrodynamics. Section 3 derives, by means of techniques from electrodynamics, the functional forms of the generalized potential vorticity and of the generalized potential vorticity flux that replace the conventional interior potential vorticity and potential vorticity flux when the inhomogeneous boundary conditions for potential vorticity dynamics are replaced by homogeneous boundary conditions. For the generalized potential vorticity, a conservation law holds that reduces, in the quasigeostrophic limit, to the conservation law for Bretherton's (1966) generalized quasigeostrophic potential vorticity. In section 4, the conservation law for generalized potential vorticity, derived

in coordinate-independent form in section 3, is expanded in isentropic coordinates.

Section 5 discusses the baroclinic induction of generalized potential vorticity at boundaries, a nongeostrophic effect. An analysis of a simulated Boussinesq flow demonstrates that the formation of a wake with lee vortices in flows past a mountain with a free-slip surface can be described in terms of generalized potential vorticity dynamics and the baroclinic induction of generalized potential vorticity at the mountain surface. A wake with lee vortices can form by separation of a generalized potential vorticity sheet from the mountain surface, similar to the separation of a friction-induced vorticity sheet from an obstacle, except that the generalized potential vorticity sheet can be induced by baroclinicity at the surface.

Section 6 summarizes the conclusions. The appendix lists the notation and symbols used in this paper.

The analyses presume an ideal-gas atmosphere with the planet's surface as the only dynamically relevant boundary. However, the concepts and mathematical techniques presented are easily adaptable, for example, to ocean flows with lateral boundaries and with a more complex thermal equation of state, irrespective of the fact that the more complex thermal equation of state of seawater implies that potential vorticity is not necessarily materially conserved in adiabatic and frictionless ocean flows (cf. McDougall 1988).

2. Potential vorticity and potential vorticity flux as sources of flow quantities

a. Field equations

The potential vorticity is the pseudoscalar function

$$P = \frac{\boldsymbol{\omega}_a \cdot \nabla \theta}{\rho} \quad (1)$$

of absolute vorticity $\boldsymbol{\omega}_a$, potential temperature θ , and density ρ . The absolute vorticity is the curl $\boldsymbol{\omega}_a = \nabla \times \mathbf{u}_a$ of the three-dimensional absolute velocity $\mathbf{u}_a = \mathbf{u} + \boldsymbol{\Omega} \times \mathbf{r}$, or the sum $\boldsymbol{\omega}_a = \boldsymbol{\omega}_r + 2\boldsymbol{\Omega}$ of the relative vorticity $\boldsymbol{\omega}_r = \nabla \times \mathbf{u}$ and the vorticity $\nabla \times (\boldsymbol{\Omega} \times \mathbf{r}) = 2\boldsymbol{\Omega}$ of a planetary rotation with constant angular velocity $\boldsymbol{\Omega}$. The potential vorticity P is a conserved quantity with a conservation law of the flux form

$$\partial_t(\rho P) + \nabla \cdot (\rho \mathbf{J}) = 0, \quad (2)$$

with a potential vorticity flux¹

$$\mathbf{J} = \mathbf{u}P - \rho^{-1}Q\boldsymbol{\omega}_a + \rho^{-1}\nabla\theta \times \mathbf{F} \quad (3)$$

in which $Q = D\theta/Dt$ is a diabatic heating rate and \mathbf{F} a

¹ This potential vorticity flux \mathbf{J} differs from the quantity $\rho\mathbf{J}$ that Haynes and McIntyre (1987) call potential vorticity flux by a density factor ρ . We refer to the quantity $\rho\mathbf{J}$ as the potential vorticity flux density. For a uniqueness property of the potential vorticity flux (3), see Bretherton and Schär (1993).

frictional force per unit mass (Haynes and McIntyre 1987). Diabatic heating Q and frictional forces \mathbf{F} contribute to the potential vorticity flux \mathbf{J} and redistribute potential vorticity within a flow, but in the interior of the flow, they do not create or destroy potential vorticity (Haynes and McIntyre 1987, 1990).

Since the divergence of the absolute vorticity vanishes, $\nabla \cdot \boldsymbol{\omega}_a = 0$, the product $\rho P = \boldsymbol{\omega}_a \cdot \nabla \theta$ of density and potential vorticity is the divergence $\rho P = \nabla \cdot \mathbf{D}$ of a vector field \mathbf{D} . The density and potential vorticity determine the vector field \mathbf{D} up to a nondivergent component. For example, one might choose $\mathbf{D} = \theta \boldsymbol{\omega}_a$, or $\mathbf{D} = \mathbf{u}_a \times \nabla \theta$. The difference between these two choices for \mathbf{D} is the nondivergent vector field $\nabla \times (\theta \mathbf{u}_a)$. For both choices for \mathbf{D} , the product ρP is the divergence, and hence the source, of \mathbf{D} .

That the product ρP is both a conserved quantity and the divergence of a vector field can be expressed through field equations that make some properties of potential vorticity manifest and that will be convenient in analyzing the role of boundaries in potential vorticity dynamics. Upon substitution of the divergence $\rho P = \nabla \cdot \mathbf{D}$, the conservation law (2) for potential vorticity becomes $\nabla \cdot (\partial_t \mathbf{D} + \rho \mathbf{J}) = 0$ (cf. Bretherton and Schär 1993). It follows that the potential vorticity flux density $\rho \mathbf{J}$ has the form

$$\rho \mathbf{J} = -\partial_t \mathbf{D} + \nabla \times \mathbf{H},$$

where \mathbf{H} is the vector potential of the sum $\partial_t \mathbf{D} + \rho \mathbf{J}$. The facts that ρP is both a conserved quantity and the divergence of a vector field \mathbf{D} can thus be expressed through the first two Maxwell equations:

$$\nabla \cdot \mathbf{D} = \rho P, \quad (4a)$$

$$\nabla \times \mathbf{H} - \partial_t \mathbf{D} = \rho \mathbf{J}. \quad (4b)$$

The quantity ρP corresponds to the charge density in electrodynamics, the field \mathbf{D} to the electric displacement field, the field \mathbf{H} to the magnetic field, and the potential vorticity flux density $\rho \mathbf{J}$ to the current density. We refer to the quantity ρP as the potential vorticity density.² In the Maxwell equations (4), the potential vorticity P and the potential vorticity flux \mathbf{J} appear as sources of the fields \mathbf{D} and \mathbf{H} , just as in electrodynamics charges and currents appear as sources of the electric displacement field and of the magnetic field. The conservation law (2) for potential vorticity follows from the Maxwell equations by adding the time derivative of Eq. (4a) to the divergence of Eq. (4b).

b. Gauge invariance

The potential vorticity P , the potential vorticity flux \mathbf{J} , and the density ρ do not determine the fields \mathbf{D} and

² We prefer the term ‘‘potential vorticity density’’ to Haynes and McIntyre’s (1990) term ‘‘amount per unit volume of potential vorticity substance.’’ The word substance connotes an independence of other quantities and a subsistence in itself that are not characteristic of the potential vorticity.

\mathbf{H} uniquely. The Maxwell equations (4) are invariant under gauge transformations of the form

$$\mathbf{D} \leftarrow \mathbf{D} + \nabla \times \mathbf{A} \quad (5)$$

$$\mathbf{H} \leftarrow \mathbf{H} + \partial_t \mathbf{A} + \nabla \psi,$$

where \mathbf{A} is a vector field and ψ a scalar field. Given a potential vorticity P , a potential vorticity flux \mathbf{J} , and a density ρ , the fields \mathbf{D} and \mathbf{H} are only determined up to such gauge transformations.

The freedom in the definition of the fields \mathbf{D} and \mathbf{H} can be exploited to find gauges of the fields \mathbf{D} and \mathbf{H} that are amenable to physical interpretation or are convenient in specific contexts.

c. Gauge I

With the choice

$$\mathbf{D} = \mathbf{u}_a \times \nabla \theta, \quad (6)$$

the potential vorticity density $\nabla \cdot \mathbf{D} = \rho P$ is the divergence of a field that indicates, up to a minus sign, the absolute angular momentum of the flow along isentropes.³

A field \mathbf{H} for this choice for \mathbf{D} can be found from the Maxwell equation (4b) by expanding the time derivative $\partial_t \mathbf{D} = \partial_t \mathbf{u}_a \times \nabla \theta + \mathbf{u}_a \times \nabla (\partial_t \theta)$ and by substituting for the velocity derivative $\partial_t \mathbf{u}_a = \partial_t \mathbf{u}$ from the momentum equation

$$\partial_t \mathbf{u} + \boldsymbol{\omega}_a \times \mathbf{u} = -\frac{1}{\rho} \nabla p - \frac{1}{2} \nabla \|\mathbf{u}\|^2 - \nabla \Phi + \mathbf{F}$$

and for the potential temperature derivative $\partial_t \theta$ from the thermodynamic equation

$$\partial_t \theta + \mathbf{u} \cdot \nabla \theta = Q. \quad (7)$$

Using the differential $d\hat{h} = \rho^{-1} dp + T d\hat{s}$ of the specific enthalpy $\hat{h} = c_p T$ and the relation $\nabla \hat{s} = c_p \nabla \log \theta$ between gradients of specific entropy \hat{s} and gradients of potential temperature θ , one can cast the momentum equation in the form (cf. Batchelor 1967, chapter 3.5; Schär 1993)

$$\partial_t \mathbf{u} + \boldsymbol{\omega}_a \times \mathbf{u} = c_p T \nabla \log \theta - \nabla B + \mathbf{F}, \quad (8)$$

with Bernoulli function

$$B = \frac{1}{2} \|\mathbf{u}\|^2 + c_p T + \Phi. \quad (9)$$

Taking the cross product of the momentum equation (8)

³ More precisely, if one expands the field $\mathbf{D} = \mathbf{u}_a \times \nabla \theta$ in isentropic coordinates, using the hydrostatic approximation and the notation and techniques that follow in section 4, one obtains $\mathbf{D} = h^{-1}(v_a, -u_a, 0)^a$. Within the approximations of the primitive equations (in which the distance from any point in the atmosphere to the center of the planet is taken to be equal to the planet radius a), this field \mathbf{D} is related to the absolute angular momentum density $\rho \mathbf{L}_a = \rho a(-v_a, u_a, 0)^a$ of the flow along isentropes by $\rho \mathbf{D} = -a^{-1} \rho \mathbf{L}_a$. The divergence of the field \mathbf{D} becomes in isentropic coordinates $\nabla \cdot \mathbf{D} = h^{-1}(\partial_x v_a - \partial_y u_a)$.

and the potential temperature gradient $\nabla\theta$ leads to the Maxwell equation (4b) with potential vorticity flux (3) and field

$$\mathbf{H} = -B\nabla\theta - (\partial_t\theta)\mathbf{u}_a. \quad (10)$$

The explicit time derivative $\partial_t\theta$ in this expression for the field \mathbf{H} could be expanded by substitution from the thermodynamic equation (7); however, with the explicit time derivative $\partial_t\theta$, it is evident that the second term in the expression for the field \mathbf{H} does not appear in a representation of the field in isentropic coordinates. In isentropic coordinates, the field \mathbf{H} is proportional to the Bernoulli function times the cross-isentropic basis vector $\nabla\theta$, and the curl $\nabla \times \mathbf{H}$ involves only derivatives of the Bernoulli function along isentropes.⁴

In a steady state, the Maxwell equation (4b) reduces to $\rho\mathbf{J} = \nabla \times \mathbf{H}$, which becomes, in this gauge, $\rho\mathbf{J} = -\nabla \times B\nabla\theta = \nabla\theta \times \nabla B$. That is, as noted by Schär (1993), the potential vorticity flux is directed along lines of intersection between surfaces of constant potential temperature θ (isentropes) and surfaces of constant Bernoulli function B . The Bernoulli function B is the streamfunction of the potential vorticity flux \mathbf{J} along isentropes. Since the Bernoulli function is a measure of the specific energy of the moving fluid, along-stream gradients of the Bernoulli function imply energy dissipation. According to the relation $\rho\mathbf{J} = \nabla\theta \times \nabla B$ between potential vorticity fluxes and Bernoulli function, energy dissipation indicated by gradients of the Bernoulli function along streamlines on isentropes are associated, in a steady state, with across-stream potential vorticity fluxes. Some practical implications of this generalization of Bernoulli's theorem are discussed by Schär (1993), Schär and Durran (1997), and, in an oceanic context, by Marshall et al. (2001).

