Note to editors: Crucial new terms are flagged up using the macro `\newterm{...}` in the LaTeX source, defined here to print as *italic*. Possible cross-references to other articles are flagged `\crossref{...}`. For visibility, `\crossref{...}` is temporarily defined to print as sans-serif. Therefore, *sans-serif italic* signals both a new term and a possible cross-reference. I have used `\openingscarequote...\closingscarequote` in a few places. This bends the house-style rules, but unfortunately we don't live in an ideal world with perfectly logical terminology. A standard example, that of the variable solar 'constant', is enough to make the point. Another is the slow 'manifold'. It is not a manifold.

Dynamic Meteorology MS 140

Potential vorticity[1]

Michael E. McIntyre

Emeritus Professor,
University of Cambridge,
Department of Applied Mathematics and Theoretical Physics,
Wilberforce Road,
Cambridge CB3 0WA,
UK.
www.atm.damtp.cam.ac.uk/people/mem/
m.e.mcintyre@damtp.cam.ac.uk

Synopsis: The significance of the potential vorticity (PV) for atmosphere–ocean dynamics was first explored by Carl-Gustaf Rossby in the 1930s. Reviewed here are its key properties including invertibility, material invariance, and the impermeability theorem — the last two suggesting mixability along stratification surfaces. These properties easily explain the once-mysterious anti-friction or 'negative viscosity' of strongly nonlinear atmosphere–ocean eddy fields, outside the scope of linear theory and homogeneous turbulence theory. Invertibility implies that eddy fluxes of momentum are intimately related to isentropic eddy fluxes of PV, including those due to strongly nonlinear disturbances, as summarized by the quasigeostrophic Taylor identity.

[1]Article in press for the 2nd edition of the *Encyclopedia of Atmospheric Science*, edited by Gerald North, Fuqing Zhang and John Pyle (Elsevier, 2012), finalized 24 July 2012.

Potential vorticity[1]

Michael E. McIntyre

University of Cambridge, Department of Applied Mathematics and Theoretical Physics, Wilberforce Road, Cambridge CB3 0WA, UK.
www.atm.damtp.cam.ac.uk/people/mem/

1 The fundamental definition

The idea of the potential vorticity (PV) as a material invariant central to stratified, rotating fluid dynamics was first introduced and explored by Carl-Gustaf Rossby in the 1930s. Material invariance means constancy on a fluid particle. The potential vorticity, a scalar field, will be denoted here by P and can be defined in several ways, as shown shortly. We have

$$DP/Dt \;=\; 0 \tag{1}$$

for dissipationless flow, where D/Dt is the material derivative. For such flow we also have material invariance of the potential temperature θ,

$$D\theta/Dt \;=\; 0\;. \tag{2}$$

Rossby's idea, as it originally emerged from his papers of 1936, 1938 and 1940, was to introduce a vorticity-like quantity that is related to the vertical component of vorticity in the same way that potential temperature is related to temperature. In his 1938 and 1940 papers he recognized, moreover, that 'vertical' can more accurately be replaced by 'normal to stratification surfaces', i.e., in the atmosphere, normal to isentropic or constant-θ surfaces.

Equivalent to this is the idea, clearly emerging on page 252 of the 1938 paper, that P is exactly proportional to the absolute *Kelvin circulation* C_Γ, Eq. (7) below, around an infinitesimally small closed material contour Γ lying on an isentropic surface. The exact material-invariance property (1) is then obvious from Kelvin's *circulation theorem*, as generalized by V. Bjerknes, since (2) ensures that the material contour Γ remains on the isentropic surface.

Rossby's idea is today recognized as having central and far-reaching importance for understanding the dynamical behavior not only of planetary

[1] Article in press for the 2nd edition of the *Encyclopedia of Atmospheric Science*, edited by Gerald North, Fuqing Zhang and John Pyle (Elsevier, 2012), finalized 24 July 2012.

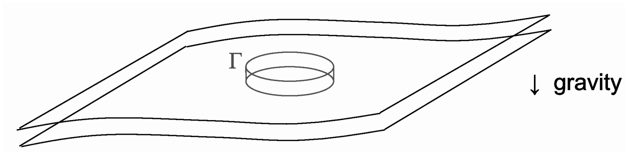

Figure 1: Sketch showing the material mass element defined by a small isentropic contour Γ and a pair of neighboring isentropic (stratification) surfaces with potential temperatures θ and $\theta + \mathrm{d}\theta$. The exact PV is the mass-normalized Kelvin circulation around Γ, in the limit of an infinitesimally small element (see text). In a layer model, the two surfaces are taken instead as the layer boundaries.

atmospheres and oceans but also of the radiative interiors of solar-type stars. It is especially important for understanding *balanced flow* and thence a vast range of basic dynamical processes, such as Rossby-wave propagation and breaking and its many consequences including, in the Earth's atmosphere, global-scale teleconnections, anti-frictional phenomena such as jet stream self-sharpening, and the genesis of cyclones, anticyclones and storm tracks, answering the child's age-old question of where the wind comes from.

The relation $P \propto C_\Gamma$ provides the simplest and most fundamental way to define P exactly, not only for continuously stratified systems but also for single-layer shallow-water or 'equivalent barotropic' models and their multi-layer extensions. For continuous stratification, today's standard definition of P chooses the constant of proportionality to be $\mathrm{d}\theta$, the potential-temperature increment between a pair of neighboring isentropic surfaces (see Fig. 1), divided by the mass of the small material fluid element lying between those surfaces and having perimeter Γ. Mass conservation is assumed throughout this article.

For the single-layer and multi-layer models one need only replace the pair of isentropic surfaces by layer boundaries. Then for finite layer thickness the proportionality constant can be chosen as simply the reciprocal of the mass of the material element, or of its volume when the usual incompressible-flow assumption is made. Then from Stokes' theorem P becomes absolute vorticity divided by layer thickness, the formula first presented in Rossby's 1936 paper.

For continuous stratification Rossby derived an approximate formula adequate for use with synoptic-scale observational data. With the foregoing choice of proportionality constant, Rossby's formula is

$$P \;\approx\; g\left\{\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)_{\theta} \;+\; f\right\}\left|\frac{\partial \theta}{\partial p}\right| \tag{3}$$

where g is the gravitational acceleration, p is pressure, and f is the Coriolis parameter, a function of latitude. To obtain (3) from the exact relation $P \propto C_\Gamma$ one must assume that the mass and pressure fields are related hydrostatically and that the slopes of isentropic surfaces are small in comparison with unity. In practice these conditions usually hold to more than sufficient accuracy. The horizontal coordinates x, y in (3) are local Cartesian coordinates in a tangent-plane representation, with corresponding horizontal velocity components u, v relative to the Earth. The formula converts to spherical or other coordinates in the same way as the ordinary vertical vorticity.

