Nonuniqueness of Attribution in Piecewise Potential Vorticity Inversion

JOSEPH EGGER
Meteorological Institute, University of Munich, Munich, Germany

THOMAS SPENGLER
Geophysical Institute, University of Bergen, and Bjerknes Centre for Climate Research, Bergen, Norway

(Manuscript received 7 February 2017, in final form 19 December 2017)

ABSTRACT

Piecewise potential vorticity inversion (PPVI) seeks to determine the impact of observed potential vorticity (PV) anomalies on the surrounding flow. This widely used technique is based on dividing a flow domain $D$ into subdomains $D_1$ and $D_2 = D - D_1$. The influence of PV in $D_1$ on the flow in $D_2$ is assessed by removing all PV anomalies in $D_2$ and then inverting the modified PV in $D_2$. The resulting flow with streamfunction $c_1$ is attributed to the PV anomalies in $D_1$. The relation of PV in $D_1$ to $c_1$ in $D_2$ is not unique, because there are many PV distributions in $D_1$ that induce the same $c_1$. There is, however, a unique solution if the ageostrophic circulation is included in the inversion procedure.

The superposition principle requires that the sum of inverted flows with PV $\neq 0$ in $D_2$ and the complementary ones with PV $\neq 0$ in $D_1$ equal the inverted flow for the complete observed PV in $D$. It is demonstrated, using two isolated PV balls as a paradigmatic example, that the superposition principle is violated if the ageostrophic circulation is included in PPVI, because the ageostrophic circulation cannot be associated with only one of the anomalies.

Inversions of Ertel's PV are carried out using Charney's balance condition. PPVI is not unique in that case, because many different PV fields can be specified in $D_1$, which all lead to the same inverted flow in $D_2$. The balance condition assumes vanishing vertical velocity $w$ so that uniqueness cannot be established by including $w$ in the inversion, as was possible in the quasigeostrophic case.

1. Introduction

Potential vorticity (PV) has been rarely used in diagnostics before the 1970s and 1980s even though its usefulness was demonstrated in a series of papers (e.g., Kleinschmidt 1955; Danielsen 1968). It was presumably the complicated structure of PV that precluded widespread use of this quantity, where PV

$$ Q = \frac{\mathbf{\omega}_a \cdot \nabla \theta}{\rho} $$

(1)

is a nonlinear function with absolute vorticity $\mathbf{\omega}_a$, potential temperature $\theta$, and density $\rho$. As a matter of fact, lack of data led Kleinschmidt (1955, p. 96) to speculate about the “existence of a non-conservative force” when analyzing the PV of a cutoff low.

The general attitude toward PV changed when Thorpe (1985) and Hoskins et al. (1985) demonstrated that $Q$ can be inverted [potential vorticity inversion (PVI)]; that is, all flow variables can be derived from $Q$ on the basis of hydrostatic equilibrium, a balance condition, and suitable boundary conditions. The balance conditions are an essential part of PVI, ranging from geostrophic balance to rather complicated digital filter schemes (Arbogast et al. 2008), where most conditions are nonlinear, including the frequently used scheme of Charney (1955). The invertibility of PV was well known in quasigeostrophic theory, where PV is a linear function so that inversion is also a linear problem (e.g., Hakim et al. 1996). Quasigeostrophic models invert PV as part of the time integration procedure (e.g., Bleck 1974). Inversion of $Q$, however, requires iterative techniques because of the nonlinearity of $Q$. Following Hoskins et al. (1985), the flow state deduced from the PV field by inversion is referred to as “induced” (see also McIntyre 2012), although a causal link does not exist (Spengler and Egger 2012).

PVI aims to recover that atmospheric state that has been used to determine $Q$ in the first place. This effort
provides little new information except with respect to the quality of the balance condition. However, inversion of prescribed PV anomalies yields useful insight into the atmospheric state associated with such anomalies (Hoskins et al. 1985; Davis 1992; Bishop and Thorpe 1994). Piecewise inversion of PV (PPVI) “is perhaps the most useful diagnostic application of the invertibility principle” (Davis and Emanuel 1991, p. 1936). This technique seeks to determine an influence of PV anomalies on the surrounding flow. Assume a flow domain D, where a specific subdomain D1 of D is selected, which contains PV anomalies of dynamic interest. The target region D2 = D − D1 is separated from the source region D1 by a boundary S. A new PV field is then defined, which keeps the observed PV in D1 but removes all observed PV anomalies in D2. Inversion of this new PV distribution yields a flow field in D2, which is attributed to the PV in D1. For example, the inverted winds in D2 are believed to represent the contribution of PV in D1 to the advection of PV anomalies in D2. A complementary procedure keeps PV as observed in D2 and removes that in D1. The superposition principle (Bishop and Thorpe 1994) requires that the flow inverted from the complete PV in D is the sum of those obtained by PPVI and its complementary version. Attribution is meaningless if parts of the wind field cannot be assigned to PV either in D1 or D2. Although most published examples of PPVI rely on Q, quasigeostrophic PPVI is of interest as well. Thus, both cases will be considered.

