QUARTERLY JOURNAL
OF THE
ROYAL METEOROLOGICAL SOCIETY

Vol. 111 OCTOBER 1985 No. 470

Quart. J. R. Met. Soc. (1985), **111**, pp. 877–946 551.509.3:551.511.2:551.511.32

On the use and significance of isentropic potential vorticity maps

By B. J. HOSKINS[1], M. E. McINTYRE[2] and A. W. ROBERTSON[3]

[1] *Department of Meteorology, University of Reading*
[2] *Department of Applied Mathematics and Theoretical Physics, University of Cambridge*
[3] *Laboratoire de Physique et Chimie Marines, Université Pierre et Marie Curie, 75230 Paris Cédex 05*

(Received 12 February 1985; revised 2 July 1985)

SUMMARY

The two main principles underlying the use of isentropic maps of potential vorticity to represent dynamical processes in the atmosphere are reviewed, including the extension of those principles to take the lower boundary condition into account. The first is the familiar Lagrangian conservation principle, for potential vorticity (PV) and potential temperature, which holds approximately when advective processes dominate frictional and diabatic ones. The second is the principle of 'invertibility' of the PV distribution, which holds whether or not diabatic and frictional processes are important. The invertibility principle states that if the total mass under each isentropic surface is specified, then a knowledge of the global distribution of PV on each isentropic surface and of potential temperature at the lower boundary (which within certain limitations can be considered to be part of the PV distribution) is sufficient to deduce, diagnostically, all the other dynamical fields, such as winds, temperatures, geopotential heights, static stabilities, and vertical velocities, under a suitable balance condition. The statement that vertical velocities can be deduced is related to the well-known omega equation principle, and depends on having sufficient information about diabatic and frictional processes. Quasi-geostrophic, semi-geostrophic, and 'nonlinear normal mode initialization' realizations of the balance condition are discussed. An important constraint on the mass-weighted integral of PV over a material volume and on its possible diabatic and frictional change is noted.

Some basic examples are given, both from operational weather analyses and from idealized theoretical models, to illustrate the insights that can be gained from this approach and to indicate its relation to classical synoptic and air-mass concepts. Included are discussions of (a) the structure, origin and persistence of cutoff cyclones and blocking anticyclones, (b) the physical mechanisms of Rossby wave propagation, baroclinic instability, and barotropic instability, and (c) the spatially and temporally nonuniform way in which such waves and instabilities may become strongly nonlinear, as in an occluding cyclone or in the formation of an upper air shear line. Connections with principles derived from synoptic experience are indicated, such as the 'PVA rule' concerning positive vorticity advection on upper air charts, and the role of disturbances of upper air origin, in combination with low-level warm advection, in triggering latent heat release to produce explosive cyclonic development. In all cases it is found that time sequences of isentropic potential vorticity and surface potential temperature charts—which succinctly summarize the combined effects of vorticity advection, thermal advection, and vertical motion without requiring explicit knowledge of the vertical motion field—lead to a very clear and complete picture of the dynamics. This picture is remarkably simple in many cases of real meteorological interest. It involves, in principle, no sacrifices in quantitative accuracy beyond what is inherent in the *concept* of balance, as used for instance in the initialization of numerical weather forecasts.

CONTENTS

1. INTRODUCTION AND HISTORICAL REVIEW
   1(*a*) Early ideas
   1(*b*) Rossby and Ertel
   1(*c*) Subsequent developments
   1(*d*) The invertibility principle for potential vorticity
2. ISENTROPIC POTENTIAL VORTICITY MAPS FROM ROUTINE ANALYSES
   2(*a*) Preliminaries