However, as Rossby pointed out, the quantity within braces is not the ordinary vertical vorticity. The subscript θ is crucial. It signifies that the horizontal differentiations of the horizontal velocity components are to be carried out with θ held constant. That is, one stays on a single isentropic surface, just as one does when calculating C_Γ. Rossby explains this point very clearly on, for instance, page 253 of his 1938 paper. The resulting quantity, bearing a superficial resemblance to the ordinary vertical vorticity, can more aptly be called Rossby's *isentropic vorticity.* Within the approximations involved in (3), this isentropic vorticity is the same as the component of the vorticity vector normal to the isentropic surface. It can differ substantially from the vertical vorticity.

Such differences are commonplace in balanced flows with strong vertical shear ($\partial u/\partial z$, $\partial v/\partial z$) where z is geometric altitude or pressure altitude. That is, they are commonplace in balanced flows with high baroclinicity. Examples include tropopause jet streams. Baroclinicity means tilting of isentropic surfaces relative to isobaric surfaces, usually the cross-stream tilting that balances the vertical shear as indicated by the so-called thermal wind equation. A natural measure of baroclinicity is $1/\mathrm{Ri}$ where $\mathrm{Ri} = N^2/(\partial|\mathbf{u}|/\partial z)^2$, the gradient Richardson number, where $N^2 = g\theta^{-1}\partial\theta/\partial z$, the square of the buoyancy frequency. The shear and cross-stream tilting effects were shown to make substantial contributions to the right-hand side of (3) in, for instance, the 1950s work of R. J. Reed, F. Sanders and E. F. Danielsen on observational data describing tropopause fronts and jet streams, in which air of stratospheric origin was recognized by its relatively high values of P. Slopes are geometrically small but Ri values low enough for the subscript θ to be important in (3).

Equations (1)–(3) provide a remarkably succinct description of how dissipationless processes affect the component of absolute vorticity normal to an isentropic surface. There are two distinct effects. The first is that the normal component of absolute vorticity increases through vortex stretching

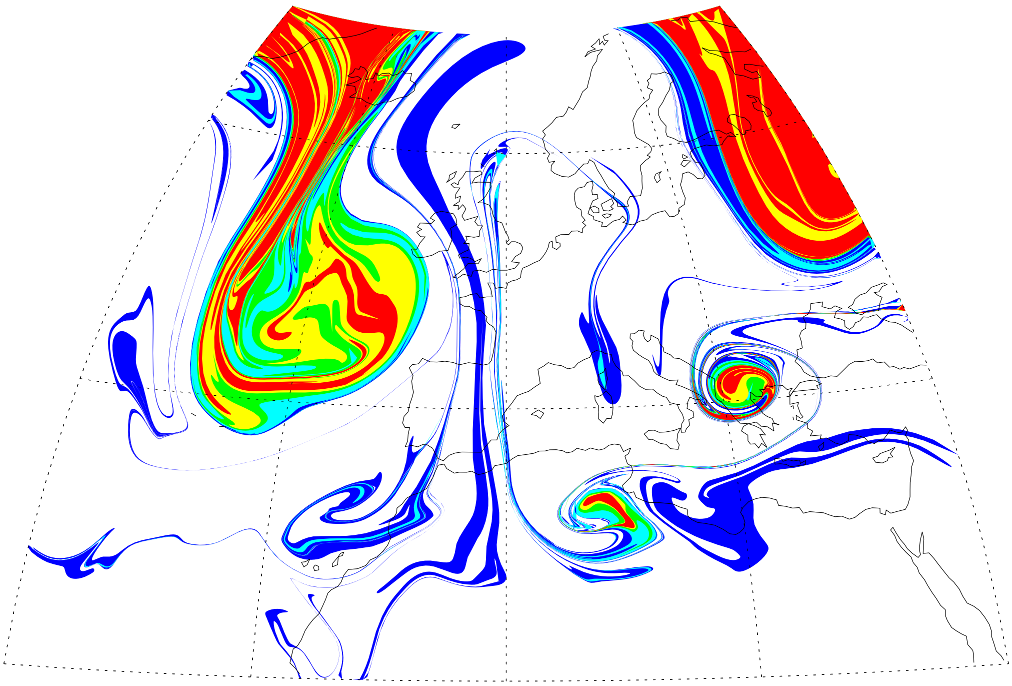

Figure 2: Estimated isentropic distribution of the (Rossby–Ertel) PV on the 320 K isentropic surface on 14 May 1992 at 1200 UT (Greenwich mean time), derived from observations as explained in the text. Over Europe the 320 K surface lies near jetliner cruising altitudes $z \sim 10\,\mathrm{km}$. The estimate used data from the operational weather-prediction analyses of the European Centre for Medium Range Weather Forecasts (ECMWF). Values from 1 PVU upwards are colored rainbow-wise from dark blue to red, with contour interval 1 PVU, where 1 PVU = $10^{-6}\mathrm{m}^2\mathrm{s}^{-1}\mathrm{K\,kg}^{-1}$. Courtesy W. A. Norton (personal communication); further details in Appenzeller et al. (1996). Figure 15b on p. 1450 of that paper checks that the wind field does, as expected from PV inversion, exhibit the usual tropopause jet structure around the periphery of the large high-PV region on the left. See *PV mixability and strong jets* below.

if the isentropic surfaces move apart. This is a generalization of angular momentum conservation, i.e., a generalization of the ballerina effect or ice-skater's spin. The second is that the normal component of absolute vorticity is preserved if the isentropic surfaces do nothing but tilt away from the horizontal.

The generalized ballerina effect often contributes to the spin-up of cyclonic vortices, such as the small vortex over the Balkans in Fig. 2. The colors mark air with different estimated values of P, on the $\theta = 320\,\mathrm{K}$ isentropic surface at geometric altitudes around 10 km, with the warmest colors marking the highest P-values. The vortex over the Balkans has a core of high-P air that has undergone stretching, while moving equatorward out of the stratosphere. The cyclonic, i.e. counterclockwise, rotation of the core relative to the surrounding air shows up as a tendency of the surrounding colored filaments to be wound up into spirals.

The estimated isentropic distribution of P shown in Fig. 2 was derived from an initial coarse-grain estimate from operational weather-forecasting analyses together with an assumption that material invariance, (1) with (2), holds to sufficient accuracy over 4 days. A highly accurate tracer advection technique, contour advection, was used. It was first introduced into the atmospheric-science literature by W. A Norton, R. A. Plumb and D. W. Waugh following work of N. J. Zabusky and D. G. Dritschel. The pattern thus revealed, reminiscent of cream on coffee, illustrates the typical advective effects of the layerwise-two-dimensional flow characteristic of mesoscale and larger-scale flow regimes heavily constrained by stable stratification. Such regimes can often be considered to be balanced flows, whose isentropic distributions of P contain nearly all the information about the dynamics. This will be made precise in the section on *PV inversion* below.