The statement of Bracegirdle and Gray (2009, p. 882) that “the difference between the unmodified and the modified fields then yields a measure of the contribution of the anomaly to the flow and its interaction with other anomalies” aptly describes the currently held view on PPVI. For example, Wu and Emanuel (1995, p. 69) investigate the influence of PV anomalies near a hurricane and state that “one can identify the influence of each PV perturbation on hurricane movement” according to PPVI. PPVI has also been applied to the interaction of the stratosphere and troposphere (Hartley et al. 1998) and to cyclone growth with strong latent heat release (Pomroy and Thorpe 2000; Ahmadi-Givi et al. 2004). The contribution of inverted winds to PV advection (Teubler and Riemer 2016) and atmospheric diagnostics like frontogenesis functions have been investigated this way (Baxter et al. 2011). However, some questions remain with respect to the interpretation of PPVI. Hakim (2008, p. 2950) points out that “multiple states may share the same PV distribution.” In addition, there are problems with the uniqueness of attribution, because the inversion of nonlinear functions is not unique (Davis 1992; Hakim 2008). Moreover, the contributions of two separate patches of PV cannot be exactly superposed. Although these effects of nonlinearity are well understood (see also Birkett and Thorpe 1997), their interpretation with respect to attribution is not yet resolved.

As discussed carefully by Thorpe (1997), an ageostrophic circulation can also be attributed to specific PV anomalies by solving the related omega equation (see also Clough et al. 1996). This equation contains the geostrophic temperature and vorticity advection, which are nonlinear terms. Although it is common to restrict PVI to the balanced flow, both the geostrophic and the ageostrophic circulation should be attributed to PV. As the ageostrophic flow has not been discussed in the context of PPVI so far, it is a main aim of our study to shed light on its role in PPVI.

2. Quasigeostrophic PPVI

a. Green’s function and PPVI

The main points of our argument can be made by accepting an f-plane atmosphere with constant mean density ρ and Brunt–Väisälä frequency N. The domain D may be infinite or located between a rigid lid and the bottom. The subdomain D1 with volume V1 has a closed surface S as boundary, which may contain some part B of the boundary of D. Observations provide the pressure field pobs. Specification of anomalies may sometimes be difficult in PPVI, but this problem is not central to our discussion so that we simply assume a background state where pobs is a perturbation. The PV qobs can be computed using

\[ \nabla^2 p = \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) p = \bar{p}qf, \]  

(2)

where \( f = f/N \) and \( p_{\text{obs}} \) must be inserted in (2). Then

\[ G(r_1, r_2) = -\frac{1}{4\pi |r_1 - r_2|}, \]  

(3)

is the Green’s function for quasigeostrophic PV with modified position vectors \( r_1 \) and \( r_2 \) in an infinite domain. As pointed out by Davis and Emanuel (1991), and many others, the Green’s function technique appears to support the view that \( p_1 \) in \( D_2 \) away from \( S \) can be attributed to PV in \( D_1 \), because

\[ p_1(r_2) = \bar{p}f \int_{V_1} q_{\text{obs}}(r_1) G(r_1, r_2) \, dV_1, \]  

(4)

with \( r_1 \) and \( r_2 \) located in \( D_1 \) and \( D_2 \), respectively. Thus, the observed PV in \( D_1 \) appears to determine the
inverted pressure in $D_2$. However, partial integrations lead to

$$p_1(r_2) = \int_S \left\{ \nabla[p_{\text{obs}}(r_1)G(r_1, r_2)] - \frac{p_{\text{obs}}(r_1)(r_1 - r_2)}{4\pi|r_1 - r_2|^3} \right\} \cdot n\, dS,$$

(5)

with normal vector $n$ on $S$, $r_1$ pointing onto $S$, and $r_2$ in $D_2$. Hence, $p_1$ in (5) is a pressure field with vanishing PV that is completely determined by the observed values of pressure and normal pressure gradient at $S$. Thus, attribution of $p_1$ to PV in $D_1$ appears not to be meaningful, because $q_{\text{obs}}$ is not part of (5). In fact, an infinitely large number of fields of pressure and related PV can be prescribed in $D_1$, which all satisfy (5), provided these pressures and pressure gradients at $S$ coincide with the observed pressure terms in (5). It is possible to also satisfy conditions at the rigid boundaries of $D$ by introducing mirror domains, but that complicates the result without adding further insight.

More complicated cases, where a height-dependent reference density and Brunt–Vaisälä frequency are prescribed and upper and lower boundary conditions have to be satisfied, can be dealt with by assuming that a solution $p^*$, with $p^* = p_1$ in $D_2$, has been found using numerical methods in PPVI after choosing a specific $p_{\text{obs}}$. We may then replace the observed pressure distribution in $D_1$ by a new field $p_2$ such that the observed pressure and its normal derivative are not altered at $S$. After determining $q_a$ in $D_1$ from $p_{\text{ev}}$, a piecewise inversion of $q_a$ is not necessary, because the same pressure $p_1$ must result from (5) in $D_2$, as was the case prior to switching from $p_{\text{obs}}$ to $p_a$. It follows that both $q_{\text{obs}}$ and $q_a$ induce the same response in $D_1$. Thus, unique attribution of the inverted $p_1$ to PV in $D_1$ is not possible.