The Maxwell equations (4) in the gauge given by Eqs. (6) and (10) generalize to time-dependent flows the relation between potential vorticity fluxes and Bernoulli function that Schär (1993) noted for steady flows. In this gauge, the potential vorticity P and the potential vorticity flux \mathbf{J} appear as sources of orthogonal fields \mathbf{D} and \mathbf{H} that indicate, up to minus signs, the absolute angular momentum and the energy of the flow along isentropes. With the potential vorticity flux \mathbf{J} understood as source (or, because of the minus sign, as sink) of a field that indicates the energy of the flow along isentropes, it is not as surprising as it might seem that a nonconservative force \mathbf{F} and the diabatic heating rate Q contribute to a quantity \mathbf{J} that is called a flux.

⁴ More precisely, if one expands the field $\mathbf{H} = -B\nabla\theta - (\partial_t\theta)\mathbf{u}_a$ in isentropic coordinates, using the hydrostatic approximation and the notation and techniques that follow in section 4, one obtains $\mathbf{H} = -h^{-1}B\mathbf{e}^\theta$, where $\mathbf{e}^\theta = (-\partial_x z, -\partial_y z, h^{-1})^\theta$ is the contravariant basis vector $\nabla\theta$ nondimensionalized by multiplication by the scale factor h , and derivatives are to be understood as derivatives along isentropes (cf. Arfken 1985, chapter 3.8). So the field \mathbf{H} is related to the Bernoulli density ρB by $\rho\mathbf{H} = -\rho B\mathbf{e}^\theta$. The curl of the field \mathbf{H} becomes in isentropic coordinates $\nabla \times \mathbf{H} = -h^{-1}(\partial_x B, -\partial_y B, 0)^\theta$.

d. Gauge II

For the analysis of boundary conditions in section 3, the choice

$$\mathbf{D} = \theta\boldsymbol{\omega}_a \quad (11)$$

is convenient. A field \mathbf{H} for this choice for \mathbf{D} can be found, as above, from the Maxwell equation (4b) by expanding the time derivative $\partial_t\mathbf{D} = (\partial_t\theta)\boldsymbol{\omega}_a + \theta(\partial_t\boldsymbol{\omega}_a)$ and by substituting for the potential temperature derivative $\partial_t\theta$ from the thermodynamic equation (7) and for the vorticity derivative $\partial_t\boldsymbol{\omega}_a = \partial_t\boldsymbol{\omega}_r$ from the vorticity equation

$$\partial_t\boldsymbol{\omega}_r = \nabla \times (\mathbf{u} \times \boldsymbol{\omega}_a + c_p T \nabla \log \theta + \mathbf{F}) \quad (12)$$

belonging to the momentum equation (8). Similar algebra to the above leads to the potential vorticity flux (3) and to the field⁵

$$\mathbf{H} = \mathbf{u} \times \theta\boldsymbol{\omega}_a + c_p T \nabla \theta + \theta\mathbf{F}. \quad (13)$$

This functional form of the field \mathbf{H} does not seem to have a simple physical interpretation.

Irrespective of the gauge of the fields \mathbf{D} and \mathbf{H} , the Maxwell equations (4) are a way of arranging the momentum equation with the help of the thermodynamic equation such that the existence of a conservation law (2) for potential vorticity is immediately evident. At the same time, the Maxwell equations represent differential equations that link the potential vorticity and the potential vorticity flux to the velocity (or vorticity) and the potential temperature—a link that is necessary in any formulation of dynamics in which the conservation law for potential vorticity is taken as a fundamental equation. Together with boundary conditions and with other dynamical equations, the Maxwell equations determine the primitive variables given the potential vorticity and its flux. The Maxwell equations provide a framework within which one can analyze the role of boundaries in the potential vorticity dynamics of flows that need be neither balanced nor hydrostatic and that can be influenced by diabatic processes and friction.

3. Generalized potential vorticity in coordinate-independent form

a. Boundary conditions and boundary sources

Since potential temperature can fluctuate at the surface, the boundary conditions for the fields \mathbf{D} and \mathbf{H} in the Maxwell equations (4) are generally inhomogeneous. These inhomogeneous boundary conditions at

⁵ Alternatively, the frictional term in the potential vorticity flux \mathbf{J} could have been written as $-\rho^{-1}\theta\nabla \times \mathbf{F}$, in which case the field \mathbf{H} would not contain a frictional term. However, an advantage of the functional forms (3) and (13) of the potential vorticity flux \mathbf{J} and field \mathbf{H} is that the frictional term $\rho^{-1}\nabla\theta \times \mathbf{F}$ in the potential vorticity flux has no cross-isentropic component and is therefore easier to expand in isentropic coordinates than a frictional term of the form $-\rho^{-1}\theta\nabla \times \mathbf{F}$ (cf. section 4).

the surface, or “immediately above” it, can be replaced by homogeneous boundary conditions “inside” the surface by inclusion of suitable boundary sources in the Maxwell equations (see, e.g., Morse and Feshbach 1953, chapter 7). Since the potential vorticity P and the potential vorticity flux \mathbf{J} are sources in the Maxwell equations, the boundary sources can be viewed as boundary contributions to the potential vorticity and to the potential vorticity flux.

For the fields \mathbf{D} and \mathbf{H} in Gauge II, we specify the homogeneous boundary conditions $\mathbf{D}_b = 0$ and $\mathbf{H}_b = 0$ inside the surface. The subscript b denotes quantities inside the surface, and the subscript s will denote quantities immediately above the surface. Taking the fields \mathbf{D}_b and \mathbf{H}_b to be zero inside the surface can be viewed as a consequence of taking the potential temperature θ_b and with it, for consistency, all other flow quantities in the thermodynamic equation (7) and momentum equation (8) to be zero inside the surface.⁶ Additionally, we specify the potential vorticity density ρP and the potential vorticity flux density $\rho \mathbf{J}$ to be zero inside the surface. The homogeneous boundary conditions for the fields \mathbf{D} and \mathbf{H} must be compensated by boundary contributions to the potential vorticity and to the potential vorticity flux.

b. Boundary contributions to the potential vorticity

At the interface between two media, the normal component of the electric displacement field has a discontinuity proportional to a surface charge density at the interface (Jackson 1975, section I.5). Analogously, the normal component of the field \mathbf{D} has, at the surface, a discontinuity proportional to a surface density of potential vorticity.

Given the fields \mathbf{D}_b inside the surface and \mathbf{D}_s immediately above the surface with upward normal \mathbf{n} , one can compute the surface potential vorticity that is required to force the normal component of the field \mathbf{D} from $\mathbf{n} \cdot \mathbf{D} = \mathbf{n} \cdot \mathbf{D}_b = 0$ inside the surface to $\mathbf{n} \cdot \mathbf{D} = \mathbf{n} \cdot \mathbf{D}_s$ immediately above the surface. Integrating the divergence equation (4a) over an infinitesimally small volume enclosing the surface and using Gauss’s theorem, one finds that the homogeneous boundary condition $\mathbf{n} \cdot \mathbf{D} = \mathbf{n} \cdot \mathbf{D}_b = 0$ must be compensated by including on the right-hand side of the divergence equation

(4a) a source with a surface density of potential vorticity equal to $\mathbf{n} \cdot (\mathbf{D}_s - \mathbf{D}_b) = \mathbf{n} \cdot \mathbf{D}_s$. One can think of this surface density of potential vorticity as the across-surface integral of the potential vorticity density ρS that belongs to the singular surface potential vorticity

$$S = \frac{\mathbf{n} \cdot \mathbf{D}_s}{\rho} \delta(z - z_s).$$

The surface potential vorticity S is concentrated in a delta-function potential vorticity sheet on the surface at $z = z_s(x, y)$. The Gauge II expression $\mathbf{D} = \theta \boldsymbol{\omega}_a$ leads to the surface potential vorticity

$$S = \frac{\boldsymbol{\omega}_a \cdot \mathbf{n}}{\rho} \theta \delta(z - z_s). \quad (14)$$

The divergence equation (4a) with inhomogeneous boundary condition $\mathbf{n} \cdot \mathbf{D} = \mathbf{n} \cdot \mathbf{D}_s$ and potential vorticity P is equivalent to a divergence equation (4a) with homogeneous boundary condition $\mathbf{n} \cdot \mathbf{D} = \mathbf{n} \cdot \mathbf{D}_b = 0$ and generalized potential vorticity

$$P_g = P + S. \quad (15)$$

Replacing the inhomogeneous boundary condition for the field \mathbf{D} by a homogeneous boundary condition is compensated by putting an idealized potential vorticity sheet with surface potential vorticity S on the surface.

c. Boundary contributions to the potential vorticity flux

At the interface between two media, the tangential component of the magnetic field has a discontinuity proportional to a current density at the interface (Jackson 1975, section I.5). Analogously, the tangential component of the field \mathbf{H} has, at the surface, a discontinuity proportional to a surface density of potential vorticity flux.

The surface flux of potential vorticity that is required to force the tangential component $\mathbf{n} \times \mathbf{H}$ of the field \mathbf{H} from $\mathbf{n} \times \mathbf{H} = \mathbf{n} \times \mathbf{H}_b = 0$ inside the surface to $\mathbf{n} \times \mathbf{H} = \mathbf{n} \times \mathbf{H}_s$ immediately above the surface can be found by integrating the second Maxwell equation $\nabla \times \mathbf{H} - \partial_t \mathbf{D} = \rho \mathbf{J}$ over an area perpendicular to and including the surface. In the limit of an infinitesimally small area, the area integral of the time derivative $\partial_t \mathbf{D}$ vanishes because $\partial_t \mathbf{D}$ is finite at the surface. Using Stokes’s theorem, one finds that the homogeneous boundary condition $\mathbf{n} \times \mathbf{H} = \mathbf{n} \times \mathbf{H}_b = 0$ must be compensated by including on the right-hand side of the Maxwell equation (4b) a source with a surface density of potential vorticity flux equal to $\mathbf{n} \times (\mathbf{H}_s - \mathbf{H}_b) = \mathbf{n} \times \mathbf{H}_s$. Analogous to the line of reasoning that led to the surface potential vorticity, one can think of the surface density of potential vorticity flux as the across-surface integral of the potential vorticity flux density $\rho \mathbf{K}$ that belongs to the singular potential vorticity flux