2 Ertel's formula

For continuous stratification it is a simple exercise in vector calculus to show, via Stokes' theorem, that Rossby's fundamental relation $P \propto C_\Gamma$ is exactly equivalent to

$$P \;=\; \rho^{-1}\boldsymbol{\zeta}^{\mathrm{a}}\cdot\nabla\theta \tag{4}$$

when the constant of proportionality is chosen as before. Here ρ is the mass density, ∇ is the three-dimensional gradient operator, and $\boldsymbol{\zeta}^{\mathrm{a}}$ is the absolute vorticity vector, the curl of the three-dimensional velocity field viewed in an inertial frame. In the Earth's rotating frame, $\boldsymbol{\zeta}^{\mathrm{a}}$ is the three-dimensional relative vorticity added vectorially to twice the Earth's angular velocity vector $\mathbf{\Omega}$. The formula (4) was first published in 1942 by Hans Ertel, who had visited Rossby at MIT in 1937. The formula has attracted much attention in the mathematical fluid-dynamics community and has been generalized in various ways.

In strongly stratified flows like that of Fig. 2 we have $N^2 \gg 4|\mathbf{\Omega}|^2$. Also, the small-slope approximation is valid, making $\nabla\theta$ nearly vertical. In (4), the scalar multiplication by $\nabla\theta$ picks out f, the latitude-dependent vertical component of $2\mathbf{\Omega}$, to good approximation. This is the fundamental reason why f and its latitudinal variation often suffice to capture the main effects of the Earth's rotation $\mathbf{\Omega}$, including the so-called beta effect.

Under the small-slope and hydrostatic approximations, $\rho^{-1}|\nabla\theta|$ is approximately equal to $g|\partial\theta/\partial p|$ in (3). The contributions to (3) and (4) from $2\mathbf{\Omega}$ therefore agree. It is straightforward to show that the remaining contributions also agree in these circumstances provided that, for consistency with the hydrostatic approximation, the vertical component of velocity is neglected when taking the curl of the relative velocity field to form the relative vorticity.

The small-slope and hydrostatic approximations are usually so good that (3) and (4) give practically indistinguishable results when evaluated from typical meteorological datasets, and from the output of numerical weather forecasting models. So (3) and (4) are often treated as equivalent for practical purposes, both being called 'exact' when distinguishing them from the much less accurate formulae for the material invariants possessed by certain approximate balanced models, such as quasigeostrophic theory and semi-geostrophic theory. Their material invariants are also called potential vorticities but are defined by formulae that differ substantially from (3) and (4), for instance (15) below. Unlike (3) and (4) these formulae cannot be considered quantitatively accurate. The potential vorticity in its quantitatively accurate sense will be referred to as the *Rossby–Ertel potential vorticity* or simply, for brevity, the PV, whether defined by (3) or (4) or by any other formula accurately equivalent to $P \propto C_\Gamma$.

To check that (4) is accurately, indeed exactly, equivalent to $P \propto C_\Gamma$ and materially invariant for dissipationless flow, we note first that (4) can be rewritten exactly as

$$P \;=\; \sigma^{-1}\boldsymbol{\zeta}^{\mathrm{a}}\cdot\mathbf{n} \qquad (5)$$

where $\sigma = \rho/|\nabla\theta|$, and $\mathbf{n} = \nabla\theta/|\nabla\theta|$, the upward-directed unit normal to the isentropic surface $\mathcal{S}$, say, on which P is being evaluated. The scalar field

σ, a stratification-related mass density, is a strictly positive quantity. Under the small-slope approximation it is the mass density in isentropic coordinates. With the definition just given, $\sigma\,\mathrm{d}\theta$ is exactly the mass per unit area between neighboring isentropic surfaces, such as those sketched in Fig. 1, whose θ values differ by $\mathrm{d}\theta$. Thus if $\mathrm{d}A$ is the area element of integration on the surface $\mathcal{S}$, then $\sigma\,\mathrm{d}A\,\mathrm{d}\theta$ is exactly the mass element of integration.

For dissipationless flow we have (2) as well as mass conservation, hence

$$\iint_{\mathcal{S}(\Gamma)} \sigma\,\mathrm{d}A \;=\; \text{constant} \tag{6}$$

where $\mathcal{S}(\Gamma)$ denotes any simply-connected portion of $\mathcal{S}$ enclosed by a material contour Γ. Here Γ can, but need not, be small. By definition its Kelvin circulation is

$$C_\Gamma \;=\; \oint_\Gamma \mathbf{u}^{\mathrm{a}}\cdot\mathrm{d}\mathbf{x} \;=\; \text{constant} \tag{7}$$

for dissipationless flow, where $\mathbf{u}^{\mathrm{a}}$ is the three-dimensional velocity field in the inertial frame. From Stokes' theorem and (5) we have exactly

$$C_\Gamma \;=\; \iint_{\mathcal{S}(\Gamma)} \boldsymbol{\zeta}^{\mathrm{a}}\cdot\mathbf{n}\,\mathrm{d}A \;=\; \iint_{\mathcal{S}(\Gamma)} P\,\sigma\,\mathrm{d}A \tag{8}$$

and if, as before, we now take Γ to be small — more precisely, if we take the greatest diameter of Γ to be arbitrarily small in comparison with all lengthscales of the flow — then P is simply (8) divided by (6). This verifies not only the material invariance of P but also the equivalence of (4) and (5) to $P \propto C_\Gamma$ for small Γ, with the choice of proportionality constant made earlier.

For completeness we sketch the alternative derivation given by Ertel, written using the three-dimensional velocity field $\mathbf{u}$ relative to the rotating frame. One takes the scalar product of $\nabla\theta$ with the frictionless three-dimensional vorticity equation, the curl of the nonhydrostatic equation for $D\mathbf{u}/Dt$, and then makes use of $\nabla(D\theta/Dt) = 0$, from (2). Note that $D/Dt = \partial/\partial t + \mathbf{u}\cdot\nabla$ and that the three-dimensional gradient operator ∇ acts on $\mathbf{u}$ as well as on θ. The baroclinic term in the vorticity equation, proportional to $\nabla p \times \nabla\rho$, is annihilated when the scalar product with $\nabla\theta$ is taken, because the thermodynamics says that θ is a function of p and ρ alone (the standard approximation to this function implying that $\theta \propto T/p^{\kappa}$, $\kappa \approx 2/7 \approx 0.286$, with temperature $T \propto p/\rho$). The result is a conservation relation in the general sense of the term, in 'flux form',

$$\frac{\partial}{\partial t}(\rho P) \;+\; \nabla\cdot(\rho\mathbf{u}P) \;=\; 0 \tag{9}$$

with P defined by (4) or (5). Putting this together with the corresponding equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \tag{10}$$

expressing mass conservation, we immediately obtain Eq. (1) for dissipationless flow.