Altogether, it follows that PPVI provides an association of a flow with vanishing PV in $D_2$ to the pressure and pressure gradients at $S$. There is no impact of PV in $D_1$ on the inverted flow in $D_2$. One may modify $p_{\text{obs}}$ in $D_1$ arbitrarily as long as the boundary values on $S$ remain the same. Hence, there is no uniqueness of attribution. However, there is an integral constraint, because integration of (2) over the volume $V_1$ of $D_1$ yields

$$\hat{q}_1 = V_1^{-1} \int q'dV_1 = (\nabla f V_1)^{-1} \int_S \nabla p_{\text{obs}} \cdot n\, dS.$$

(6)

Thus, $\hat{q}_1$ cannot be altered by prescribing a field $p_n$ in $D_1$. Complications arise if $S$ contains parts of the boundary of $D$, though we will not discuss them here as they are of minor importance for our argument.

As attribution seeks to determine an influence of PV in $D_1$ on the flow in $D_2$, one looks at PV in $D_1$ and argues about its role with respect to the flow in $D_2$. Such an argument becomes fairly meaningless if rather different PV fields in $D_1$ induce the same response and only the pressure field at $S$ determines the response.

In the context of attribution, it is also common to invoke the electrostatics analogy (Birkett and Thorpe 1997), where (5) is valid as well but PV is replaced by electric charge density and $p_1$ by electric potential. Removing all charges in $D_2$ would not affect the electric field induced in $D_2$ by the charges in $D_1$. The charges in $D_1$ exert a force in $D_2$ that would accelerate charged matter there. Thus, there is a clear physical impact of the charges in $D_1$ on $D_2$. However, no force is exerted by PV in $D_1$ on PV in $D_2$, and there are no related accelerations. The physical interpretation in the electrostatic case differs profoundly from that in PPVI despite the fact that both cases are based on solutions of the Laplacian with sources. Moreover, an important part of the induced circulation is not represented at all in the electrostatic case, because the ageostrophic circulation has no counterpart in electrostatics. We have to conclude that the electrostatics analogy does not help to clarify the issue of attribution in PPVI. In other words, the meanings of “induction” and “attribution” are completely different in both cases.

b. Ageostrophic circulation

Although Thorpe (1997) and Clough et al. (1996) calculated the ageostrophic circulation of PV anomalies, the role of this circulation in PPVI has not been discussed yet. Its importance in the context of PPVI will be outlined in this section.

The ageostrophic vertical velocity $w$ with respect to the coordinate $z$ follows from the omega equation:

$$\nabla^2 w + \varepsilon^2 \frac{\partial^2 w}{\partial z^2} = \frac{f}{N^2} \left\{ \frac{\partial}{\partial z} \left( v_g \cdot \nabla (\nabla^2 \psi) \right) - \nabla^2 \left( v_g \cdot \nabla \frac{\partial \psi}{\partial z} \right) \right\},$$

(7)

with geostrophic wind $v_g$, the right-hand side combines the geostrophic advections of temperature and vorticity. Note that advection by a constant mean zonal flow does not contribute to (7). It is fairly easy to derive $w$ in domains with vanishing PV, where $\nabla^2 \psi = -\varepsilon^2 \frac{\partial^2 \psi}{\partial z^2}$ with streamfunction $\psi = p/f$. Thus,

$$v_g \cdot \nabla (\nabla^2 \psi) = -\varepsilon^2 v_g \cdot \nabla \left( \frac{\partial^2 \psi}{\partial z^2} \right) = -\varepsilon^2 \frac{\partial}{\partial z} \left( v_g \cdot \nabla \frac{\partial \psi}{\partial z} \right),$$

(8)

where the vanishing of temperature advection by the thermal wind has been taken into account. Hence,

$$w = -\frac{f}{N^2} v_g \cdot \nabla \frac{\partial \psi}{\partial z}$$

(9)
is a solution of (7) for vanishing PV, where we do not impose any boundary conditions. However, (9) is invalid for flows with \( \partial^2 \psi / \partial z^2 = 0 \), because (8) is not satisfied in that case.

For the case with \( q_{\text{obs}} \) in \( D_1 \) and vanishing PV in \( D_2 \), the full solution for \( w_1 \) in \( D_2 \) is, however, not fully determined by (9), because the continuity of \( w_1 \) and \( \partial w_1 / \partial z \) has to be imposed at \( S \). Thus, while the inverted \( p_1 \) in \( D_2 \) does not depend on \( q_{\text{obs}} \) in \( D_1 \), each solution for \( p_1 \) induces a specific vertical velocity \( w_1 \) in \( D_2 \). These induced ageostrophic circulations are not necessarily the same just because the induced geostrophic winds are equal. We present a rather simple example in the next section, where the ageostrophic circulation related to PPVI can be obtained analytically and demonstrate that uniqueness can be established in this case.

c. Illustrative example

A rather simple example of quasigeostrophic PPVI is presented in order to illustrate our argument. We consider a two-dimensional model atmosphere of depth \( H (-H/2 \leq z \leq H/2) \) and zonal extent \( L (-L/2 \leq x \leq L/2) \), with a flat bottom at \( z = -H/2 \) and a lid on top at \( z = H/2 \). This model is based on the same formalism as in the previous section and assumes zonal periodicity. Perturbations have only zonal wavenumber \( 2\pi/L \). All equations can be solved analytically.