⁶ One could also specify a boundary condition of constant potential temperature $\theta_b \neq 0$ inside the surface, leading to fields $\mathbf{D}_b = \theta_b \boldsymbol{\omega}_a$ and $\mathbf{H}_b = \mathbf{u} \times \theta_b \boldsymbol{\omega}_a + c_p T(\theta_b/\theta) \nabla \theta + \theta_b \mathbf{F}$ inside the surface. Such an alternative boundary condition corresponds to specifying homogeneous boundary conditions $\mathbf{D}'_b = 0$ and $\mathbf{H}'_b = 0$ for the fields $\mathbf{D}' = (\theta'/\theta) \mathbf{D}$ and $\mathbf{H}' = (\theta'/\theta) \mathbf{H}$, obtained by replacing the potential temperature θ by the fluctuation $\theta' = \theta - \theta_b$ about the constant potential temperature θ_b . Equivalently, the fields \mathbf{D}' and \mathbf{H}' can be obtained from the fields \mathbf{D} and \mathbf{H} by a gauge transformation of the form (5) with $\mathbf{A} = -\theta_b \mathbf{u}_a$. The ambiguities due to the invariance of the Maxwell equations under such gauge transformations are discussed in section 3e.

$$\mathbf{K} = \frac{\mathbf{n} \times \mathbf{H}_s}{\rho} \delta(z - z_s).$$

This surface potential vorticity flux is concentrated on the surface and has only components tangential to the surface. Substituting for the field \mathbf{H} from the Gauge II expression (13) and using the fact that, at the surface, the normal component $\mathbf{n} \cdot \mathbf{u}$ of the velocity \mathbf{u} vanishes, one obtains the surface potential vorticity flux

$$\mathbf{K} = \mathbf{uS} + \mathbf{K}_{bc} + \mathbf{K}_f \quad (16a)$$

with baroclinic component

$$\mathbf{K}_{bc} = \rho^{-1} c_p T (\mathbf{n} \times \nabla \theta) \delta(z - z_s) \quad (16b)$$

and frictional component

$$\mathbf{K}_f = \rho^{-1} \theta (\mathbf{n} \times \mathbf{F}) \delta(z - z_s). \quad (16c)$$

The surface potential vorticity flux \mathbf{K} consists of an advective component \mathbf{uS} , of a baroclinic component \mathbf{K}_{bc} that is directed along lines of intersection between isentropes and the surface, and of a frictional component \mathbf{K}_f .⁷ The baroclinic component \mathbf{K}_{bc} has its origin in the contribution of the baroclinicity vector $\rho^{-2} \nabla \rho \times \nabla p$ to the vector potential \mathbf{H} . With a no-slip boundary condition at the surface, the velocity \mathbf{u} along the surface vanishes, and the surface potential vorticity flux \mathbf{K} consists only of the nonadvective components \mathbf{K}_{bc} and \mathbf{K}_f . The advective component of the surface potential vorticity flux might, nevertheless, be of practical relevance. For example, numerical atmosphere models typically use not a no-slip boundary condition, but a drag-law boundary condition, so the advective component of the surface potential vorticity flux might not be negligible in a generalized potential vorticity budget of such a model. The advective component of the surface potential vorticity flux is proportional to the surface heat flux, and horizontal heat fluxes are significant down to the lowest levels of typical general circulation models and down to the lowest atmospheric levels for which observational data are available (cf. Held and Schneider 1999).

The Maxwell equation (4b) with inhomogeneous boundary condition $\mathbf{n} \times \mathbf{H} = \mathbf{n} \times \mathbf{H}_s$ and potential vorticity flux \mathbf{J} is equivalent to a Maxwell equation (4b) with homogeneous boundary condition $\mathbf{n} \times \mathbf{H} = \mathbf{n} \times \mathbf{H}_b = 0$ and generalized potential vorticity flux

$$\mathbf{J}_g = \mathbf{J} + \mathbf{K}. \quad (17)$$

Replacing the inhomogeneous boundary condition for the field \mathbf{H} by a homogeneous boundary condition is compensated by the added flux \mathbf{K} in the idealized potential vorticity sheet on the surface.

⁷ If the frictional term in the interior potential vorticity flux \mathbf{J} would have been written as $-\rho^{-1} \theta \nabla \times \mathbf{F}$, the field \mathbf{H} and, with it, the surface potential vorticity flux \mathbf{K} would not contain frictional components (cf. footnote 5).

d. Conservation of generalized potential vorticity

The generalized potential vorticity P_g and flux \mathbf{J}_g contain the boundary contributions to the potential vorticity balance. The original Maxwell equations (4) with inhomogeneous boundary conditions at the surface and with the interior potential vorticity P and flux \mathbf{J} as sources are equivalent to Maxwell equations with homogeneous boundary conditions inside the surface and with the generalized potential vorticity P_g and flux \mathbf{J}_g as sources. Adding the time derivative of the first Maxwell equation (4a) to the divergence of the second equation (4b) yields the conservation law

$$\partial_t (\rho P_g) + \nabla \cdot (\rho \mathbf{J}_g) = 0 \quad (18)$$

for the generalized potential vorticity P_g .

The conservation law (18) for generalized potential vorticity is similar to the conservation law (2) for interior potential vorticity. In contrast to interior potential vorticity, however, generalized potential vorticity is not, in general, materially conserved in adiabatic and frictionless flows. The baroclinic component (16b) of the surface potential vorticity flux can redistribute generalized potential vorticity nonadvectively along the surface, even in adiabatic and frictionless flows.

Nevertheless, the integral of the generalized potential vorticity density ρP_g over the volume of the atmosphere is conserved. Integrating the conservation law (18) over the volume V of the atmosphere and using Gauss's theorem with the boundary condition $\mathbf{n} \cdot (\rho \mathbf{J}_g) = 0$ inside the surface and with the assumption that the flux density $\rho \mathbf{J}_g$ vanish at the top of the atmosphere, one obtains

$$\partial_t \int_V \rho P_g \, d\mathbf{x} = 0,$$

or, equivalently,

$$\partial_t \int_V \rho P \, d\mathbf{x} = -\partial_t \int_V \rho S \, d\mathbf{x}.$$

Any increase in the volume-integrated interior potential vorticity density ρP is compensated by a decrease in the volume-integrated surface potential vorticity density ρS , and vice versa. The conservation law (2) for interior potential vorticity implies that the volume-integrated interior potential vorticity density ρP can change only if the integral of the normal flux component

$$\mathbf{n} \cdot (\rho \mathbf{J}) = -Q(\mathbf{n} \cdot \boldsymbol{\omega}_a) + \mathbf{n} \cdot (\nabla \theta \times \mathbf{F})$$

over the surface area A is nonzero,

$$\begin{aligned} \partial_t \int_V \rho P \, d\mathbf{x} &= - \int_A [Q(\mathbf{n} \cdot \boldsymbol{\omega}_a) - \mathbf{n} \cdot (\nabla \theta \times \mathbf{F})] \, dA \\ &= -\partial_t \int_V \rho S \, d\mathbf{x}. \end{aligned} \quad (19)$$

Since the normal component of the velocity vanishes at

the surface, only diabatic and frictional processes at the surface can effect changes in the volume-integrated interior potential vorticity density ρP . Given that changes in the volume-integrated interior potential vorticity density ρP are compensated by opposing changes in volume-integrated surface potential vorticity density ρS , diabatic and frictional processes at the surface can be viewed as converting surface potential vorticity into interior potential vorticity, and vice versa. The surface integral of the normal flux component $\mathbf{n} \cdot (\rho \mathbf{J}_g) = \mathbf{n} \cdot (\rho \mathbf{J})$ indicates the conversion rate.

e. Alternative generalized potential vorticity functionals

The functional forms of the generalized potential vorticity $P_g = P + S$ and of the generalized potential vorticity flux $\mathbf{J}_g = \mathbf{J} + \mathbf{K}$ are not unique because the functional forms of the surface potential vorticity S and of the surface potential vorticity flux \mathbf{K} depend on the gauge of the fields \mathbf{D} and \mathbf{H} .

We chose Gauge II because, in this gauge, the functional forms of the surface potential vorticity (14) and of the surface potential vorticity flux (16) resemble the functional forms of the interior potential vorticity (1) and of the interior potential vorticity flux (3). For example, the surface potential vorticity (14) is proportional to the absolute vorticity component $\boldsymbol{\omega}_a \cdot \mathbf{n}$ normal to the surface, while the interior potential vorticity (1) is proportional to the absolute vorticity component $\boldsymbol{\omega}_a \cdot \nabla \theta$ normal to isentropes. And like the interior potential vorticity flux (3), the surface potential vorticity flux (16) contains an advective component, which legitimizes its interpretation as a flux.

Alternatively, however, we could have chosen a gauge in which, for example, a field $\mathbf{D}' = \theta' \boldsymbol{\omega}_a$ is defined with the potential temperature fluctuation $\theta' = \theta - \theta_0$ about a constant reference potential temperature θ_0 in place of the absolute potential temperature θ (cf. footnote 6). Such a gauge suggests itself when a reference potential temperature is given, such as is the case in Boussinesq flows (see below). One obtains the alternative gauge \mathbf{D}' and \mathbf{H}' from Gauge II by a transformation of the form (5) with $\mathbf{A} = -\theta_0 \mathbf{u}_a$. Replacing inhomogeneous with homogeneous boundary conditions in this alternative gauge results in a surface potential vorticity $S' = (\theta'/\theta)S$ and a surface potential vorticity flux $\mathbf{K}' = (\theta'/\theta)\mathbf{K}$. All of the above statements about the conservation of generalized potential vorticity and about the structure of the surface potential vorticity and the surface potential vorticity flux remain valid if the surface potential vorticity S and the flux \mathbf{K} are replaced by the alternative functionals S' and \mathbf{K}' . This arbitrariness in the definition of the surface potential vorticity and of the surface potential vorticity flux means that their absolute values by themselves carry no dynamical significance.

Other gauges are possible and may be convenient in some contexts. However, the conservation law (18) for

generalized potential vorticity does not depend on the gauge chosen. The gauge invariance of the Maxwell equations (4) translates into gauge invariance of the conservation law (18) for generalized potential vorticity.

f. Adiabatic and frictionless Boussinesq flows

The surface potential vorticity S and the surface potential vorticity flux \mathbf{K} for adiabatic and frictionless Boussinesq flows can be derived in a similar manner as in the general case. In the Boussinesq approximation, the density ρ in the potential vorticity (1) is taken to be equal to a constant reference density ρ_0 , and the potential vorticity is defined with the potential temperature fluctuation $\theta' = \theta - \theta_0$ about a constant reference potential temperature θ_0 in place of the absolute potential temperature θ :

$$P = \frac{\boldsymbol{\omega}_a \cdot \nabla \theta'}{\rho_0}. \quad (20)$$

Correspondingly, the surface potential vorticity becomes

$$S = \frac{\boldsymbol{\omega}_a \cdot \mathbf{n}}{\rho_0} \theta' \delta(z - z_s). \quad (21)$$

The surface potential vorticity flux for adiabatic and frictionless Boussinesq flows consists of the advective component $\mathbf{u}S$ and of a baroclinic component \mathbf{K}_{bc} . Which form the baroclinic component \mathbf{K}_{bc} takes can be seen by going back to the derivation of the surface potential vorticity flux in the general case. The baroclinic component has its origin in the baroclinicity vector $-\nabla \times \rho^{-1} \nabla p$. The baroclinic component resulted from writing the term $-\theta(\nabla \times \rho^{-1} \nabla p)$ in the expansion of the time derivative $\partial_t \mathbf{D}$ as the curl of the vector field $c_p T \nabla \theta$, making this field $c_p T \nabla \theta$ part of the field \mathbf{H} [Eq. (13)], and taking the tangential component $\mathbf{n} \times \mathbf{H}$ to determine the surface potential vorticity flux. In the Boussinesq approximation, the baroclinicity vector is $\nabla \times (g\theta'/\theta_0)\mathbf{k}$, with vertical unit vector \mathbf{k} , and the term $-\theta(\nabla \times \rho^{-1} \nabla p)$ in the general case becomes $\theta'(\nabla \times (g\theta'/\theta_0)\mathbf{k}) = \nabla \times (g\theta'^2/2\theta_0)\mathbf{k}$ (cf. Salmon 1998, chapter 2.16). Consequently, the baroclinic component of the surface potential vorticity flux becomes

$$\mathbf{K}_{bc} = \frac{1}{\rho_0} \frac{g\theta'^2}{2\theta_0} (\mathbf{n} \times \mathbf{k}) \delta(z - z_s). \quad (22)$$