A corollary of material invariance and mass conservation is the existence of so-called *Casimir invariants*. They are important in theories that make explicit the Hamiltonian mathematical structure of the dissipationless dynamics, and in associated theorems on instability and on wave–mean interaction. Note first that we have not only constancy of (8) but also

$$\iint_{\mathcal{S}(\Gamma)} \varphi_1(P)\, \sigma \,\mathrm{d}A = \text{constant} \tag{11}$$

where $\varphi_1(P)$ is an arbitrary function and Γ is again arbitrary. This is because each mass element has a single value of P and therefore a single value of $\varphi_1(P)$. Extending $\mathcal{S}(\Gamma)$ to span the whole fluid domain and integrating over all surfaces $\mathcal{S}$, with arbitrary θ-weighting, we obtain

$$\iiint \varphi_2(P, \theta)\, \sigma \,\mathrm{d}A\, \mathrm{d}\theta = \text{constant} \tag{12}$$

with $\varphi_2(P, \theta)$ another arbitrary function, where the integral is taken over the whole fluid domain. These domain integrals (12) are the Casimir invariants. They are exactly constant for any dissipationless flow whatever.

3 PV units and the extratropical tropopause

Rossby's original choice of proportionality constant differed from today's standard choice. As noted in his 1940 paper, Rossby chose the physical dimensions of P to be the same as those of ordinary vorticity, namely $(\text{time})^{-1}$, drawing on the analogy with potential temperature. (See text between his Eqs. (11) and (13).) However, the usual practice today is to tolerate the slightly looser analogy and different physical units implied by (3)–(5), for the sake of having simpler formulae. The standard PV unit used today is $10^{-6}\mathrm{m}^2\mathrm{s}^{-1}\mathrm{K}\,\mathrm{kg}^{-1}$, abbreviated PVU.

By a strange accident, cross-sections of the atmosphere show P values typically around 2 PVU at the extratropical tropopause, and this has proved extremely useful as a way of defining the tropopause outside a tropical band of latitudes, say outside $\pm 20^\circ$ or so. More precisely, the extratropical tropopause is often marked by steep isentropic gradients of P with values ranging from about 1 to 4 PVU. The shape of the 2-PVU contour in Fig. 2, dividing dark blue from light blue, gives no more than a slight hint of the complicated three-dimensional shape of the tropopause, where it intersects the 320 K isentropic surface at the instant shown. The instantaneous tropopause is a highly convoluted surface with an overall poleward-downward slope, so that the white areas in Fig. 2 are in the troposphere and the main colored areas are in the stratosphere.

Airborne measuring instruments flown along the 320 K surface and crossing from white through dark blue into light blue and warmer-colored areas would see changes in chemical composition characteristic of the transition from tropospheric to stratospheric air. Indeed, such changes have often been observed in association with finer-scale, filamentary structures of the kind seen in the figure, beginning with the pioneering work of D. W. Waugh and R. A. Plumb in the early 1990s using chemical data from NASA's ER-2 aircraft.

The usefulness of the PV as an extratropical tropopause marker is an accident because, for one thing, it depends on the choice of θ as the thermodynamical material invariant that satisfies (2) and appears in the definitions (3)–(5). There is no fundamental reason for that choice. Everything in the dynamical theory works just as well with other thermodynamical material invariants such as the specific entropy, or indeed any other smooth, monotonic function of θ. The PV thus redefined is sometimes called a *modified PV*. Isentropic distributions of P like that in Fig. 2 remain the same after such modification, apart from changes to the units and to the numerical values assigned to each color. Notice, however, that the normalizing factors for those changes depend on θ and are therefore different on each isentropic surface.

4 PV inversion and generalized PV

Any flow that can be considered balanced whether geostrophically or at higher accuracy (see Dynamic Meteorology: Balanced Flow) satisfies what is now called the *invertibility principle* for PV. The principle says that, to an accuracy limited only by the accuracy of the balance relation, one can capture all the dynamical information about the flow by specifying only the following:

1. the mass under each isentropic surface $\mathcal{S}$,
2. the isentropic distributions of P, on all the surfaces $\mathcal{S}$, and

3. the distributions of θ on the lower boundary and on the upper boundary if present.

By implication there exists, then, a nonlocal diagnostic operator, the *PV inversion operator* associated with the given balance relation. Its input is the foregoing information at some instant. Its output is the remaining dynamical information at the same instant including the p, ρ, T, and $\mathbf{u}$ fields. Very often $\mathbf{u}$ is dominated by its horizontal component, the weaker vertical component nevertheless being dynamically significant thanks to its role in the generalized ballerina effect, and in moving and tilting isentropic surfaces.

The idea of PV inversion is implicit in textbook descriptions of, for instance, the Rossby-wave mechanism. The idea is used at the point in the argument where the horizontal component of $\mathbf{u}$ is deduced diagnostically from the disturbance PV field associated with PV-contour undulations. Sometimes the term *induced velocity*, borrowed from aerodynamics, is used. In this context it means the velocity field deduced from the PV field by inversion.

What are PV inversion operators like, qualitatively? A partial answer is that calculating the horizontal component of $\mathbf{u}$ is like calculating the electric field $\mathbf{E}$ induced by a certain electric charge distribution, and then taking the horizontal component of $\mathbf{E}$ and rotating it counterclockwise through a right angle, for instance from northward to westward. The electric charges correspond to isentropic anomalies in P and boundary anomalies in θ. Thus, for instance, the positive isentropic anomaly in P over the Balkans in Fig. 2 corresponds to a positive electric charge, inducing an outward-pointing $\mathbf{E}$ field and hence a cyclonic or counterclockwise velocity field around it. This provides us with a way of saying what the terms *vortex*, *cyclone*, and *anticyclone* really mean. For instance the vortex over the Balkans, an upper-air cyclone, is nothing but a positive isentropic anomaly in P together with its induced velocity field.

Because of the balance relation, these velocity fields are accompanied by p, ρ, and T fields that to a first approximation satisfy the thermal wind equation; for instance the upper-air cyclone has a warm T anomaly above it and a cold T anomaly beneath. Conversely, an upper-air anticyclone has a cold T anomaly above, a fact crucial to lower-stratospheric polar ozone chemistry. Flow through such a cold anomaly cannot advect the negative PV anomaly beneath, but can give rise to fast cloud formation and accelerated chemical processing.

Similar statements about vortices apply to the distributions of θ at, say, the lower boundary surface. (In practical terms, taking friction into account, this translates to 'just above the planetary boundary layer'.) A surface cyclone or *heat low* is nothing but a positive, i.e. warm, lower-boundary anomaly in θ together with its induced velocity field, and conversely for a surface anticyclone.

Severe cyclonic storms in the extratropical atmosphere often arise from the vertical alignment of warm lower-boundary anomalies in θ and positive upper-air isentropic anomalies in P like the large cyclonic anomaly seen on the left of Fig. 2. Helped by such vertical alignment, the induced velocities can add up to give storm-force winds. Furthermore, the development of such a situation by upper-air positive-P advection along with near-surface warm advection, and poleward upgliding along sloping isentropes, induces large-scale upward motion. Such upward motion is described by any sufficiently accurate PV inversion operator. Alternatively, it can be computed via the so-called omega equation. The large-scale upward motion may trigger latent heat release, creating or intensifying isentropic anomalies in P. Especially in moist air over the extratropical oceans, the upshot can be the sudden explosive marine cyclogenesis feared and respected by sailors: "Three days from land a great tempest arose..."