The selection of the target region \( D_2 \) depends on the flow situation, but it is a common choice to investigate the “impact” of an upper layer on the atmosphere below (e.g., Davis and Emanuel 1991; Wu and Emanuel 1995; Bracegirdle and Colucci 1996; Hartley et al. 1998; Bracegirdle and Gray 2009). We consider this case here by associating \( D_1 \) with the layer \( 0 \leq z \leq H/2 \) and \( D_2 \) with that below. Thus, the separating surface \( S \) is located at \( z = 0 \). PPVI is thought to quantitate the impact of PV anomalies in \( D_1 \) on the lower layer \( D_2 \).

Let the inverted pressure \( p_1 = p_{\text{obs}} + p_{D_1}^* \) in \( D_1 \) and \( p_1 = p_{D_2}^* \) in \( D_2 \), where \( q = 0 \) is solved by \( p_{D_1}^* \) in \( D_1 \) and by \( p_{D_2}^* \) in \( D_2 \). Thus, the PV related to \( p_1 \) equals \( q_{\text{obs}} \) in \( D_1 \) and vanishes in \( D_2 \), but the additional pressure perturbations \( p_{D_1}^* \) and \( p_{D_2}^* \) are needed to satisfy the matching conditions with continuity of \( p_1 \) and \( \partial p_1 / \partial z \) at \( z = 0 \). On the other hand, the boundary conditions for \( p_{D_1}^* \) \( (p_{D_2}^*) \) at \( z = -H/2 \) \( (z = H/2) \) are not unique, as is typical of PPVI (e.g., Davis and Emanuel 1991). For example, one may prescribe vanishing \( p_{D_1}^* \) \( \text{or} \) \( T_{D_1}^* \) at the upper and lower boundaries, where \( \partial p_1 / \partial z \) is proportional to temperature in the quasigeostrophic context. The temperature condition is chosen in what follows as it avoids introducing temperature anomalies at the boundaries, which are often interpreted as PV anomalies (e.g., Hoskins et al. 1985). Thus, \( T_{D_1} = 0 \) at \( z = H/2 \) and \( T_{D_2} = 0 \) at \( z = -H/2 \).

The observed pressure

\[
p_{\text{obs}} = P_0 \left[ \cos(kx + mz) \cos(ny) \right] F(z) = \psi_{\text{obs}} \mathbf{f} \tag{10}
\]

is prescribed in \( D \) with wavenumbers \( k, m, \) and \( n \); streamfunction \( \psi \); and a constant amplitude \( P_0 \). The factor \( F(z) \) has to satisfy \( F(0) = 1 \) and \( dF/dz(0) = 0 \). The simplest choice in our context is \( F = \cos(\gamma z) \) with wavenumber \( \gamma \), but \( F = \cosh(\gamma z) \) is another simple possibility. The observed PV \( q_{\text{obs}} \) follows from (10) by applying (2), which is shown in Fig. 1a for \( F = \cos(\gamma z) \) and \( \gamma = 2\pi/H \). The PV extrema at \( S (z = 0) \) are seen also at \( z = \pm 3H/4 \) and are shifted zonally. PPVI tries to clarify the impact of the upper extrema in \( D_1 \) \( (0 \leq z \leq H/2) \) on those in \( D_2 \) \( (-H/2 \leq z \leq 0) \). The PV
contours are straight slanted lines for $\gamma = 0$ (Fig. 1b). A standard guess would predict that PPVI yields rather different responses in $D_2$ for these choices of $\gamma$. Amplitudes increase upward and downward if we choose the cosh function for $F$.

The geostrophic wind in $D$ is represented by $\psi$, and the ageostrophic $w$ follows from (7) and (10). After simple manipulations, we obtain

$$w = [A + B \cos(2\gamma z)] + C \cosh(2ne^{-1}z) \sin(2ny)$$

(11)

with $\psi_0 = P_0/\bar{\rho} f$, 

$$A = -\frac{nk \psi_0^2}{g},$$

(12)

$$B = -\frac{n^3k \psi_0^2}{g(4n^2 + 4\gamma^2)},$$

(13)

$$C = -\frac{A + B \cosh(\gamma H)}{\cosh(nH)}.$$  

(14)

The ageostrophic meridional velocity is

$$v_{ag} = \left[-\gamma \frac{B}{n} \sin(2\gamma z) + C \sinh(2ne^{-1}z)\right] \cos(2ny),$$

(15)

and the ageostrophic circulation consists of rolls with zonally oriented axes. Neither $w$ nor $v_{ag}$ depend on $x$.

As outlined above, PPVI accepts $q_{obs}$ in $D_1$, but pressure fields $p_{D_1}^b$ and $p_{D_2}^b$ of vanishing PV must be added in the respective layers to ensure continuity of $p_1$ and $\partial p_1/\partial z$ at $S$. The result of this matching procedure is

$$p_{D_1}^b = P_0 \left[ \frac{-\cos(kx)}{2 \cosh(\alpha H/2)} - \frac{\sin(kx)m}{2a \sinh(\alpha H/2)} \right] \times \cos(ny) \cosh \left[ \alpha \left( z - \frac{H}{2} \right) \right],$$

(16)

$$p_{D_2}^b = P_0 \left[ \frac{\cos(kx)}{2 \cosh(\alpha H/2)} - \frac{\sin(kx)m}{2a \sinh(\alpha H/2)} \right] \times \cos(ny) \cosh \left[ \alpha \left( z + \frac{H}{2} \right) \right],$$

(17)

with $\alpha = (k^2 + n^2)^{1/2} e^{-1}$ as in the Eady model. Thus, $p_1 = p_{obs} + p_{D_1}$ in $D_1$ and $p_1 = p_{D_2}$, and PPVI claims that $p_{D_2}^b$ is induced by $q_{obs}$ in $D_1$. This interpretation requires also $p_{D_1}^b$ to be induced by $q_{obs}$. We are not aware that this result has been discussed in the literature.