2(*b*) Vertical structure and time-variability
2(*c*) Development of a North Atlantic cutoff cyclone
2(*d*) A minor blocking episode
2(*e*) The conceptual duality between cutoff cyclones and blocking anticyclones
3. SOME SIMPLE EXAMPLES, FOLLOWING KLEINSCHMIDT
4. ON THE CANCELLATION OF HORIZONTAL ADVECTION BY VERTICAL MOTION
5. ANOMALIES AT THE LOWER BOUNDARY, AND THE INVERTIBILITY PRINCIPLE FOR GENERAL, TIME-DEPENDENT FLOW
5(*a*) Surface and near-surface anomalies
5(*b*) Quasi-geostrophic theory
5(*c*) Semi-geostrophic theory and Salmon's generalization
5(*d*) Inversion by nonlinear normal-mode initialization
6. ROSSBY WAVES AND SHEAR INSTABILITIES
6(*a*) Rossby wave propagation and the scale effect
6(*b*) Baroclinic and barotropic shear instabilities
6(*c*) Lateral and vertical Rossby wave propagation
6(*d*) The nonlinear saturation of baroclinic instabilities
6(*e*) Further remarks about cyclogenesis in the real atmosphere
7. THE MAINTENANCE AND DISSIPATION OF CUTOFF CYCLONES AND BLOCKING ANTICYCLONES
8. FURTHER REMARKS ABOUT CUTOFF SYSTEMS AND AIR MASSES
9. CONCLUDING REMARKS

ACKNOWLEDGEMENTS

APPENDIX: THE COMPUTATION OF VERTICAL MOTION

1. INTRODUCTION AND HISTORICAL REVIEW

(*a*) *Early ideas*

Circulation and vorticity have been recognized as fundamental concepts in meteorology and oceanography for many years, dating back to the pioneering work of V. Bjerknes (1898a, b, 1901, 1902); see also, e.g., Eliassen (1982). The three-dimensional vorticity equation, as it appears in textbooks on general fluid dynamics, may be written for frictionless motion relative to a coordinate system rotating with angular velocity $\boldsymbol{\Omega}$ as

$$\frac{D}{Dt}(\boldsymbol{\zeta}_a/\rho) = (\boldsymbol{\zeta}_a/\rho) \cdot \nabla \mathbf{u} - (1/\rho)\nabla(1/\rho) \times \nabla p \tag{1}$$

where the absolute vorticity

$$\boldsymbol{\zeta}_a = 2\boldsymbol{\Omega} + \boldsymbol{\zeta} \tag{2}$$

and

$$\boldsymbol{\zeta} = \nabla \times \mathbf{u}, \tag{3}$$

the relative vorticity. D/Dt is the material rate of change, and ∇ is the three-dimensional gradient operator. We denote the three-dimensional velocity by $\mathbf{u} = (u, v, w)$ to distinguish it from the horizontal wind velocity $\mathbf{v} = (u, v, 0)$. The first term on the right-hand side of (1) is the stretching–twisting term, and the second the so-called solenoid term.

The absolute circulation around a material circuit Γ moving with the fluid is

$$C_a = C + 2\Omega A, \tag{4}$$

where

$$C = \oint_\Gamma \mathbf{u} \cdot d\mathbf{l} \tag{5}$$

is the relative circulation and A is the area bounded by a projection of the circuit onto a plane normal to $\boldsymbol{\Omega}$. Circulation is another measure of the rotational character of the air motion (which is the aspect of the air motion usually of interest in dynamical meteorology) and is equal to the integral of the vorticity over a surface bounded by the circuit. Following Bjerknes, we may write the circulation theorem for frictionless motion as

$$dC_a/dt = -\oint_\Gamma (1/\rho)\, dp. \tag{6}$$

The stretching–twisting term, the first term on the right of (1), has been absorbed into the behaviour of the material circuit and does not appear explicitly in (6).

When the usual meteorological approximations are made, i.e. neglecting vertical accelerations and the horizontal component of the rotation $\boldsymbol{\Omega}$, taking the geoid to be spherical, and replacing the distance of an air parcel from the centre of the earth by a constant representative value of the earth's radius, the equations are unaltered except that

(i) where $\mathbf{u}$ appears explicitly in (3) and (5) it is replaced by the horizontal wind vector $\mathbf{v}$;
(ii) only the vertical component f of $2\boldsymbol{\Omega}$ is used in (2) and (4); and
(iii) the plane involved in the definition of the projected area A in (4) is horizontal, rather than perpendicular to $\boldsymbol{\Omega}$.