It hardly needs saying that, whenever the invertibility principle holds to sufficient accuracy, it gives us a vastly simplified conceptual view of the dynamical evolution. The dynamical system is completely specified by a PV inversion operator together with the remarkably simple prognostic equations (1) and (2) or their diabatic, frictional generalizations. Those equations provide us with the simplest way to cope with the bedrock mathematical difficulty of fluid dynamics, the advective nonlinearity.

Since P and θ are scalar fields, keeping track of them using pictures like Fig. 2, actual or mental, is a far simpler task than keeping track of the evolving p, ρ, T, and $\mathbf{u}$ fields in three dimensions, including the nonlocal effects mediated by the p field under the constraints imposed by the balance relation. The nonlocal effects are all encapsulated in the PV inversion operator. The foregoing points, implicit in Rossby's work, were articulated with increasing clarity by Jule G. Charney and Aleksandr M. Obukhov in the late 1940s and by Ernst Kleinschmidt in the early 1950s. They allow us to make sense not only of Rossby-wave propagation, cyclogenesis, and anticyclogenesis but also, for instance, of aerodynamical ideas like vortex rollup — the idea that a strong isentropic anomaly in PV can roll 'itself' up into a nearly circular vortex, as in the Balkans example of Fig. 2.

In 1966 Francis P. Bretherton pointed out that an even greater conceptual simplification is possible. The single prognostic equation (1) is enough to determine the dissipationless evolution by itself, provided that we consider the PV field $P(\mathbf{x}, t)$ to contain delta-function contributions at the upper and lower boundaries, with strengths determined by the θ distributions at the boundaries. Ignoring frictional boundary-layer phenomena, we may relate this to the idea that isentropic surfaces $\mathcal{S}$ intersecting the lower boundary, say, can be imagined to continue along the boundary in an infinitesimally thin layer of infinite $|\nabla\theta|$ hence infinite P. In the electrostatic analogy, surface θ distributions correspond to surface charge distributions — electric charge per unit area rather than per unit volume. The PV field with surface θ distributions included may be called the *generalized PV field*, containing all the information in the second and third numbered items above.

5 Some illustrations

The idea of PV inversion can be illustrated in a simple way by considering the theoretical limiting case of infinite sound speed and infinite stable stratification. The buoyancy frequency N and gradient Richardson number both tend to infinity. The isentropic surfaces $\mathcal{S}$ become rigid and horizontal — horizontal in the billiard-table sense, with the sum of the gravitational and centrifugal potentials constant. The balance relation degenerates to a statement that the flow on each $\mathcal{S}$ is strictly horizontal and strictly incompressible. Then, in the rotating frame, we have $\mathbf{u} = \hat{\mathbf{z}} \times \nabla_H \psi$ for some streamfunction ψ, where $\hat{\mathbf{z}}$ is a unit vertical vector, and, from (5),

$$P \;=\; \sigma^{-1}(f + \nabla_H^2 \psi) \qquad (13)$$

with σ now strictly constant. Here ∇_H is the two-dimensional horizontal gradient operator and ∇_H^2 the corresponding Laplacian, so that $\nabla_H^2\psi$ is the relative vorticity. We may regard (13) as a Poisson equation to be solved for ψ when P is given. Solving it is a well defined, and well behaved, operation, given suitable boundary conditions such that the P field on each $\mathcal{S}$ satisfies (8) with Γ taken as the horizontal domain boundary; see also (16) below. Symbolically, in the rotating frame,

$$\mathbf{u} \;=\; \hat{\mathbf{z}} \times \nabla_H \psi \quad \text{with} \quad \psi \;=\; \nabla_H^{-2}(\sigma P - f) \;, \qquad (14)$$

expressing PV invertibility in the limiting case. The PV inversion problem now resembles an electrostatics problem in two, rather than three, dimensions. The charge distribution corresponds to σ times the PV anomaly

$(P - \sigma^{-1} f)$, with $-\psi$ in the role of the electric potential. In this limiting case, as in general, PV inversion is a diagnostic, nonlocal operation.

Notice that our limiting case is degenerate in another sense as well. The altitude z now enters the problem only as a parameter. There is no derivative $\partial/\partial z$ anywhere in the problem, either in the horizontal Laplacian or in the material derivative $D/Dt = \partial/\partial t + \mathbf{u} \cdot \nabla$ in (1), with $\mathbf{u}$ strictly horizontal. Not only is the flow layerwise-two-dimensional, but the layers are completely decoupled from each other. For the validity of this picture there is, therefore, an implicit restriction on magnitudes of $\partial/\partial z$, i.e. an implicit restriction on the smallness of vertical scales in the limit, with the further implication that the picture cannot be uniformly valid for all time.

More realistically, when N and Ri are large but finite, and when f is finite, $\partial/\partial z$ reappears in the problem and brings back vertical coupling. The flow remains layerwise-two-dimensional in the sense that notional 'PV particles' move along each isentropic surface $\mathcal{S}$ — see *impermeability theorem* below — but the surfaces $\mathcal{S}$ themselves are no longer quite horizontal, nor quite rigid. Aside from the vertical advection that moves and tilts the surfaces $\mathcal{S}$, all the vertical coupling comes from the PV inversion operator. The two-dimensional inverse Laplacian in (14) is replaced by an inverse elliptic operator that qualitatively resembles a three-dimensional inverse Laplacian when a stretched vertical coordinate Nz/f is used; thus the vertical coupling for flows of horizontal scale L is effective over a height scale of the order of the corresponding Rossby deformation height fL/N.

For finite N and Ri there are tradeoffs between accuracy and simplicity. The mathematically simplest though least accurate three-dimensional PV inversion operator is that arising in the standard Charney–Obukhov quasigeostrophic theory, an asymptotic theory whose approximations are valid away from the equator, for large Ri and small Rossby number $\mathrm{Ro} \sim \mathrm{Ri}^{-1/2}$, where Ro can be defined as f_0^{-1} times a typical relative-vorticity value with f_0 a constant representative value of the Coriolis parameter f. The price paid for the mathematical simplicity includes resorting to a strange double subterfuge in which, first, we retain only the purely horizontal velocity field $\mathbf{u} = \hat{\mathbf{z}} \times \nabla_H \psi$ even though vertical motion is now significant and, second, abandon P, the exact, Rossby–Ertel PV, which is advected by vertical as well as by horizontal velocities, in favour of a so-called *quasigeostrophic potential vorticity*, q, advected by the horizontal velocity only. For background

$\rho = \rho_0(z)$ and $N = N_0(z)$ we may define

$$q \;=\; f \;+\; \nabla_H^2\psi \;+\; \frac{1}{\rho_0}\frac{\partial}{\partial z}\left(\frac{\rho_0 f_0^2}{N_0^2}\frac{\partial\psi}{\partial z}\right)\,, \tag{15}$$

noting the agreement with (13) in the limit $N_0 \to \infty$, apart from the factor σ^{-1}. Omission of that factor is part of the subterfuge, making vertical advection implicit. The generalized ballerina effect is now hidden inside the last term of (15). The isobaric anomalies in T and θ, measuring small displacements and tilting of the isentropic surfaces $\mathcal{S}$, are proportional to $\partial\psi/\partial z$. For instance if $\theta_0(z)$ denotes the background potential temperature, so that $N_0^2(z) = g\,\mathrm{d}\ln\theta_0/\mathrm{d}z$, then we have $\theta - \theta_0(z) = g^{-1}\theta_0 f_0\,\partial\psi/\partial z$ to within the approximations of the theory.