The coefficients of $\sin(kx)$ are the same in $p_{D_1}^b$ and $p_{D_2}^b$, while those of $\cos(kx)$ are of opposite sign. It is easy to check that continuity requirements at $z = 0$ are satisfied. Both $p_{D_1}^b$ and $p_{D_2}^b$ are displayed in Fig. 2. PPVI induces a high with an upright vertical axis west of the observed high with its center at $x = 0$. Note that $T_{D_{1/2}}^b = 0$ at the lids.

It is the main point here that neither $p_{D_1}^b$ nor $p_{D_2}^b$ depend on $\gamma$. Any value of $\gamma$ is compatible with the induced geostrophic flow $\psi_{D_1}^b = p_{D_1}^b/\bar{\rho} f$ in $D_2$. In particular, both PV patterns in Fig. 1 induce the same pressure $p_{D_2}^b$ in $D_2$. This is a clear demonstration of the nonuniqueness of PPVI, which is predicted by (5).

The vertical wind component $w$ has been evaluated as well, but the resulting expressions are rather lengthy. These complications stem mainly from (16) and (17), because there are now additional winds and temperatures. There is, however, no contribution by vorticity advection in the omega equation. The complete formula for $w$ contains 25 terms, but the main message can be conveyed by writing down just that part of the vertical motion that is directly due to $p_{obs}$ in $D_1$. Thus, (11) is accepted in $D_1$ with corresponding values of $w$ and $dw/\partial z$ at $S$. A matching of vertical motions is achieved by adding functions $e^{2ne^{-1}z}$. The procedures are quite similar to those applied for (16) and (17), and the result is

$$w_2 = \left[ A + B \cos(\gamma H) \right] \left\{ \frac{[\cosh(ne^{-1}H)]^2 + [\sinh(ne^{-1}H)]^2}{2 \sinh(ne^{-1}H) \cosh(ne^{-1}H)} \sin(2ny) \sinh \left[ 2ne^{-1} \left( z + \frac{H}{2} \right) \right] \right\},$$

(18)
Unlike the geostrophic streamfunction \( \psi_g \), the ageostrophic circulation in \( D_2 \) depends on \( \gamma \). Thus, PPVI is unique in this case if the ageostrophic circulation is included.

d. Superposition

Although we have shown in the preceding section that incorporation of the ageostrophic circulation in PPVI may lead to uniqueness in PPVI, there is now the problem of superposability. The omega equation [see (7)] is nonlinear, and superposition is not satisfied. One cannot even argue that this is a minor problem. This will be demonstrated by analyzing a situation that is paradigmatic for PPVI. We assume two spherical PV anomalies with constant PV \( q_1 \) and \( q_2 \) with centers \((\delta, 0, 0)\) and \((-\delta, 0, 0)\) and radii \( R_1 \) and \( R_2 \) embedded in an atmosphere of vanishing PV. Thus, the streamfunction is \( \psi = \psi_1 + \psi_2 \) with

\[
\psi_1 = -R_1^3 q_1 / 3d_1, \\
\psi_2 = -R_2^3 q_2 / 3d_2,
\]

with \( d_{1/2} = [(x - \delta)^2 + y^2 + \varepsilon^{-2}z^2]^{1/2} \) (e.g., Bishop and Thorpe 1994) at a location outside the spheres. This situation has been used by Birkett and Thorpe (1997) to analyze the superposability problem, but these authors did not calculate the ageostrophic circulation.

Let the plane \( x = 0 \) serve as a separating surface \( S \) with half-space \( x > 0 \) as \( D_1 \) so that \( \psi = \psi_1 F(x) \) for \( x > 0 \) in \( D_1 \) and \( \psi = \psi_1 \) in \( D_2 \) satisfies the boundary conditions at \( S \) if \( F(0) = 1, \partial F/\partial x(0) = 0 \). Because PV is given by the PV ball in \( D_1 \) for \( F = 1, \psi = \psi_1 \) is the solution to the paradigmatic problem for PPVI. However, PPVI is not unique because PV in \( D_1 \) depends on \( F \), while the flow in \( D_2 \) \((x < 0)\) is not affected by the choice of \( F \). The related geostrophic wind describes the motion of the second sphere. This information is quite useful when considering the interaction of PV anomalies (e.g., Egger 2009).