(*b*) *Rossby and Ertel*

In general, the complexity of the foregoing equations means that detailed argument from them is difficult. Rossby (1939) took a key step by realizing that in practice the vertical component of absolute vorticity $\zeta_a = f + \mathbf{k} \cdot (\nabla \times \mathbf{v})$, $\mathbf{k}$ being a unit vertical vector, is the most important for the large-scale atmospheric flow. He realized furthermore that many features of the flow could be surprisingly well modelled by assuming conservation of ζ_a in two-dimensional horizontal motion—the familiar barotropic model of large-scale atmospheric dynamics. The streamfunction for this flow is obtainable at any time by inversion of the Laplacian operator linking it and the vorticity; we shall refer to this as the 'invertibility principle' for vorticity in the barotropic model. Rossby's insight, which clarified and simplified some earlier insights from what Rossby referred to as "a remarkable paper by J. Bjerknes" (1937), led to a number of important developments, for instance the theory of Rossby wave motion, the use of constant absolute vorticity trajectories, and, a decade later, the conception and execution of the first practicable numerical forecasting models.

Rossby (1940) took a further key step by noticing that if h is the depth of a material fluid column in the barotropic model, then

$$\zeta_a/h = \text{constant} \tag{7}$$

following the fluid column. This describes, in an ingeniously simple way, the two processes that often dominate the vorticity budget, namely the creation of vorticity by vertical stretching of vortex tubes, and the horizontal advection of absolute vorticity. This is the simplest version of the modern concept of 'potential vorticity'.

In the late 1930s, Rossby and his co-workers (e.g. Rossby 1937a, b; Namias

1940; for some more history see Namias 1983) had made systematic use of isentropic charts for displaying the motion of the atmosphere and had recognized their power in elucidating the Lagrangian, air-mass structure of the large-scale flow given that diabatic processes were of secondary importance. The use of isentropic charts had previously been advocated by Shaw (1930). It was therefore natural for Rossby to move straight from Eq. (7), the conservation of ζ_a/h, to the proof of a similar result for an adiabatic, frictionless atmosphere consisting of a finite number of layers of constant potential temperature, θ. Making the hydrostatic approximation, and expressing the layer thickness $\Delta = -\delta p/g$ as the mass per horizontally projected area, he showed (in the same paper, 1940) that, for each such layer,

$$P_R = (f + \zeta_\theta)/\Delta = \text{constant} \tag{8}$$

following an air parcel. Here ζ_θ is the relative 'isentropic vorticity'

$$\zeta_\theta = \mathbf{k} \,.\, \nabla_\theta \times \mathbf{v} = (\partial v/\partial x)_\theta - (\partial u/\partial y)_\theta \tag{9}$$

where, as before, (u, v) are the horizontal wind components and dx, dy are horizontal increments of distance, but differentiations are carried out on the isentropic surface $\theta =$ constant (*op. cit.*, p. 72). This is another ingenious mathematical device of great simplifying power. Rossby pointed out that "it is possible to derive corresponding results also for an atmosphere in which the potential temperature varies continuously with elevation", and introduced the term 'potential vorticity' for the value of ζ_θ which the isentropic fluid layer would have if it were moved to a standard latitude and its mass per unit area, Δ, brought to a standard value. The term 'potential vorticity' is today used for quantities like (7) and (8) instead, but the underlying concept is the same, namely the idea that there is a potential for creating vorticity by changing latitude and by adiabatically changing the separation of isentropic layers.