The most efficient way of describing the relation between q and P is to say that $\nabla_H q$, the local horizontal or isobaric (constant-z) gradient of q, is proportional to $(\nabla_H P)_\theta$, i.e. proportional to the corresponding isentropic gradient of P. Isobaric eddy fluxes of q are correspondingly related to isentropic eddy fluxes of P.

From (15) we see that the electrostatic analogy holds, qualitatively, in three dimensions, with stretched vertical coordinate $N_0 z/f_0$. The electric charge distribution is $q - f$. This can include Bretherton delta functions. If we impose $\partial\psi/\partial z = 0$ at the lower boundary, for instance, when inverting (15) to get ψ from q, then a delta-function contribution to the last term of (15) can accommodate finite $\partial\psi/\partial z$ just above the boundary, hence a nonvanishing θ anomaly there.

Three-dimensional inversions far more accurate than quasigeostrophic are now being used in weather forecasting as well as in research and development. The most accurate possible PV inversion operators are mathematically complicated because accurate balance relations $\mathbf{u} = \mathbf{u}^{\mathrm{B}}$ are mathematically complicated, as discussed in the article on balanced flow. This difficulty can, however, be sidestepped using the forecast-initialization components of today's numerical data-assimilation technology.

6 The quasi-westward ratchet

The single time derivative acting on the generalized PV field in (1) and (2) exposes another fundamental point about the balanced dynamics. This point is well hidden within the equations expressing Newton's laws of motion in terms of the p, ρ, T, and $\mathbf{u}$ fields. The single time derivative shows for instance why all the different types of Rossby waves, including internal and topographic (surface-θ) Rossby waves, exhibit one-way phase propagation. The Earth's rotation imposes a handedness or chirality upon the wave dynamics as seen in the rotating frame. In this regard the Rossby-wave mechanism is quite unlike classical wave mechanisms, where the governing equations always contain even numbers of time derivatives, making the propagation time-reversible.

On the global or planetary scale, P has an isentropic gradient whose sign, in a coarse-grain view, is usually set by the sign of the planetary-scale gradient in f. From the Antarctic to the Arctic, f and P go from large negative to large positive values. Planetary-scale Rossby waves feel this gradient. As a result, they exhibit westward, never eastward, phase propagation relative to the mean flow. And in all cases of Rossby waves, planetary-scale or smaller, the sense of the relative phase propagation is *quasi-westward* — meaning *as if westward* — defined to be such that high or predominantly high generalized PV values are on the right. Thus for instance topographic Rossby waves, dependent on a surface gradient in the Bretherton delta function, propagate with warm surface air on the right where 'warm' is measured by θ.

The same chirality accounts for the ratchet-like, one-way character of related processes such as the self-sharpening of jet streams and the irreversible transport of angular momentum due to the dissipation of Rossby waves in the stratosphere, producing a persistent westward or retrograde mean force there, hence the *gyroscopic pumping* — always poleward and never equatorward — that drives the global-scale stratospheric circulations and chemical transports usually discussed under the headings Brewer–Dobson circulation and wave–driven circulation.

(If a zonally symmetric mean force keeps pushing air westward, then Coriolis effects keep turning it poleward — a persistent mechanical pumping action. The best-known example is Ekman pumping, the special case in which the zonal force happens to be frictional, as in classic spindown.)

7 PV mixability and strong jets

One of the mechanisms involved in the dissipation of Rossby waves is *wave breaking*, the irreversible deformation of otherwise-wavy PV contours. This definition of breaking is motivated by fundamental results in wave–mean interaction theory, namely the so-called nonacceleration theorems, which are corollaries of Kelvin's circulation theorem applied to initially-zonal material contours.

Rossby wave breaking gives rise to the irreversible mixing of PV along the isentropic surfaces $\mathcal{S}$. This can happen on a spectacularly large scale in some cases, as in the wintertime stratospheric surf zone commonly observed. Such mixing is a strongly nonlinear phenomenon and, because it tends to be highly inhomogeneous spatially, with surf zones adjacent to wavy PV contours, it often lies outside the scope of homogeneous turbulence (spectral cascade) theory. The idea of PV mixing does, however, explain the ubiquity of such quintessentially inhomogeneous phenomena as the strong jet streams observed in the atmosphere and oceans. The jet that flows along the poleward border of the stratospheric surf zone is just one example among many.

A strong jet, in the sense at hand, is nothing but a narrow core of concentrated isentropic gradients of P together with its induced velocity fields. The properties of PV inversion operators ensure that these induced velocity fields are always jet-like, flowing quasi-eastward, i.e. flowing with high PV on the left. For instance, in the westernmost part of Fig. 2 a strong jet flows southward over the Atlantic, with its core at the edge of the large colored region corresponding to high-PV stratospheric air. The jet continues around the periphery of that region past Spain toward the British Isles. Maximum wind speeds reach values of the order of $50\,\mathrm{m\,s^{-1}}$ in this case.

Once such a jet structure has formed it has a tendency to be self-sustaining, or self-sharpening. The concentrated core gradients form a waveguide or duct for Rossby waves whose dispersion properties make them liable to breaking on one or both flanks of the jet, while leaving the core intact. PV mixing adjacent to the core weakens the surrounding PV gradients and strengthens the core's PV gradients, automatically sharpening or re-sharpening the core and the jet velocity profile. Mixing across the core is strongly inhibited, thanks to the combined effects of the shear and the core's Rossby-wave quasi-elasticity.

The inhibition applies to chemical tracers as well as to PV. Countless observations of chemical tracers verify this, going back to Edwin F. Danielsen's classic 1968 aircraft observations of nuclear bomb-test debris showing distinct isotopic signatures to either side of a strong tropopause jet core. So a strong jet core can be identified with what is sometimes called a PV barrier but more aptly an eddy-transport barrier, recognizing the complementary role of the shear in the jet flanks first noted in the doctoral thesis work of M. N. Juckes. These phenomena clearly have a role in keeping the stratosphere and troposphere chemically distinct and the tropopause sharp.

The idea that the PV is mixable along the isentropic surfaces $\mathcal{S}$ merits closer examination. In using it we are setting up an analogy with chemical mixing. How far can we push that analogy? Despite its evident power to handle some kinds of strongly nonlinear phenomena, including strong-jet formation, the analogy is not always apt because the PV is not a passive tracer. Self-organizing, dynamically active phenomena like vortex rollup, and vortex merging, illustrate that isentropic anomalies in P can, in some situations, transport themselves against mean isentropic gradients of P, contrary to the mixing idea. Furthermore, there are rotational force fields that can systematically widen the range of P values on a surface $\mathcal{S}$. If we think of isentropic anomalies in P as electric-charge anomalies, this is like pair production. Such rotational force fields include those due to dissipating gravity waves.