It is straightforward to obtain \( w \) with (19) according to (9) so that

\[
w = \frac{88^2 R_1^2 R_2^2 q_1 q_2 f^{-1} xyz}{3(d_1 d_2)^3},
\]

with \( w = 0 \) if there is only one anomaly. The vertical motion is upward for \( q_1 q_2 > 0 \) if \( xyz > 0 \) and downward in the remaining space. It is antisymmetric with respect to the planes \( x = 0, y = 0, \) and \( z = 0 \). In particular, with \( q_1 = q_2 > 0 \), we have a rotating pair with ascent (descent) ahead (behind) each sphere for \( z > 0 \) and reversed values of \( w \) for \( z < 0 \). Similar patterns of vertical motion have been observed in numerical experiments with baroclinic dipoles (e.g., Pallás-Sanz and Viúdez 2007), but we are not aware that \( w \) has been calculated analytically for the basic situation of a dipole as in (19).

Another interesting issue is the advection of temperature by the flows obtained in PPVI. The salient features of this problem can be addressed by looking at the case of the two PV balls. The total streamfunction \( \psi = \psi_1 + \psi_2 \) and vertical motion \( w \) yield the temperature tendency

\[
\frac{\partial}{\partial t} \frac{\partial \psi}{\partial z} = -\left( \mathbf{v} + \mathbf{v}_s \right) \cdot \nabla \frac{\partial \psi}{\partial z} - \frac{N^2}{\mathcal{F}} w, \tag{21}
\]

where PPVI wishes to single out the contribution of specific PV anomalies to such tendencies. Thus,

\[
\left( \frac{\partial}{\partial t} \frac{\partial \psi}{\partial z} \right)_1 = -\mathbf{v}_s \cdot \nabla \frac{\partial \psi}{\partial z} \tag{22}
\]

describes the impact of the PV ball 1 on the temperature tendency according to PPVI. A corresponding equation is valid for the other ball. The sum of both equations gives the correct first term on the right-hand side of (21) but misses the second term because \( w_1 = w_2 = 0 \). Thus, the total temperature tendency resulting from PPVI is incorrect, while the total tendency of total PV is correct.

Strictly speaking, (20) is valid only for a point vortex model where the product \( q_i R_i^3 \) has to be assumed finite. Assuming \( q_i \approx 3 \times 10^{-5} \text{ s}^{-1}, R_i \approx 5 \times 10^5 \text{ m}, \) and \( \delta = 2R_i, \) we obtain \( w \approx 5 \times 10^{-1} \text{ ms}^{-1} \) for \( xyz \approx R_i, d \approx R_i. \) The velocity is large near the centers but decreases rapidly with increasing distance. Altogether, we find that the pair has an ageostrophic circulation, but the vertical velocity associated with both spheres is not the sum of the vanishing vertical velocities derived from each in isolation. Hence, attribution of the ageostrophic circulation to selected PV anomalies is impossible. This result is an immediate consequence of the nonlinearity of the omega equation, which has not been taken into account in PPVI yet.

3. General form of PV

Inversion of nonlinear \( Q \) [see (1)] is more demanding than that of \( q \) and can only be performed iteratively. In particular, there is no simple solution to the condition \( Q' = 0 \) in a domain. Geostrophic balance is nearly always replaced by more realistic balances, like that of Charney (1955). It is common to choose the complementary PPVI strategy in the general case by inverting first all anomalies in \( D \) and to invert next with \( Q' = 0 \) in \( D_1 \) and an “observed” anomaly in \( D_2 \). The difference of both results is thought to represent the impact of \( Q' \) in \( D_1 \) on the flow in \( D_2 \) (e.g., Ahmadi-Givi et al. 2004). All published piecewise inversions rely on numerical methods so that only continuity of the fields can be imposed at \( S \).
The approach of Davis and Emanuel (1991) is presumably the most popular in PPVI and is chosen here to discuss the uniqueness of PPVI. The vertical coordinate is the Exner function \( H = c_p (p/p_0)^{\rho/c_p} \) with constant \( p_0 \) so that \( S \) is an isobaric surface. That has the disadvantage that the location of \( S \) changes in space when PPVI is carried out. Thus, the shape of \( D_1 \) changes slightly, as does the PV field in \( D_1 \), but that has never been seen as problematic. One could avoid this problem by choosing height as a vertical coordinate. It is, however, preferable to stick to the well-known formulation of Davis and Emanuel (1991) in pressure coordinates, where

\[
Q = \frac{g R I}{p c_p} \left[ f + \nabla^2 \psi \right] \frac{\partial^2 \phi}{\partial \Pi^2} = \frac{\partial^2 \psi}{\partial x \partial \Pi} - \frac{\partial^2 \phi}{\partial y \partial \Pi} - \frac{\partial^2 \psi}{\partial y \partial \Pi} = \frac{\partial^2 \phi}{\partial y \partial \Pi}.
\]

where \( \psi \) is the streamfunction of the flow on an isobaric surface, and \( \phi \) is the geopotential. Davis and Emanuel (1991) use the balance condition of Charney (1955)

\[
\nabla^2 \psi = f \nabla^2 \psi + 2 \left( \frac{\partial^2 \psi}{\partial x \partial y} \right)^2 - \frac{\partial^2 \psi}{\partial x \partial y} \frac{\partial^2 \psi}{\partial y^2},
\]

(24)

where we have chosen an \( f \)-plane geometry to be consistent with previous sections.