The baroclinic component of the surface potential vorticity flux is quadratic in potential temperature fluctuations θ' and hence would not appear in a linearized Boussinesq system. Instead of being directed along lines of intersection between isentropes and the surface, as the baroclinic component (16b) in the general case, the baroclinic component of the Boussinesq surface potential vorticity flux is directed along lines of constant surface elevation. At a flat surface ($\mathbf{n} = \mathbf{k}$), the baroclinic component vanishes, and the surface potential vorticity

flux reduces to the advective flux $\mathbf{K} = \mathbf{u}S$. Since the Boussinesq approximation is often an adequate approximation for atmospheric flows near the surface—say, within the planetary boundary layer—the vanishing of the baroclinic component \mathbf{K}_{bc} of the surface potential vorticity flux for Boussinesq flows over a flat surface suggests that this component is only important if topography exerts a significant influence on the flow. The baroclinic component \mathbf{K}_{bc} of the surface potential vorticity flux and topographic effects will be discussed in more detail in section 5.

g. Quasigeostrophic Boussinesq flows over flat surface

Taking the quasigeostrophic limit of the Maxwell equations for Boussinesq flows, one finds that, for quasigeostrophic Boussinesq flows over a flat surface ($z_s = 0$), the normal component $\boldsymbol{\omega}_a \cdot \mathbf{n}$ of the absolute vorticity in the surface potential vorticity (21) must be approximated by a constant reference value f_0 of the Coriolis parameter f , so that the surface potential vorticity becomes

$$S = \frac{f_0}{\rho_0} \theta' \delta(z).$$

This surface potential vorticity corresponds to that boundary contribution to the quasigeostrophic potential vorticity with which Bretherton (1966) replaced an inhomogeneous thermodynamic boundary condition.

The baroclinic component (22) of the surface potential vorticity flux vanishes for Boussinesq flows over a flat surface, and, for adiabatic and frictionless quasigeostrophic flows, the surface potential vorticity flux reduces to the advective flux

$$\mathbf{K} = \mathbf{u}_g S = \frac{f_0}{\rho_0} \mathbf{u}_g \theta' \delta(z),$$

where the advecting velocity \mathbf{u}_g is the geostrophic velocity. For quasigeostrophic Boussinesq flows over a flat surface, the surface potential vorticity flux is proportional to the geostrophic surface heat flux. This surface potential vorticity flux corresponds to the boundary contribution to the quasigeostrophic potential vorticity flux discussed by Bretherton (1966).

4. Generalized potential vorticity in isentropic coordinates

We assume that for each time t and at each point in the (x, y) plane, the potential temperature θ is a strictly monotonic function of height z , so that the instantaneous thermal stratification is everywhere statically stable ($\partial_z \theta > 0$), and the potential temperature can be used as the vertical coordinate in an isentropic coordinate system. We adopt the hydrostatic approximation and carry out the analysis within the framework of the primitive equations.

We will determine the generalized potential vorticity and the components of the generalized potential vorticity flux in isentropic coordinates by expanding the coordinate-independent expressions of sections 2 and 3 in isentropic coordinates. Expressions for the interior potential vorticity (1) and for the interior potential vorticity flux (3) in isentropic coordinates are well known; they are usually derived from the equations of motion in isentropic coordinates [see, e.g., Salmon (1998, chapter 2.18); Andrews et al. (1987, chapter 3.8)]. The technique of expanding the coordinate-independent expressions in isentropic coordinates has the advantage of being applicable to the singular surface potential vorticity and its flux, without it being necessary to go back to the equations of motion to deduce the representation of these quantities in isentropic coordinates.

Isentropic coordinates are nonorthogonal, so contravariant and covariant vector components must be distinguished (see, e.g., Arfken 1985, chapter 3). We use the notation $(a^x, a^y, a^\theta)^\theta$ for the contravariant components of a vector \mathbf{a} in isentropic coordinates. The contravariant horizontal components $a^x = \mathbf{a} \cdot \mathbf{i}$ and $a^y = \mathbf{a} \cdot \mathbf{j}$ are equal to the local Cartesian components of the vector \mathbf{a} , the local Cartesian unit vectors \mathbf{i} and \mathbf{j} being directed eastward and northward.⁸ The contravariant cross-isentropic component $a^\theta = \mathbf{a} \cdot \nabla \theta$ is the scalar product of the vector \mathbf{a} and the potential temperature gradient $\nabla \theta$.

a. Generalized potential vorticity

The generalized potential vorticity takes a particularly simple form in isentropic coordinates. In the primitive equations, the planetary vorticity $2\boldsymbol{\Omega}$ is approximated by its local vertical component $f\mathbf{k}$, and in the hydrostatic approximation, horizontal derivatives of the vertical velocity in the relative vorticity $\boldsymbol{\omega}_r = \nabla \times \mathbf{u}$ are neglected compared with vertical derivatives of the horizontal velocity. The absolute vorticity hence becomes the sum $\boldsymbol{\omega}_a = f\mathbf{k} + \nabla \times \mathbf{v}$ of the planetary vorticity $f\mathbf{k}$ and the relative vorticity $\nabla \times \mathbf{v}$ of the horizontal flow $\mathbf{v} = (u, v, 0)$. The relative vorticity of the horizontal flow can be represented in isentropic coordinates as

$$\nabla \times \mathbf{v} = h^{-1}(-\partial_\theta v, \partial_\theta u, \partial_x v - \partial_y u)^\theta, \quad (23)$$

where the horizontal derivatives ∂_x and ∂_y are to be understood as derivatives along isentropes, and the scale factor

$$h = \partial_\theta z, \quad (24)$$

an inverse measure of static stability, is the Jacobian $h = \partial(x, y, z)/\partial(x, y, \theta)$ of the transformation from Cartesian coordinates to isentropic coordinates. By the rep-

⁸ As horizontal coordinates, we use local Cartesian coordinates in what follows. The transformation of the horizontal coordinates from local Cartesian coordinates to spherical coordinates is straightforward.

resentation (23) of the relative vorticity in isentropic coordinates, the interior potential vorticity density $\rho P = \boldsymbol{\omega}_a \cdot \nabla \theta$ —the contravariant cross-isentropic component of the absolute vorticity—is $\boldsymbol{\omega}_a \cdot \nabla \theta = h^{-1}(f + \zeta_\theta)$, where $\zeta_\theta = \partial_x v - \partial_y u$ is the relative vorticity of the horizontal flow \mathbf{v} along isentropes. In the hydrostatic approximation, the density is $\rho = -g^{-1} \partial_z p$, and the product of density ρ and scale factor h is the isentropic density

$$\rho_\theta = \rho h = -g^{-1} \partial_\theta p.$$

Combining the density ρ and the scale factor h in the potential vorticity density $\boldsymbol{\omega}_a \cdot \nabla \theta = h^{-1}(f + \zeta_\theta)$ yields the well-known interior potential vorticity

$$P = \frac{f + \zeta_\theta}{\rho_\theta} \mathcal{H}(\theta - \theta_s). \quad (25)$$

The step function

$$\mathcal{H}(\theta - \theta_s) = \begin{cases} 1 & \text{if } \theta > \theta_s \\ 0 & \text{if } \theta < \theta_s \end{cases}$$

indicates that, according to our conventions, the interior potential vorticity P contributes to the generalized potential vorticity P_g only above the surface, on isentropes with potential temperature θ greater than the surface potential temperature $\theta_s(x, y, t)$.

An isentropic-coordinate representation of the boundary contribution S to the generalized potential vorticity $P_g = P + S$ can be found in a similar way. Under the assumption of static stability, the delta function $\delta(z - z_s)$ transforms according to $\delta(z - z_s) = h^{-1} \delta(\theta - \theta_s)$, where the scale factor h is to be evaluated immediately above the surface (since the surface potential vorticity and its flux only contain quantities immediately above the surface; cf. sections 3b and 3c). Combining the density ρ in the surface potential vorticity (14) with the scale factor h from the transformation of the delta function yields the isentropic-coordinate expression

$$S = \frac{\boldsymbol{\omega}_a \cdot \mathbf{n}}{\rho_\theta} \theta \delta(\theta - \theta_s) \quad (26)$$

for the surface potential vorticity.

Within the approximations of the primitive equations and under the assumption of static stability, the generalized potential vorticity (15) is the sum of the interior potential vorticity (25) and the surface potential vorticity (26) in isentropic coordinates.

b. Generalized potential vorticity flux

The generalized potential vorticity flux in isentropic coordinates can likewise be found by expanding the vectors and differential operators of the coordinate-independent interior potential vorticity flux (3) and surface potential vorticity flux (16). The isentropic-coordinate representation of the term $\rho^{-1} Q \boldsymbol{\omega}_a$ in the interior potential vorticity flux (3) follows by expanding the

relative vorticity $\nabla \times \mathbf{v}$ with the help of the expression (23); the frictional force per unit mass \mathbf{F} is assumed to have only horizontal components F^x and F^y ; and products of the scale factor h and the density ρ are combined to the isentropic density ρ_θ . The cross-isentropic components of the advective flux $\mathbf{u}P$ and of the diabatic term $\rho^{-1} Q \boldsymbol{\omega}_a$ cancel because, by the thermodynamic equation (7), the contravariant cross-isentropic component $\mathbf{u} \cdot \nabla \theta$ of the velocity is the heating rate Q , and the contravariant cross-isentropic component $\boldsymbol{\omega}_a \cdot \nabla \theta$ of the absolute vorticity is the potential vorticity density ρP . Combining all terms, one finds the well-known interior potential vorticity flux (cf. Haynes and McIntyre 1987)

$$\mathbf{J} = (u, v, 0)^\theta P + \mathbf{J}_Q + \mathbf{J}_F, \quad (27a)$$

with diabatic flux

$$\mathbf{J}_Q = \rho_\theta^{-1} Q (\partial_\theta v, -\partial_\theta u, 0)^\theta \mathcal{H}(\theta - \theta_s) \quad (27b)$$

and frictional flux

$$\mathbf{J}_F = \rho_\theta^{-1} (-F^y, F^x, 0)^\theta \mathcal{H}(\theta - \theta_s). \quad (27c)$$

Even in the presence of diabatic heating and friction, the interior potential vorticity flux has no cross-isentropic component. Therefore, the impermeability theorem holds: interior potential vorticity can only be redistributed along isentropes but cannot be transferred across isentropes (Haynes and McIntyre 1987).

In order to represent the surface potential vorticity flux (16) in isentropic coordinates, we use for the unit normal vector at the surface $z = z_s(x, y)$ the explicit representation

$$\mathbf{n} = \mu \nabla(z - z_s) = \mu (\mathbf{k} - \nabla z_s),$$

with normalization factor

$$\mu = (1 + \|\nabla z_s\|^2)^{-1/2}.$$

The hydrostatic approximation is only justifiable if the horizontal scale of the topography $z_s(x, y)$ is much greater than the vertical scale, such that $\mu \approx 1$. For consistency with the hydrostatic approximation, we should set the normalization factor μ equal to one. But with the understanding that the hydrostatic approximation would be inappropriate if the normalization factor μ were significantly less than one, we retain the normalization factor μ in the following equations as a marker of where topographic effects can play a role.