Nevertheless, the mixing idea seems to work well in situations such as Rossby wave breaking in which a large-scale flow advects smaller-scale PV anomalies, in a manner that becomes increasingly passive-tracer-like as the large-scale strain or deformation fields shrink the advected scales. Once this advective scale-shrinkage takes hold, it goes exponentially fast on the timescale of the large-scale straining. The passive-tracer-like behavior is possible because PV inversion is relatively insensitive to small-scale PV anomalies.

Scenarios of PV transport along, rather than across, the moving surfaces $\mathcal{S}$ can remain valid even when Eqs. (1) and (2) are replaced by their diabatic and frictional generalizations. More precisely, P can be regarded as the amount per unit mass of a notional chemical substance consisting of charged particles that are permanently trapped on the moving surfaces $\mathcal{S}$. Net charge is conserved: one can have pair production and mutual annihilation, but no net creation or destruction except where a surface $\mathcal{S}$ intersects a boundary. In this picture the surfaces $\mathcal{S}$ are impermeable to the PV particles even when they are permeable to air undergoing diabatic heating or cooling — a behavior very different from that of a real chemical. The corresponding mathematical statement is sometimes called the impermeability theorem for PV.

The theorem is simple to prove, along with the conservation of net charge, by repeating the derivation that led to the flux-form conservation equation (9) but with arbitrary diabatic heating and external forces included. This reveals first that the resulting equation is still of the form $\partial(\rho P)/\partial t + \nabla \cdot (\ ) = 0$, i.e. that it is still a conservation equation in flux form — there are no source and sink terms — and second that the flux itself, the vector field acted on by the three-dimensional divergence operator, naturally takes a form such that it always represents zero transport across moving surfaces $\mathcal{S}$. Thus the surfaces $\mathcal{S}$ behave as if they were impermeable to the charged particles of PV-substance.

Of course one can always make the surfaces $\mathcal{S}$ look permeable by adding an identically nondivergent vector field to the flux. But that is arguably a needless complication, for the reasons discussed in the paper by C. S. Bretherton and C. Schär in the Further Reading list.

It is important to remember when using the analogy with chemicals that P is the amount of PV substance or PV charge per unit mass. It is the chemical mixing ratio, so called, not the amount per unit volume, to which P is analogous. Clearly, an inert chemical lacking sources or sinks can be diluted or concentrated. An extreme example is the formation of tropical cyclones, in which, in terms of the foregoing picture, PV charge is advected inwards along the surfaces $\mathcal{S}$ and greatly concentrated near the center of the cyclone. Although such processes cannot create net PV charge, they can and do create strong isentropic anomalies in P, whose inversion may yield hurricane-force winds.

8 The inhomogeneity of PV mixing

Why does PV mixing have such a strong propensity to be inhomogeneous? Part of the answer has already been indicated, namely the self-organizing properties of strong jets as eddy-transport barriers. One can add that the inhomogeneity reflects not only the dispersion properties of jet-guided Rossby waves, but also, arguably, a generic positive-feedback mechanism sometimes called the 'PV Phillips effect'. It can operate at the earlier stages of self-organization. Wherever large-scale isentropic gradients of P are weakened by PV mixing, Rossby-wave quasi-elasticity is weakened, facilitating further mixing. On the borders of such a region, the gradients are strengthened and mixing is inhibited. If shear and Rossby-waveguide ducting become important at the borders, then mixing is inhibited still further as eddy-transport barriers form.

There is yet another reason to expect PV mixing to be inhomogeneous. It is especially clear in the case of surfaces $\mathcal{S}$ that span the globe and are therefore topologically spherical, as in the stratosphere and upper troposphere (and also in the solar interior). If we extend the surface integrals in Eqs. (8) to the entire sphere, there is no enclosing contour Γ and we have

$$\iint_{\mathcal{S}} P\,\sigma\,\mathrm{d}A \;=\; 0\;, \tag{16}$$

stating that on each topologically spherical $\mathcal{S}$ there are equal numbers of positively and negatively charged PV particles, regardless of whether the flow is forced, dissipating, or dissipationless. This is consistent with the charge-conservation and impermeability theorems. The integral relation (16) imposes a severe constraint on the possible evolution of the isentropic distributions of PV on each such $\mathcal{S}$, hence on the possible evolution of the flow. That constraint is enough in itself to make uniform or homogeneous mixing highly improbable, as the following argument shows.

Consider a hypothetical situation in which the mixing is uniform, as if the distribution of P on a surface $\mathcal{S}$ were subject to a uniform horizontal diffusivity. Under the constraint (16), in which σ is strictly positive, the perfectly mixed state toward which the distribution of P would then relax can only be a state in which $P = 0$ everywhere on $\mathcal{S}$. But invertibility says that the entire surface $\mathcal{S}$ would then have to be at rest relative to the stars, apart from oscillations representing imbalance such as sound waves and inertia–gravity waves. In a rapidly rotating system like the Earth's atmosphere, with strong Coriolis effects and Rossby numbers typically small, such a state of rest would be overwhelmingly improbable. It would require a redistribution of angular momentum that would not only have an implausibly large magnitude but would also need to take a very special form.

9 The Taylor identity

The hypothetical situation just sketched is an implausible extreme case, but it illustrates another fundamental fact. Almost any isentropic redistribution of PV, or other modification to the PV field, will be accompanied by changes in the distribution of angular momentum.

The PV mixing associated with breaking Rossby waves is just one piece of what might be called a wave–turbulence jigsaw in which wave propagation has just as crucial a role as wave breaking, through wave-induced transport of angular momentum such as that giving rise, as already mentioned, to the gyroscopic pumping of the Brewer–Dobson and other global-scale mean circulations. A by-product is that eddy fluxes of momentum often look anti-frictional, exhibiting the so-called 'negative viscosity' that was once regarded as a great enigma of atmospheric science, but is now recognized as a natural consequence of the interplay between wave generation, wave propagation, and wave breaking.

The way in which the jigsaw fits together is reflected in a central result from quasigeostrophic theory, which for historical reasons might be called the Taylor–Charney–Stern–Bretherton–Eady–Green identity. It is traceable back to a seminal 1915 paper by G. I. Taylor that applies to the limiting case (14). For brevity it will here be called the Taylor identity. It interrelates the eddy fluxes of momentum and PV. The standard form of the identity is for disturbances to a zonal-mean state. Using overbars and primes to denote the zonal mean and fluctuations about it, which can have arbitrary amplitude, we readily find from (15) that

$$\overline{v'q'} \;=\; \frac{1}{\rho_0}\left(\frac{\partial F}{\partial y} + \frac{\partial G}{\partial z}\right) \tag{17}$$

where

$$(F,\, G) \;=\; \rho_0\left(-\overline{u'v'},\; \frac{f_0 g}{N_0^2\theta_0}\,\overline{v'\theta'}\right)\;, \tag{18}$$

the so-called Eliassen–Palm (EP) flux or effective stress (minus the effective eddy momentum flux). This quantifies the Rossby-wave-induced momentum transport. Here $(u',\, v') = (-\partial\psi'/\partial y,\, \partial\psi'/\partial x)$, the eastward and northward components of $\hat{\mathbf{z}} \times \nabla_H \psi'$, and $g\theta' = \theta_0 f_0 \partial\psi'/\partial z$. The vertical component of the EP flux is the same as the pressure-fluctuation-induced form stress defined in oceanography (sometimes less aptly called 'form drag'), the mean zonal force per unit area across an undulating stratification surface, whose vertical displacement is $-g\theta'/N_0^2\theta_0$. The Taylor identity has special importance not least because of its validity for strongly nonlinear flows, such as breaking Rossby waves. No small-amplitude assumption is needed.