In practice, a mean state must be defined, and \( Q \) in (23) must be replaced by \( Q' = Q - Q_0 \), where (24) must be rewritten accordingly. The simplest choice, \( \phi = \overline{\phi} (\Pi) \) and \( \overline{\phi} = 0 \), will be made in what follows. However, as discussed by Davis and Emanuel (1991), there is some freedom in defining the mean state.

Assume now that a specific \( Q_{obs} \) has been inverted in \( D \) with a result of \( \psi_0, \phi_0 \). Both versions of PPVI can be carried out. We may prescribe \( Q' = 0 \) in \( D_2 \) and \( Q' = Q_{obs} \) in \( D_1 \) with an inversion result of \( \psi_1, \phi_1 \). As stated, it is customary to conduct complementary PPVI by assuming \( Q' = 0 \) in \( D_1 \) and \( Q_{obs} \) in \( D_2 \) with a result of \( \psi_2, \phi_2 \). The difference \( \psi_0 - \psi_2, \phi_0 - \phi_2 \) represents the impact of \( Q_{obs} \) in \( D_1 \) on the flow in \( D_2 \). However, while \( \psi_0 - \psi_2 = \psi_1 \), \( \phi_0 - \phi_2 = \phi_1 \) in linear PPVI, there is no guarantee that this simple relation is also correct in nonlinear PPVI.

Inserting \( \psi_0 = \psi_1 + \psi_2 \) and \( \phi_0 = \phi_1 + \phi_2 \) in (23), we find the residuum in \( D_1 \)

\[
Q_1^* = \frac{g R T}{p c_p} \left( \nabla^2 \psi_1 \frac{\partial^2 \phi_1}{\partial \Pi^2} + \nabla^2 \psi_2 \frac{\partial^2 \phi_2}{\partial \Pi^2} - \frac{\partial^2 \psi_1}{\partial x \partial \Pi} \frac{\partial^2 \phi_2}{\partial y \partial \Pi} - \frac{\partial^2 \psi_2}{\partial y \partial \Pi} \frac{\partial^2 \phi_1}{\partial x \partial \Pi} \right).
\]

(25)

where we took into account that \( \psi_1, \phi_1 \) have the PV \( Q_1 \) and that the PV of \( \psi_2, \phi_2 \) vanishes in \( D_1 \). Moreover, (24) is not satisfied by \( \psi_1 + \psi_2 \). Instead, the nonlinear term in (24) does not vanish and has to be satisfied by an additional geopotential perturbation. In other words, \( \psi_1 + \psi_2 \neq \psi_0 \), and attribution fails. After all, \( \psi_1 (\psi_2) \) is the streamfunction induced by \( Q_1 (Q_2) \), but the sum \( Q = Q_1 + Q_2 \) does not induce the sum of the streamfunctions induced by \( Q_1 \) and \( Q_2 \) in separation.

Last but not least, let us demonstrate that the type of nonuniqueness found for quasigeostrophic flows exists also in the general case. Assume that \( \psi_1 \) and \( \phi_2 \) are available. We construct now alternative solutions \( \psi_2 + \delta \phi, \phi_1 + \delta \phi \) with \( \delta \phi = \delta \phi = 0 \) in \( D_1 \). Insertion of these modified fields in (23) gives a new PV field in \( D_2 \) with \( Q' = 0 \) in \( D_1 \). We have to make sure that the modified fields satisfy the balance equation [see (24)]. Thus,

\[
\nabla^2 \delta \phi = f \nabla^2 \delta \psi + 2 \left( \frac{\partial^2 \delta \psi}{\partial x \partial y} \right)^2 + 2 \frac{\partial^2 \delta \psi}{\partial y \partial x} \frac{\partial^2 \delta \psi}{\partial x \partial y} + 2 \frac{\partial^2 \delta \psi}{\partial x^2} \frac{\partial^2 \delta \psi}{\partial y^2} - 2 \left( \frac{\partial^2 \delta \phi}{\partial x^2} \frac{\partial^2 \delta \phi}{\partial y^2} + \frac{\partial^2 \delta \phi}{\partial x \partial y} \frac{\partial^2 \delta \phi}{\partial y \partial x} \right)
\]

(26)

has to be solved in \( D_2 \). This can be achieved by assuming \( \delta \psi = \psi \), where \( \psi \) is an arbitrary function of \( x, y \), and \( \Pi \), which vanishes outside of \( D_1 \) and has \( \tau = 0 \) and \( \partial \tau/\partial n = 0 \) at \( S \). All these conditions can be satisfied, because we are free with respect to the choice of \( \tau \). Then (26) must be solved in \( D \) with \( \delta \psi \) replaced by \( \tau \). Hence, (26) is a linear equation for \( \delta \psi \) with \( \delta \phi = 0 \) at \( S \). Inserting this solution in (24) yields a new PV field in \( D_2 \) with unaltered flow in \( D_1 \). Thus, attribution is not unique with respect to \( \psi \) and \( \phi \) in this fairly general case, and many different fields of \( Q' \) in \( D_2 \) lead to the same “response” in \( D_1 \).

Vertical motion is excluded by (24), and the inverted wind does not contain a vertical circulation. Thus, uniqueness cannot be established by including the vertical flow component in the inversion. This conclusion depends on the balance condition in (24), which implies that horizontal divergence vanishes and \( \omega = 0 \) in the inverted flow.