With the explicit representation of the normal vector \mathbf{n} , the surface potential vorticity flux can be expanded in isentropic coordinates term-by-term. The delta functions in the surface potential vorticity flux are transformed in the same way as above: $\delta(z - z_s) = h^{-1} \delta(\theta - \theta_s)$. And horizontal derivatives of the potential temperature θ at constant height z are transformed into horizontal derivatives of the height z at constant potential temperature θ by means of the relation $\partial_{x_i} \theta|_z = -h^{-1} \partial_{x_i} z|_\theta$ for $x_i = x, y$. One obtains the surface potential vorticity flux

$$\mathbf{K} = (u, v, Q)^\theta S + \mathbf{K}_{bc} + \mathbf{K}_f, \quad (28a)$$

with baroclinic component

$$\mathbf{K}_{bc} = \frac{\mu E}{\rho_\theta} (-\theta_s^y, \theta_s^x, 0)^\theta \theta \delta(\theta - \theta_s) \quad (28b)$$

and frictional component

$$\mathbf{K}_f = \frac{\mu}{\rho_\theta} (-F^y, F^x, F^x \theta_s^y - F^y \theta_s^x)^\theta \theta \delta(\theta - \theta_s), \quad (28c)$$

where

$$\theta_s^x = -h^{-1} \partial_x (z - z_s) \quad \text{and} \quad \theta_s^y = -h^{-1} \partial_y (z - z_s) \quad (29)$$

are the derivatives of the surface potential temperature $\theta_s(x, y, t)$ with respect to x and y . In the baroclinic component (28b), the specific enthalpy $c_p T$ has been written as the product $c_p T = \theta E$ of potential temperature θ and Exner function $E = c_p (p/p_0)^\kappa$, p_0 being a constant reference pressure.

Within the approximations of the primitive equations and under the assumption of static stability, the generalized potential vorticity flux (17) is the sum of the interior potential vorticity flux (27) and the surface potential vorticity flux (28) in isentropic coordinates. Since both the advective component (28a) and the frictional component (28c) of the surface potential vorticity flux have cross-isentropic components, the impermeability theorem does not hold for the generalized potential vorticity flux in Gauge II. The generalized potential vorticity in Gauge II can be transferred across isentropes by diabatic heating and friction at the surface. However, whether the generalized potential vorticity flux has a cross-isentropic component is gauge-dependent; for example, the generalized potential vorticity flux has no cross-isentropic component in Gauge I.

c. Conservation of generalized potential vorticity

The conservation law (18) for generalized potential vorticity becomes in isentropic coordinates

$$\partial_t(\rho_\theta P_g) + \partial_x(\rho_\theta J_g^x) + \partial_y(\rho_\theta J_g^y) + \partial_\theta(\rho_\theta J_g^\theta) = 0, \quad (30)$$

where $(J_g^x, J_g^y, J_g^\theta)^\theta$ are the contravariant components of the generalized potential vorticity flux \mathbf{J}_g .⁹

⁹ The conservation law (30) results from the representation

$$\nabla \cdot \mathbf{a} = \frac{1}{h} [\partial_x(ha^x) + \partial_y(ha^y) + \partial_\theta(ha^\theta)]$$

of the divergence of a vector $\mathbf{a} = (a^x, a^y, a^\theta)^\theta$ in isentropic coordinates (cf. Arfken 1985, chapter 3.9). Since the scale factor h can be viewed as the Jacobian $h = \partial(x, y, z, t)/\partial(x, y, \theta, t)$ of the four-dimensional transformation from (x, y, z, t) -coordinates to (x, y, θ, t) -coordinates, and since it is the Jacobian h of the transformation from Cartesian to isentropic coordinates that appears in the representation of the divergence in isentropic coordinates, the explicit time derivative $\partial_t(\rho P)$ can be viewed as being part of a four-dimensional divergence operator and can, like the space derivatives, be written as $h^{-1} \partial_t(h\rho P) = h^{-1} \partial_t(\rho_\theta P)$; hence the conservation law (30) in isentropic coordinates.

5. Baroclinic induction of generalized potential vorticity

In adiabatic and frictionless flows, the surface potential vorticity flux (28) is the sum $\mathbf{K} = (u, v, 0)^\theta S + \mathbf{K}_{bc}$ of the advective component $(u, v, 0)^\theta S$ and the baroclinic component \mathbf{K}_{bc} . The presence of the nonadvective baroclinic component \mathbf{K}_{bc} of the surface potential vorticity flux implies that surface potential vorticity, and through it generalized potential vorticity, can be induced baroclinically.

a. Origin of baroclinic component of surface potential vorticity flux

The baroclinic component \mathbf{K}_{bc} of the surface potential vorticity flux arises because of differences between the interior potential vorticity $(f + \zeta_\theta)/\rho_\theta$ and the quantity $(\boldsymbol{\omega}_a \cdot \mathbf{n})/\rho_\theta$ to which the surface potential vorticity (26) is proportional. Denoting the relative vorticity component perpendicular to the surface by $\zeta_\sigma = \boldsymbol{\omega}_r \cdot \mathbf{n}$, one can write $(\boldsymbol{\omega}_a \cdot \mathbf{n})/\rho_\theta = (\mu f + \zeta_\sigma)/\rho_\theta$. In the hydrostatic approximation, the normalization factor μ is equal to one, so that the interior potential vorticity $(f + \zeta_\theta)/\rho_\theta$ and the quantity $(\boldsymbol{\omega}_a \cdot \mathbf{n})/\rho_\theta = (f + \zeta_\sigma)/\rho_\theta$ differ only by the relative vorticity factors: the interior potential vorticity

$$P = \frac{f + \zeta_\theta}{\rho_\theta} \mathcal{H}(\theta - \theta_s)$$

contains the relative vorticity ζ_θ of the flow along isentropes; the surface potential vorticity

$$S = \frac{f + \zeta_\sigma}{\rho_\theta} \theta \delta(\theta - \theta_s)$$

contains the relative vorticity ζ_σ of the flow along the surface. Because of this difference in the relative vorticity factors, the surface potential vorticity flux has a nonadvective baroclinic component in adiabatic and frictionless flows, while the interior potential vorticity flux is purely advective in such flows.

The regular part of the conservation law (30) describes the time evolution of the isentropic relative vorticity ζ_θ . In the absence of diabatic heating and friction, the interior potential vorticity flux (27) reduces to an advective flux along isentropes, so the interior potential vorticity $(f + \zeta_\theta)/\rho_\theta$ is materially conserved and is, in particular, materially conserved in adiabatic and frictionless flows along isentropes at or immediately above the surface. The singular part of the conservation law (30) describes the time evolution of the relative vorticity ζ_σ on isentropes $\theta = \theta_s(x, y, t)$ at the surface. Even in the absence of diabatic heating and friction, the surface potential vorticity flux (28) contains the nonadvective baroclinic component \mathbf{K}_{bc} , so the quantity $(f + \zeta_\sigma)/\rho_\theta$ is not, in general, materially conserved in adiabatic and frictionless flows along isentropes at the surface.

The quantity $(f + \zeta_\sigma)/\rho_\theta$ is not materially conserved

because baroclinicity at the surface can affect the relative vorticity ζ_σ of the surface flow. The contribution $\nabla \cdot (\rho \mathbf{K}_{bc})$ of the baroclinic component (16b) to the divergence of the surface potential vorticity flux density is equal, up to factors that are constant along isentropes at the surface, to the downward normal component

$$-\mathbf{n} \cdot \nabla \times (c_p T \nabla \log \theta) = \nabla \cdot \left(\frac{c_p T}{\theta} \mathbf{n} \times \nabla \theta \right)$$

of the baroclinicity vector

$$-\nabla \times \rho^{-1} \nabla p = \nabla \times c_p T \nabla \log \theta.$$

The component of the baroclinicity vector normal to the surface does not generally vanish but affects the surface potential vorticity S via the relative vorticity ζ_σ of the surface flow. In contrast, the component of the baroclinicity vector normal to isentropes vanishes, so the isentropic relative vorticity ζ_θ is not affected by baroclinicity and the flux of interior potential vorticity is purely advective in adiabatic and frictionless flows, including flows along isentropes at or immediately above the surface.

The baroclinic component of the surface potential vorticity flux, then, is due to the difference between the relative vorticities ζ_θ and ζ_σ .

b. Scale analysis for small Rossby numbers

For hydrostatic flows with small Rossby numbers, the ratio of the difference $\zeta_\theta - \zeta_\sigma$ between the relative vorticities to the relative vorticities ζ_θ and ζ_σ themselves scales like the ratio Fr^2/Ro of squared Froude number $\text{Fr} = U/(NH)$ to Rossby number $\text{Ro} = U/(fL)$, where U is a velocity scale, N the Brunt-Väisälä frequency, H a height scale, and L a length scale. Since the baroclinic component \mathbf{K}_{bc} of the surface potential vorticity flux is due to the difference $\zeta_\theta - \zeta_\sigma$ between the relative vorticities, it is of order $O(\text{Fr}^2/\text{Ro})$ compared with the advective component $(u, v, Q)^\theta S$. That is, for flows with small Rossby numbers, only if $\text{Fr}^2 \gtrsim \text{Ro}$ can surface baroclinicity lead to significant deviations from material conservation of generalized potential vorticity.

For quasigeostrophic flows on the scale of the Rossby radius $L_{\text{Ro}} = NH/f$, the Froude number Fr is of the same order as the Rossby number Ro , and the baroclinic component \mathbf{K}_{bc} of the surface potential vorticity flux is of order $O(\text{Ro})$ compared with the advective component $(u, v, Q)^\theta S$. The baroclinic component of the surface potential vorticity flux hence is negligible in quasigeostrophic scaling.¹⁰

As shown above (section 3f), the baroclinic component of the surface potential vorticity flux vanishes for Boussinesq flows over a flat surface, irrespective of

quasigeostrophic scaling. In quasigeostrophic scaling, the baroclinic component of the surface potential vorticity flux is negligible, irrespective of topography. Hence, in large-scale atmospheric flows, the baroclinic component of the surface potential vorticity flux will often be negligible.

c. Example: Wake formation in flows past a mountain

One example of a flow in which nongeostrophic effects and the baroclinic component of the surface potential vorticity flux can play a significant role is stratified flow past a mountain. Smolarkiewicz and Rotunno (1989) have demonstrated with numerical simulations that, in a stratified flow that is irrotational upstream of an isolated mountain, a wake with a pair of lee vortices can form downstream of the mountain even when the boundary condition at the mountain is a free-slip condition. A free-slip condition at the surface implies that there cannot be a usual frictional boundary layer from which vorticity could be transferred into the wake in the interior of the flow. In place of frictional processes, baroclinic effects have been linked to the induction of wake vorticity in flows past a mountain with a free-slip surface (Smolarkiewicz and Rotunno 1989; Rotunno et al. 1999; Epifanio and Durran 2002a,b). Even in the presence of baroclinicity, however, interior potential vorticity is materially conserved in adiabatic and frictionless flows, and so baroclinic effects alone cannot be responsible for the induction of potential vorticity in a wake; the interior potential vorticity would have to remain zero throughout adiabatic and frictionless flows past a mountain if it is zero upstream of the mountain. Yet, in simulations with weak frictional and thermal dissipation, a flow with zero interior potential vorticity upstream of a mountain with a free-slip surface can develop a wake with nonzero interior potential vorticity downstream of the mountain (see, e.g., Schär and Durran 1997; Rotunno et al. 1999; Epifanio and Durran 2002b). The nonzero interior potential vorticity in the wake implies that dissipative processes, however weak, must be active somewhere in the flow. But the extent to which baroclinic effects and dissipative processes play a role in the wake formation has been the subject of controversy (see, e.g., Smith 1989; Schär and Durran 1997; Rotunno et al. 1999; Epifanio and Durran 2002a,b). The generalized potential vorticity concept allows for a scenario of how a wake can form in weakly dissipative flows past a mountain.