For instance, in order to create the wintertime stratospheric surf zone, through PV mixing producing downgradient, i.e. negative, $\overline{v'q'}$, there needs to be a convergence of Rossby-wave activity from outside the surf zone, making the right-hand side of (17) negative as well, and reducing the angular momentum of the surf zone. An exquisitely precise illustration of how everything fits together is provided by the Stewartson–Warn–Warn theory of nonlinear Rossby-wave critical layers. These are narrow surf zones and well illustrate the strong inhomogeneity of the wave–turbulence jigsaw and the typical way in which (17) is satisfied.

Further Reading

Appenzeller, C., Davies H. C. and Norton, W. A. (1996). Fragmentation of stratospheric intrusions. *J. Geophys. Res.* 101, 1435–1456. (This paper, the source of Fig. 2 above, fills in many of the details of the associated weather systems and their largely-advective evolution. There are important cross-checks from satellite water-vapor imagery.)

Arbogast, P., Maynard, K. and Crepin, F. (2008). Ertel potential vorticity inversion using a digital filter initialization method. *Q. J. Roy. Meteorol.*

Soc. 134, 1287–1296. (This work, to which Dr P. Berrisford kindly drew my attention, uses the weather-forecasting technology at Météo France. The inversions produce three-dimensional velocity fields including vertical velocity. At the time of writing there is also a web page at www.cnrm.meteo.fr entitled "PV inversion as a tool for synoptic forecasting".)

Bannon, P. R., Schmidli, J. and Schär, C. (2003). On potential vorticity flux vectors. *J. Atmos. Sci.* 60, 2917–2921. (Draws on mathematical generalizations of (9) going back to Ertel and Truesdell to include multiply-buoyant fluids like seawater, for which material invariance fails. **Caution:** the authors use the terms 'source' and 'sink' in their non-chemical, non-conservational, purely causative sense.)

Bretherton, C. S. and Schär, C. (1993). Flux of potential vorticity substance: a simple derivation and a uniqueness property. *J. Atmos. Sci.* 50, 1834–1836. (The choice of flux vector that satisfies the impermeability theorem is shown to be uniquely simple: it is the only choice whose nonadvective term depends linearly on the diabatic heating rates and rotational, non-potential force fields that break material invariance.)

Bühler, O. (2009). *Waves and Mean Flows.* Cambridge: University Press. (By far the best account of wave–mean interaction fundamentals, supplying a wealth of telling examples and bringing out, for instance, the connection between nonacceleration theorems and Kelvin's circulation theorem.)

Dritschel, D. G. and McIntyre, M. E. (2008). Multiple jets as PV staircases: the Phillips effect and the resilience of eddy-transport barriers. *J. Atmos. Sci.* 65, 855–874. (This discussion of strong jets includes a historical introduction noting the seminal contributions of G. I. Taylor, N. A. Phillips, and R. E. Dickinson that were keys to solving the old 'negative viscosity' enigma; the PV Phillips effect is named after a different researcher, O. M. Phillips. One historical correction is needed: Charney and Stern appear to have been the first in print with the quasigeostrophic version (17) of the Taylor identity, in 1962.)

Haynes, P. H. (1989). The effect of barotropic instability on the nonlinear evolution of a Rossby-wave critical layer. *J. Fluid Mech.* 207, 231–266. (The most comprehensive account of the Stewartson–Warn–Warn theory and its further developments, illustrating in precise detail how the wave–turbulence jigsaw works.)

Haynes, P. H. and Anglade, J., 1997: The vertical-scale cascade of atmospheric tracers due to large-scale differential advection. *J. Atmos. Sci.* 54, 1121–1136. (Shows the robustness of exponentially fast advective scale-shrinkage of passive tracers in stratified, rotating, vertically sheared flows.)

Hoskins, B. J., McIntyre, M. E. and Robertson, A. W. (1985). On the use and significance of isentropic potential-vorticity maps. *Q. J. Roy. Meteorol. Soc.* 111, 877–946; Corrigendum 113, 402–404. (This major review presents typical examples from the real atmosphere and clarifies why it is isentropic — not horizontal or isobaric — maps, distributions, gradients and fluxes of P that are dynamically significant.)

Korty, R. L. and Schneider, T. (2007). A climatology of the tropospheric thermal stratification using saturation potential vorticity *J. Clim.* 20, 5977–5991. (Develops a proposal by K. A. Emanuel to characterize afresh the principal atmospheric air masses by means of a saturation PV, in which θ is replaced by the saturation value of the moist equivalent potential temperature θ_e^*; q.v. also for references to earlier work on varieties of 'moist PV'.)

Mestel, L. (2012). *Stellar Magnetism*, 2nd edn. Oxford: University Press. (See chapter 8 on 'late-type stars', which include solar-type stars. Recognizing the implications of PV fundamentals for stratified turbulence has led, indirectly but powerfully, to radical advances in understanding the dynamics of the stratified solar interior and the so-called tachocline.)

Rossby, C.-G. (1936). Dynamics of steady ocean currents in the light of experimental fluid mechanics. *Pap. Phys. Oc. Meteorol.* (Mass. Inst. of Technology and Woods Hole Oc. Instn.) 5(1), 1–43.

Rossby, C.-G. (1938). On the mutual adjustment of pressure and velocity distributions in certain simple current systems, II. *J. Mar. Res.* 2, 239–263.

Rossby, C.-G. (1940). Planetary flow patterns in the atmosphere. *Quart. J. Roy. Meteorol. Soc.* 66 (Suppl.), 68–87. (Rossby's three great pioneering papers seem almost forgotten today. I thank Jule G. Charney, Norman A. Phillips, George W. Platzman, and Roger M. Samelson for dispelling my historical illusions, little by little.)

Young, W. R. (2012). An exact thickness-weighted average formulation of the Boussinesq equations. *J. Phys. Oc.* 42, 692–707. (This is a major advance in the theory of residual circulations and the Taylor identity; see also transformed Eulerian mean. With the help of judiciously-chosen averages on stratification surfaces, not necessarily zonal averages, Young finds exact results that, in the case of the Taylor identity, are formally no more complicated than the standard quasigeostrophic Taylor identity, (17) above.)