4. Discussion and concluding remarks

PV is derived from observations of velocity \( \mathbf{v} \) and potential temperature. PVI after imposing a balance condition, and PPVI divides the total flow domain \( D \) into subdomains \( D_1 \) and \( D_2 \) (or more). The observed PV (\( Q_1 \) or \( q_1 \)) is prescribed in \( D_1 \), and a vanishing PV is assumed in \( D_2 \). The inversion of this PV field leads to velocity \( \mathbf{v}_1 \) and potential temperature \( \theta_1 \). Both fields are thought to describe the contribution of PV in \( D_1 \) to the flow in \( D \) and, in particular, to that in \( D_2 \). The complementary inversion leads to \( \mathbf{v}_2 \) and \( \theta_2 \). These
techniques can also be applied in the electrostatic case and would indeed describe the impact of the charges in $D_1$ with corresponding forces on charges in $D_2$, and attribution describes a physical effect in this situation. This differs significantly from the situation in PPVI, where PV in $D_1$ neither exerts forces on PV in $D_2$ nor on the flow and attribution does not describe a physical impact in PPVI. Instead, PPVI yields information on the flows with $Q_{\text{obs}}$ in $D_1$ and vanishing $Q$ in $D_2$. This information may be interesting but does not allow us to state an influence of PV in $D_1$ on the flow in $D_2$.

The superposition principle requires that the results of both inversions sum up to the result of PVI in $D$. This principle is strictly satisfied in linear cases such as quasigeostrophic inversion, but there is no guarantee that it holds for nonlinear cases. If superposition fails, the total observed flow cannot be restituted from the PV in $D_1$ and $D_2$ in separation and $\mathbf{v} \neq \mathbf{v}_1 + \mathbf{v}_2$.

Vertical motion is usually excluded in PPVI. It is thought to be sufficient to evaluate a streamfunction $\psi$ in quasigeostrophic PPVI and an additional geopotential in the general case [see (24)]. It is shown here that this type of PPVI is not unique in the sense that there are many PV fields in $D_1$ that lead to the same inverted flow in $D_2$. Furthermore, the inverted flow in $D_2$ does not depend on PV in $D_1$. It is sufficient to know the pressure and its normal derivative at $S$. PPVI also generates flows in $D_1$, a feature that appears not to have been discussed in the literature. This flow is also not unique and cannot be attributed to PV in $D_1$ cannot be interpreted as a physical effect of PV in $D_1$.

Uniqueness can be established when the vertical velocity $w$ is added to the variables that are to be restituted by quasigeostrophic PPVI. Inversion of $q_1$ in PPVI then leads to a result in $D_2$ that cannot be duplicated by prescribing other PV fields in $D_1$. This has been demonstrated explicitly for an example, and a more general proof has been proposed in section 2b. We chose the paradigm of PPVI with the interaction of two PV balls to investigate PPVI including $w$. Uniqueness is established indeed, but the superposition principle is not satisfied in this case, because a single PV anomaly is not associated with vertical motion. Hence, it is not possible to reconstruct the three-dimensional velocity pattern on the basis of PV in separate subdomains. One may criticize our example as dealing with highly idealized PV anomalies, and we extended our discussion to PV balls of finite size. This leads to substantial modifications of $w$, but the basic conclusion is not affected by this extension to a more realistic situation, and the details need not be presented here.

The result for nongeostrophic PPVI depends on the chosen balance conditions and so does the proof of nonuniqueness. The approach of Davis and Emanuel (1991) is used almost exclusively in PPVI, so our proof covers most of the published cases of PPVI. Nevertheless, other balance conditions (e.g., Herbert 1971; Mohebalhojeh and McIntyre 2007) may require new proofs.

Piecewise prognostic inversion (Hakim et al. 1996; Teubler and Riemer 2016) goes one step beyond “static” PPVI, where, after performing PPVI, the advection of PV by the geostrophic wind attributed to PV in $D_1$ is used to calculate tendencies of the geopotential in $D$, which are attributed to PV in $D_1$. The problem with this technique is the same as with PPVI. For example, the geostrophic winds in $D_2$ can be derived from pressures at $S$, but they are not directly affected by the PV in $D_1$. Thus, the advection of PV by the wind in $D_2$, which is attributed to the PV in $D_1$, is not unique.

Our findings ask for a more cautious interpretation of PPVI because of the nonuniqueness of attribution. Although the inclusion of the ageostrophic circulation leads to uniqueness in the quasigeostrophic case, superposition is not satisfied. Moreover, PV does not exert a force, and PPVI does not inform us about a physical impact of PV in one domain on the flow in another. Instead, PPVI provides the associated flows with the modified PV. This is interesting information in itself, but the evolution of these flows in time is needed to establish causality. Spengler and Egger (2012) looked into this problem for a spherical PV anomaly and calculated the evolving flows. This approach could be extended to investigate the reaction to more realistic initial fields of PV. In particular, our results on nonuniqueness can be exploited by specifying various PV fields in $D_1$ modified such that the results of PPVI are the same in $D_2$. Such experiments would shed new light on the impact of isolated PV anomalies in time.

Acknowledgments. We thank two anonymous reviewers for their comments. In particular, the constructive and extended comments by one referee helped greatly to clarify some issues and improve the paper.

REFERENCES