Figure 1 shows the time evolution of the surface potential vorticity S and of the interior potential vorticity P on two isentropes during the spinup of a mountain wake from a potential-flow initial condition. The figure is based on a simulation by Rotunno et al. (1999) of a Boussinesq flow past an isolated mountain. In the simulation, the planetary vorticity is zero, and the Brunt-Väisälä frequency N , the velocity $\mathbf{u} = (U, 0, 0)$, and the surface potential temperature θ_s are uniform far up-

¹⁰ Strictly speaking, it is the divergence $\nabla \cdot (\rho \mathbf{K}_{bc}) = \partial_x(\rho_\theta K_{bc}^x)|_\theta + \partial_y(\rho_\theta K_{bc}^y)|_\theta$ that is negligible in quasigeostrophic scaling. The baroclinic component \mathbf{K}_{bc} may have a nondivergent component that is irrelevant for the transfer of surface potential vorticity.

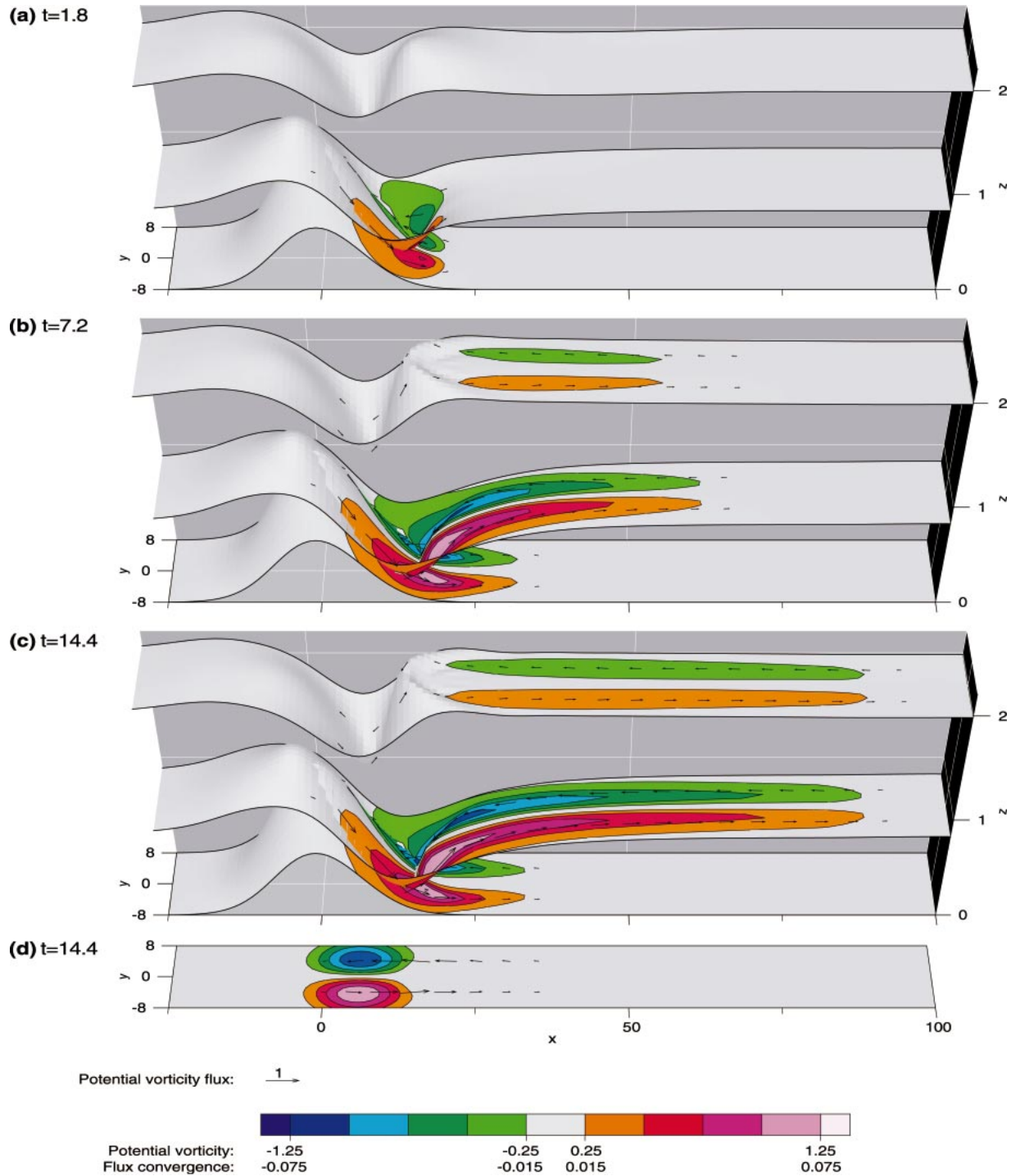


FIG. 1. Generalized potential vorticity in simulated Boussinesq flow past a mountain. The flow impinges along the x axis (from the left) upon a radially symmetric mountain at the coordinate origin. (a), (b), (c) Colored contours indicate the interior potential vorticity (20) on the isentropes $\theta' = 0.8$ and $\theta' = 2.0$ and the surface potential vorticity (21) at the mountain surface for three different times t after the start of the simulation from a potential-flow initial condition. (d) Colored contours indicate, projected onto the (x, y) plane, the convergence $-\nabla \cdot (\rho_0 \mathbf{K}_{bc})$ of the baroclinic component (22) of the surface potential vorticity flux. Vectors indicate the magnitude and direction of the advective interior potential vorticity flux $(u, v, 0)^T P$ along the isentropes and of the advective surface potential vorticity flux $\mathbf{u}S$. Quantities are given in units of the scales listed in Table 1. The delta function $\delta(z - z_c)$ in the surface potential vorticity was replaced by the inverse height scale $1/H$, so that the surface potential vorticity is finite and of magnitude comparable with the interior potential vorticity.

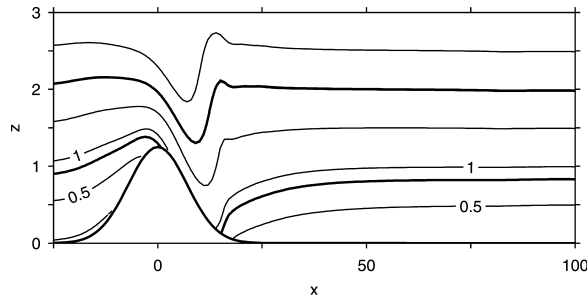


FIG. 2. Isentropes in the $y = 0$ symmetry plane at time $t = 14.4$. The thin lines represent isentropes ($\theta' = \text{const}$) with a contour interval of $\Delta\theta' = 0.5$. The thick lines represent the surface and the isentropes $\theta' = 0.8$ and $\theta' = 2.0$ on which the generalized potential vorticity is shown in Fig. 1. Potential temperature fluctuations θ' are given in units of the scale $\Theta = \theta_0(N^2H/g)$ (cf. Table 1), so that a unit potential temperature fluctuation $\theta' = 1$ corresponds to a downward displacement of an isentrope by one height scale H .

stream of the mountain, so that the interior potential vorticity P , the surface potential vorticity S , and the generalized potential vorticity $P_g = P + S$ are zero there. The flow impinges along the x axis upon a radially symmetric Gaussian mountain at the coordinate origin. The mountain is of height $h_M = 1.25H$, where $H = U/N$ is a height scale, and its slopes are gentle in that the horizontal scale $L = 10H$ of the mountain is considerably greater than the mountain height h_M . The only dissipative processes in the simulation are viscous momentum dissipation and thermal diffusion with constant viscosity ν_e and constant thermal diffusivity $\kappa_e = \nu_e$. The Reynolds number $\text{Re} = UL/\nu_e = 500$ is chosen such that the simulated flow is only weakly dissipative, yet remains laminar. The boundary condition at the mountain is a free-slip condition, namely, the vanishing of the tangential stress $-\nu_e(\mathbf{n} \cdot \nabla)\mathbf{u}$ and of the normal heat flux $-\kappa_e(\mathbf{n} \cdot \nabla)\theta'$. [See Rotunno et al. (1999, section 4) for a detailed description of the simulation.]

In the simulation, the formation of a wake with non-zero interior potential vorticity is the result of five processes: 1) modification of the thermal stratification in the vicinity of the mountain by gravity waves, 2) baroclinic induction of a surface potential vorticity dipole at the leeward slope of the mountain, 3) downslope advection of the surface potential vorticity dipole and accumulation of surface potential vorticity in a region of large gradients in surface potential temperature, 4) dissipative conversion of the surface potential vorticity dipole into an interior potential vorticity dipole, and 5) separation of the interior potential vorticity dipole from the surface and advection into the wake along isentropes that intersect the surface.

1) The way in which gravity waves modify the thermal stratification in the vicinity of the mountain above a boundary layer is qualitatively well described by adiabatic and inviscid linear theory. Figure 2 shows isentropes in the $y = 0$ symmetry plane of the flow at time $t = 14.4$, a time at which a nearly steady

TABLE 1. Scales used for nondimensionalization of quantities in Figs. 1 and 2. The fundamental scales used for the nondimensionalization are the horizontal scale L of the mountain and the velocity U and Brunt-Väisälä frequency N of the flow far upstream of the mountain.

Quantity	Scale
t	L/U
x, y, z	$H = U/N$
θ'	$\Theta = \theta_0(N^2H/g)$
P, S	$N\Theta/(\rho_0L)$
\mathbf{J}, \mathbf{K}	$UN\Theta/(\rho_0L)$
$\nabla \cdot (\rho_0\mathbf{K})$	$UN\Theta/L^2$

state has been reached in the simulation (see Table 1 for the scales used for nondimensionalization of the time and other quantities). As Fig. 2 shows, the gravity waves in the simulation are not overly steep and do not break, whence adiabatic and inviscid linear theory can account qualitatively for the modification of the thermal stratification above a boundary layer. For example, for a mountain that is sufficiently high ($h_M \geq H$), such as the mountain in the simulation considered here ($h_M = 1.25H$), adiabatic and inviscid linear theory predicts that, in the lee of the mountain, isentropes are deflected downward and collapse onto the mountain surface (Smith 1980, 1988). The downward deflection of isentropes in the lee of the mountain is evident in Fig. 2. However, adiabatic and inviscid flows generally do not satisfy the free-slip conditions that the tangential stress and the normal heat flux vanish. The prediction of linear theory that isentropes collapse onto the leeward slope of the mountain (which linear theory constrains to be an isentrope except at stagnation points) must be modified in a boundary layer. In the boundary layer, dissipative processes can become important even at high Reynolds numbers because the inhibition of motion normal to the surface allows the buildup of large tangential gradients in vorticity and potential temperature. Figure 2 shows that at the leeward slope of the mountain, where linear theory predicts a region of collapsed isentropes, a region of high and relatively uniform surface potential temperature forms, bounded downslope (and what cannot be inferred from the figure: also laterally) by a region of large gradients in surface potential temperature. Figures 1a–c each show one isentrope ($\theta' = 2.0$) that does not intersect the mountain surface and one isentrope ($\theta' = 0.8$) that intersects the leeward slope of the mountain in the region of large gradients in surface potential temperature. The essential characteristics of the thermal stratification in the vicinity of the mountain are established at time $t = 1.8$ (Fig. 1a) and evolve only slightly during the further spinup of the mountain wake (Fig. 1b, c).

2) The thermal stratification in the vicinity of the mountain implies that, in the region of high and relatively uniform surface potential temperature at the leeward

slope, a surface potential vorticity dipole is induced baroclinically. Since the baroclinic component (22) of the surface potential vorticity flux is quadratic in potential temperature fluctuations θ' , the baroclinic induction of surface potential vorticity is not taken into account in linear theories [cf. the analysis of Smolarkiewicz and Rotunno (1989), which shows that the baroclinic induction of wake vorticity is an effect of second order in perturbation amplitude]. The across-stream component

$$K_{bc}^y = \frac{\mu}{\rho_0} \frac{g\theta'^2}{2\theta_0} (\partial_x z_s) \delta(z - z_s)$$

dominates the baroclinic component of the surface potential vorticity flux at the leeward slope and transfers surface potential vorticity from the left (facing downstream) of the $y = 0$ symmetry plane to the right (since $K_{bc}^y < 0$ at the leeward slope). Figure 1d shows, at time $t = 14.4$ and projected onto the (x, y) plane, the convergence $-\nabla \cdot (\rho_0 \mathbf{K}_{bc})$ of the baroclinic component of the surface potential vorticity flux. Since the baroclinic component \mathbf{K}_{bc} depends only on the topography and on the surface potential temperature field θ' , which is already close to its steady state at time $t = 1.8$, the structure of the convergence $-\nabla \cdot (\rho_0 \mathbf{K}_{bc})$ at the earlier times $t = 1.8$ and 7.2 is similar to the shown convergence at time $t = 14.4$. The convergence of the diabatic and frictional components of the generalized potential vorticity flux is, in the region of the extrema of the baroclinic component, an order of magnitude smaller than the convergence of the baroclinic component.¹¹ The convergence of the baroclinic component of the surface potential vorticity flux at the leeward slope of the mountain leads to a surface potential vorticity dipole with negative surface potential vorticity to the left of the $y = 0$ symmetry plane and positive surface potential vorticity to the right.

- 3) The baroclinically induced surface potential vorticity dipole is advected downslope into the region of large

gradients in surface potential temperature near the leeward foot of the mountain. Since a flow with weak thermal diffusion as the only diabatic process effectively cannot cross isentropes at the surface, the advective surface potential vorticity flux converges and surface potential vorticity accumulates in the region of large surface potential temperature gradients. In Fig. 1, the convergence of the advective surface potential vorticity flux can be inferred from the vectors along the surface, which indicate the magnitude and direction of the advective surface potential vorticity flux. The accumulation of surface potential vorticity in the region of large surface potential temperature gradients near the leeward foot of the mountain is clearly recognizable in the succession of Figs. 1a–c.

- 4) As surface potential vorticity accumulates in the region of large surface potential temperature gradients near the leeward foot of the mountain, the magnitude of the conversion rate (19) between surface potential vorticity and interior potential vorticity increases on both sides of the $y = 0$ symmetry plane. In a steady state, a balance is established between the convergence of surface potential vorticity flux and the conversion of surface potential vorticity into interior potential vorticity. This balance is a consequence of the fact that, in a steady state, the divergence $\nabla \cdot (\rho_0 \mathbf{J}_g) = \nabla \cdot (\rho_0 \mathbf{J} + \rho_0 \mathbf{K})$ of the generalized potential vorticity flux vanishes, whence, by integration over a volume V that encloses a surface patch of area A and that has infinitesimally small faces normal to the surface, it follows that

$$-\int_V \nabla \cdot (\rho_0 \mathbf{K}) d\mathbf{x} = \int_A \mathbf{n} \cdot (\rho_0 \mathbf{J}) dA. \quad (31)$$

In the region of large surface potential temperature gradients near the leeward foot of the mountain, the baroclinic component $\mathbf{K}_{bc} \propto \theta'^2(\mathbf{n} \times \mathbf{k})$ is negligible in the convergence of the surface potential vorticity flux (cf. Fig. 1d) because the normal vector \mathbf{n} is nearly parallel to the vertical unit vector \mathbf{k} . The balance (31) between the convergence of surface potential vorticity flux and the conversion of surface potential vorticity into interior potential vorticity can therefore be approximated as

$$\begin{aligned} & -\int_V \nabla \cdot (\rho_0 \mathbf{u}S) d\mathbf{x} \\ & \approx -\int_A [Q\zeta_\sigma + \theta' \mathbf{n} \cdot (\nabla \times \mathbf{F})] dA, \end{aligned}$$

where we have combined the frictional components \mathbf{K}_F and \mathbf{J}_F of the surface potential vorticity flux and of the interior potential vorticity flux in the surface integral on the right-hand side (cf. footnotes 5 and 7). Since the region of large surface potential temperature gradients near the leeward foot of the moun-

¹¹ Scale analysis gives an indication of the relative magnitudes of the baroclinic and dissipative contributions to the convergence of the generalized potential vorticity flux. Integrating the convergence $-\nabla \cdot (\rho_0 \mathbf{J}_g)$ of the generalized potential vorticity flux over a volume V that encloses a surface patch of area A and that has infinitesimally small faces normal to the surface, one finds that, at about half-height at the leeward slope, the baroclinic contribution $-\int_V \nabla \cdot (\rho_0 \mathbf{K}_{bc}) d\mathbf{x}$ to the integral scales like $(g/\theta_0)(\Theta^2 h_M/L^2)A$, whereas the dissipative contributions $-\int_V \nabla \cdot (\rho_0 \mathbf{K}_F) d\mathbf{x} - \int_A \mathbf{n} \cdot (\rho_0 \mathbf{J}) dA = \int_A [Q\zeta_\sigma + \theta' \mathbf{n} \cdot (\nabla \times \mathbf{F})] dA$ scale like $(\nu_e \Theta U/H^2 L)A$ (cf. the scales in Table 1). The ratio of the baroclinic contribution to the dissipative contributions is therefore $N^2 H^3 h_M / (\nu_e UL) = H h_M / L^2 \cdot \text{Re} \approx 6$ (since $\text{Re} = 500$ and $L = 8h_M = 10H$ in the simulation considered here). This scaling estimate underestimates the actual ratio of the baroclinic contribution to the dissipative contributions—largely, it appears, because the scaling estimate $\zeta_\sigma \sim U/L$ for the vorticity in the dissipative term overestimates the vorticity in the region at the leeward slope where the baroclinic contribution to the convergence of the generalized potential vorticity flux is largest.

tain is anomalously warm (cf. Fig. 2), the diffusive diabatic heating rate Q tends to be negative there, so that, since the surface potential temperature anomaly θ' is positive at the leeward slope, the diabatic conversion $-Q\zeta_\sigma$ damps surface potential vorticity anomalies $S \propto \theta'\zeta_\sigma$ by converting positive surface potential vorticity into positive interior potential vorticity and negative surface potential vorticity into negative interior potential vorticity. The magnitude of the conversion rate $-Q\zeta_\sigma$ increases with increasing magnitude of the surface potential vorticity anomaly $S \propto \theta'\zeta_\sigma$. The frictional conversion $-\theta'\mathbf{n} \cdot (\nabla \times \mathbf{F})$ has a similar effect: to the extent that extrema in the relative vorticity $\zeta_\sigma = \mathbf{n} \cdot \boldsymbol{\omega}$, are approximately concomitant with the extrema in surface potential vorticity, the frictional conversion rate $-\theta'\mathbf{n} \cdot (\nabla \times \mathbf{F}) = -\nu_e \theta'\mathbf{n} \cdot (\nabla^2 \boldsymbol{\omega})$ is positive at the surface potential vorticity maximum near the leeward foot of the mountain and negative at the minimum. Dissipative conversion of surface potential vorticity into interior potential vorticity thus counteracts the accumulation of surface potential vorticity due to the convergence of the advective surface potential vorticity flux.

- 5) Since the large along-stream gradients in surface potential temperature imply convergence of the surface flow, the region of large surface potential temperature gradients is not only a region of significant dissipation, but also a region in which streamlines break away from the surface (cf. Rotunno et al. 1999, Figs. 3 and 6). The flow separates from the surface and advects with it the interior potential vorticity dipole that originated in the dissipative conversion of the surface potential vorticity dipole (Fig. 1). Since the flow is only weakly dissipative, the interior potential vorticity dipole is advected predominantly along isentropes—along isentropes that intersect the surface near the leeward foot of the mountain. On isentropes that do not intersect the surface, such as the upper isentropes in Fig. 1, interior potential vorticity cannot be induced by conversion from surface potential vorticity, but arises by vertical diffusion of potential temperature and/or vorticity anomalies from the surface and from lower-lying isentropes that intersect the surface. The interior potential vorticity anomalies are therefore weaker and appear later on isentropes that do not intersect the surface than on isentropes that do intersect the surface (Fig. 1).

The thickness of the boundary layer in which dissipative processes are significant for converting the surface potential vorticity dipole into an interior potential vorticity dipole and for separating the interior potential vorticity dipole from the surface can be estimated from the timescale of potential vorticity advection in the wake. To the extent that in the wake downstream of the mountain, the upstream velocity scale U and the length scale L are adequate flow scales, the timescale of potential vorticity advection

is L/U . Within that time, potential temperature and vorticity anomalies that give rise to a potential vorticity anomaly can diffuse a distance $\delta_0 = \sqrt{\nu_e L/U} = \sqrt{\kappa_e L/U} = \text{Re}^{-1/2} L$ away from the surface. This distance δ_0 is an estimate of the thickness of the boundary layer at the free-slip surface; it coincides with the thickness of a boundary layer at a no-slip surface (cf. Lighthill 1963). For the simulation considered here, with $\text{Re} = 500$ and $L = 10H$, the estimate of the boundary layer thickness is $\delta_0 = H/\sqrt{5}$, which is roughly consistent with the thickness of the region where Fig. 1 suggests significant absolute values of the divergence of the interior potential vorticity flux along isentropes. [The thicknesses of the regions where Fig. 6c of Rotunno et al. (1999) and Fig. 10 of Epifanio and Durran (2002b) indicate significant viscous and thermal dissipation are likewise consistent with the scaling estimate δ_0 of the boundary layer thickness.]

In this scenario, a wake with a pair of counterrotating lee vortices forms by conversion of a baroclinically induced surface potential vorticity dipole into an interior potential vorticity dipole. As in the analysis of Rotunno et al. (1999), baroclinicity is posited as fundamental for the formation of a wake with nonzero interior potential vorticity. The generalized potential vorticity perspective emphasizes the baroclinicity near the surface of the mountain and shows that the principal role of dissipative processes is to convert the surface potential vorticity dipole that develops at the leeward slope of the mountain into an interior potential vorticity dipole. The interior potential vorticity dipole subsequently separates from the surface. The description from the perspective of generalized potential vorticity resembles the classical descriptions of the separation of a vorticity sheet induced by frictional processes at a no-slip surface (cf. Lighthill 1963), with two main differences: whereas the surface vorticity sheet at a no-slip surface is induced by friction, a surface potential vorticity sheet can be induced by baroclinicity; and whereas the surface vorticity sheet at a no-slip surface involves vorticity components tangential to the surface, a surface potential vorticity sheet at a free-slip surface involves the vorticity component normal to the surface.

6. Summary

We have presented a formulation of potential vorticity dynamics that encompasses boundary effects. For arbitrary flows, the generalization of the potential vorticity concept to a sum of the conventional interior potential vorticity and a singular surface potential vorticity allows one to replace the inhomogeneous boundary conditions for potential vorticity dynamics by simpler homogeneous boundary conditions. For the generalized potential vorticity, a conservation law holds that is similar to the well-known conservation law for the interior po-

tential vorticity. The generalized potential vorticity reduces in the quasigeostrophic limit to Bretherton's (1966) generalized quasigeostrophic potential vorticity, which includes a surface potential vorticity that is proportional to surface potential temperature fluctuations. Not being limited to quasigeostrophic flows, however, the generalized potential vorticity concept can be used to describe flows such as mesoscale or planetary-scale flows, for which the quasigeostrophic approximation is inadequate.

The formal framework of generalized potential vorticity dynamics issues from field equations in which the potential vorticity and the potential vorticity flux appear as sources of flow quantities in the same way in which an electric charge and an electric current appear as sources of fields in electrodynamics. The boundary sources—the surface potential vorticity and the surface potential vorticity flux—that must be included in the field equations if the usual inhomogeneous boundary conditions for potential vorticity dynamics are replaced by simpler homogeneous boundary conditions were determined with techniques from electrodynamics. We derived functional forms of the surface potential vorticity and of its flux, pointed out ambiguities in these functional forms, and discussed the conservation law for generalized potential vorticity.

In an example, we demonstrated how the generalized potential vorticity and its conservation law can be used to analyze the dynamical role of boundaries in flows for which the quasigeostrophic approximation is inadequate. We outlined a theory of how a wake with lee vortices can form in mesoscale flows past a mountain with a free-slip surface. Even in adiabatic and frictionless flows, generalized potential vorticity is not, in general, materially conserved but can be induced by baroclinicity at a boundary. In stratified flows past a mountain, generalized potential vorticity can be induced by baroclinicity at the leeward slope of the mountain. As illustrated in a simulation of a stratified Boussinesq flow, weak dissipative processes in a boundary layer suffice to convert a baroclinically induced surface potential vorticity dipole that develops at the leeward slope of a mountain into an interior potential vorticity dipole, which separates from the surface and is advected into the wake along isentropes that intersect the surface. Thus a wake with a pair of counterrotating lee vortices can form by separation of a baroclinically induced generalized potential vorticity sheet from the surface of a mountain, even when frictional processes in the boundary layer are so weak that no friction-induced vorticity can be transferred into the wake in the interior of the flow.

This example illustrates how the generalized potential vorticity concept extends the conventional potential vorticity concept to encompass boundary effects. The generalized potential vorticity concept is applicable to arbitrary flows, even to flows that are not balanced. For balanced flows, the inversion principle known from con-

ventional potential vorticity dynamics carries over to generalized potential vorticity dynamics: like the conventional potential vorticity combined with typically inhomogeneous boundary conditions, the generalized potential vorticity combined with simpler homogeneous boundary conditions contains all relevant information about flows that satisfy fairly general balance conditions (cf. Hoskins et al. 1985; McIntyre and Norton 2000). Thus, as Bretherton's extension of the quasigeostrophic potential vorticity has proven fruitful in the analysis of boundary effects in quasigeostrophic flows, the generalized potential vorticity concept could prove fruitful in the analysis of boundary effects in more general balanced flows. And irrespective of whether a flow is balanced and the generalized potential vorticity contains all relevant information about the flow, the conservation law for generalized potential vorticity describes how potential vorticity fluctuations in the interior of the flow interact with boundaries and how the interior and surface potential vorticities evolve in time.

Acknowledgments. We thank Richard Rotunno, Vanda Grubišić, and Piotr Smolarkiewicz for their willingness to share their simulation data; Piotr Smolarkiewicz for making the data available in a format convenient for us; Robert Hallberg, Peter Haynes, David Marshall, Olivier Pauluis, and Richard Rotunno for helpful comments on drafts of this paper; and Heidi Swanson for editing the manuscript. Parts of this paper were drafted while T. Schneider was with the Atmospheric and Oceanic Sciences Program at Princeton University, where he was supported by a NASA Earth System Science Fellowship.

APPENDIX

Notation and Symbols

Most symbols follow standard meteorological and mathematical conventions. Listed here are only those symbols that are used repeatedly in different sections of this paper.

∂_ξ	Partial derivative with respect to space or time coordinate ξ
D/Dt	Material derivative $D/Dt = \partial_t + \mathbf{u} \cdot \nabla$ following the three-dimensional flow \mathbf{u}
c_p	Specific heat at constant pressure
E	Exner function $E = c_p(p/p_0)^\kappa$
f	Coriolis parameter $f = 2\ \boldsymbol{\Omega}\ \sin\phi$, where ϕ is latitude
\mathbf{F}	Frictional force per unit mass
g	Gravitational acceleration
h	Scale factor $h = (\partial_z\theta)^{-1}$ of isentropic coordinates
$\mathcal{H}(\cdot)$	Heaviside step function
$\mathbf{i}, \mathbf{j}, \mathbf{k}$	Local Cartesian unit vectors (eastward, northward, upward)

\mathbf{J}, \mathbf{J}_g	Potential vorticity flux, generalized potential vorticity flux
\mathbf{K}	Surface potential vorticity flux
\mathbf{n}	Unit normal vector at the surface (directed upward)
p, p_0	Pressure, constant reference pressure
P, P_g	Potential vorticity, generalized potential vorticity
Q	Diabatic heating rate $Q = D\theta/Dt$
\mathbf{r}	Radius vector $\mathbf{r} = \mathbf{r}(\mathbf{x})$ from center of the planet to point \mathbf{x} in the atmosphere
R	Gas constant
S	Surface potential vorticity
t	Time
T	Temperature
u, v, w	Velocity components (eastward, northward, upward)
\mathbf{u}	Three-dimensional velocity [$\mathbf{u} = (u, v, w)$ in local Cartesian coordinates]
\mathbf{u}_a	Three-dimensional absolute velocity $\mathbf{u}_a = \mathbf{u} + \boldsymbol{\Omega} \times \mathbf{r}$
\mathbf{v}	Horizontal velocity [$\mathbf{v} = (u, v, 0)$ in local Cartesian coordinates]
x, y, z	Local Cartesian coordinates (eastward, northward, upward)
\mathbf{x}	Three-dimensional position vector [$\mathbf{x} = (x, y, z)$ in local Cartesian coordinates]
$\delta(\cdot)$	Dirac delta function
$\zeta_\sigma, \zeta_\theta$	Relative vorticity of horizontal flow \mathbf{v} along surface [$\zeta_\sigma = \mathbf{n} \cdot (\nabla \times \mathbf{v})$] and along isentropes [$\zeta_\theta = \nabla\theta \cdot (\nabla \times \mathbf{v})/ \nabla\theta $]
θ, θ_0	Potential temperature $\theta = T(p_0/p)^\kappa$, constant reference temperature
θ_s^x, θ_s^y	Derivatives $\partial_{x,y}\theta_s(x, y, t)$ of surface potential temperature $\theta_s(x, y, t)$ with respect to x, y
κ	Adiabatic exponent $\kappa = R/c_p$
μ	Normalization factor $\mu = (1 + \ \nabla_{z_s}\ ^2)^{-1/2}$
Φ	Geopotential
ρ	Density
ρ_θ	Isentropic density $\rho_\theta = -(g^{-1}\partial_\theta p)\mathcal{H}(\theta - \theta_s)$ [density in (x, y, θ) space]
$\boldsymbol{\omega}_r$	Relative vorticity $\boldsymbol{\omega}_r = \nabla \times \mathbf{u}$
$\boldsymbol{\omega}_a$	Absolute vorticity $\boldsymbol{\omega}_a = \boldsymbol{\omega}_r + 2\boldsymbol{\Omega}$
$\boldsymbol{\Omega}$	Angular velocity of planetary rotation

REFERENCES

- Andrews, D. G., J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics*. International Geophysics Series, Vol. 40, Academic Press, 489 pp.
- Arfken, G., 1985: *Mathematical Methods for Physicists*. 3d ed. Academic Press, 985 pp.
- Batchelor, G. K., 1967: *An Introduction to Fluid Dynamics*. Cambridge University Press, 615 pp.
- Bretherton, C. S., and C. Schär, 1993: Flux of potential vorticity substance: A simple derivation and a uniqueness property. *J. Atmos. Sci.*, **50**, 1834–1836.
- Bretherton, F. P., 1966: Critical layer instability in baroclinic flows. *Quart. J. Roy. Meteor. Soc.*, **92**, 325–334.
- Epifanio, C. C., and D. R. Durran, 2002a: Lee-vortex formation in free-slip stratified flow over ridges. Part I: Comparison of weakly nonlinear inviscid theory and fully nonlinear viscous simulations. *J. Atmos. Sci.*, **59**, 1153–1165.
- , and —, 2002b: Lee-vortex formation in free-slip stratified flow over ridges. Part II: Mechanisms of vorticity and PV production in nonlinear viscous wakes. *J. Atmos. Sci.*, **59**, 1166–1181.
- Hallberg, R. W., and P. B. Rhines, 2000: Boundary sources of potential vorticity in geophysical circulations. *Developments in Geophysical Turbulence*, R. M. Kerr and Y. Kimura, Eds., Kluwer Academic, 51–65.
- 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.
- , and —, 1990: On the conservation and impermeability theorems for potential vorticity. *J. Atmos. Sci.*, **47**, 2021–2031.
- Held, I. M., and T. Schneider, 1999: The surface branch of the zonally averaged mass transport circulation in the troposphere. *J. Atmos. Sci.*, **56**, 1688–1697.
- 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.
- Jackson, J. D., 1975: *Classical Electrodynamics*. 2d ed. Wiley, 848 pp.
- Lighthill, M. J., 1963: Boundary layer theory. *Laminar Boundary Layers*, L. Rosenhead, Ed., Clarendon Press, 46–113.
- Marshall, J., D. Jamous, and J. Nilsson, 2001: Entry, flux, and exit of potential vorticity in ocean circulation. *J. Phys. Oceanogr.*, **31**, 777–789.
- McDougall, T. J., 1988: Neutral-surface potential vorticity. *Progress in Oceanography*, Vol. 20, Pergamon, 185–221.
- McIntyre, M. E., and W. A. Norton, 2000: Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.*, **57**, 1214–1235.
- Morse, P. M., and H. Feshbach, 1953: *Methods of Theoretical Physics*. International Series in Pure and Applied Physics, McGraw-Hill, 1978 pp.
- Rhines, P. B., 1979: Geostrophic turbulence. *Annu. Rev. Fluid Mech.*, **11**, 401–441.
- , and W. R. Young, 1982: Homogenization of potential vorticity in planetary gyres. *J. Fluid Mech.*, **122**, 347–367.
- Rotunno, R., V. Grubišić, and P. K. Smolarkiewicz, 1999: Vorticity and potential vorticity in mountain wakes. *J. Atmos. Sci.*, **56**, 2796–2810.
- Salmon, R., 1998: *Lectures on Geophysical Fluid Dynamics*. Oxford University Press, 378 pp.
- Schär, C., 1993: A generalization of Bernoulli's theorem. *J. Atmos. Sci.*, **50**, 1437–1443.
- , and D. R. Durran, 1997: Vortex formation and vortex shedding in continuously stratified flows past isolated topography. *J. Atmos. Sci.*, **54**, 534–554.
- Smith, R. B., 1980: Linear theory of stratified hydrostatic flow past an isolated mountain. *Tellus*, **32A**, 348–364.
- , 1988: Linear theory of stratified flow past an isolated mountain in isosteric coordinates. *J. Atmos. Sci.*, **45**, 3889–3896.
- , 1989: Comment on “Low Froude number flow past three-dimensional obstacles. Part I: Baroclinically generated lee vortices.” *J. Atmos. Sci.*, **46**, 3611–3613.
- 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.
- Tung, K. K., 1986: Nongeostrophic theory of zonally averaged circulation. Part I: Formulation. *J. Atmos. Sci.*, **43**, 2600–2618.