The Bjerknes circulation theorem (6) gives a simple alternative derivation of the result (8) obtained by Rossby. When (6) is applied to any material circuit Γ in an isentropic surface, ρ is a function of p only so that C_a must be constant for frictionless, adiabatic motion. However, integration of P_R over the mass of an isentropic layer bounded by any material circuit Γ gives

$$\iint P_R \Delta \, dA = \iint (f + \zeta_\theta) \, dA = C_a = \text{constant}, \tag{10}$$

where dA is the horizontal projection of the area element of integration, so that $\Delta \, dA$ is the mass element. The third step applies Stokes' theorem in two dimensions to the horizontal projection of the circuit Γ, giving the appropriate versions of (4) and (5) under the modifications (i)–(iii). Since (10) is true for any material circuit Γ, and the mass $\iint \Delta \, dA$ is constant, P_R must itself be conserved following the fluid.

Rossby's result (8) was given full hydrodynamical generality by the independent work of Ertel (1942). If θ is a function of the state variables p and ρ alone, then the three-dimensional scalar product of $\nabla\theta$ with the solenoid term in the vorticity equation (1) is zero. If, moreover,

$$\theta = \text{constant} \tag{11}$$

for an air parcel, signifying adiabatic motion, then, expressing this as $D\theta/Dt = 0$ and taking its three-dimensional gradient, and then taking the scalar product of $\nabla\theta$ with the frictionless vorticity equation (1), we obtain after some manipulation the celebrated result

$$P = \rho^{-1} \boldsymbol{\zeta}_a \,.\, \nabla\theta = \text{constant} \tag{12}$$

following an air parcel. This is Ertel's theorem for adiabatic, frictionless motion. Remarkably, no approximations are involved; P is conserved even for fully three-dimensional, nonhydrostatic motion. The generalization to include diabatic and frictional effects is straightforward; it is noted in section 7 below, together with an important integral constraint on those effects, arising from the fact that quantities like $\boldsymbol{\zeta}_a \cdot \nabla\theta$ can always be expressed as exact divergences $\nabla \cdot (\boldsymbol{\zeta}_a\theta)$.

Normally in meteorology, θ is taken to be the potential temperature, although it can equally well be taken to be the specific entropy or any other function of potential temperature. If the meteorological approximations are made, including the hydrostatic approximation, the isobaric and isentropic coordinate versions of the expression in (12) become

$$P = -g(f\mathbf{k} + \nabla_p \times \mathbf{v}) \cdot \nabla_p\theta \tag{13}$$

and

$$P = -g(f + \mathbf{k} \cdot \nabla_\theta \times \mathbf{v})/(\partial p/\partial\theta), \tag{14}$$

respectively, where $\mathbf{k}$ is again a unit vertical vector and ∇_p and ∇_θ are the three-dimensional gradient operators in xyp and $xy\theta$ space respectively. Note that the second expression, involving once more the 'isentropic vorticity' $\zeta_\theta = \mathbf{k} \cdot \nabla_\theta \times \mathbf{v}$ defined in (9), is much simpler than the first, where there is a scalar product which expands to three terms. From comparison of (14) and (8) it is apparent that what is now known as 'Ertel's potential vorticity', P, as it is used in meteorology, is merely the continuous-atmosphere extension of (8) to which Rossby referred. We shall simply call P 'the potential vorticity', or PV for short.

(c) *Subsequent developments*

Rossby's co-workers quickly seized on the importance of the PV as an air-mass tracer. Starr and Neiburger (1940) noted that, for adiabatic, frictionless flow, it provided the third Lagrangian marker necessary to identify an air parcel, the other two being potential temperature and specific humidity. They constructed isentropic potential vorticity maps ('IPV maps') for the 299 K–303 K isentropic layer over North America for 21 and 22 November 1939, and checked for three fluid elements that the conservation of PV was approximately satisfied in the free atmosphere. They contemplated possible future investigations in which the PV might be used for studying non-adiabatic processes, mixing, and even "prediction of the upper air winds" using the fact that "the transport of potential vorticity affects the wind-field". The correlation between specific humidity and potential vorticity was given further study by Spar (1943).

The further application of this new tool to elucidate the Lagrangian behaviour of the atmosphere appears to have proceeded at a slower rate. Platzman (1949) used data which Palmén and Newton (1948) had composited relative to the polar front at 80°W, to demonstrate a salient feature of the isentropic PV distribution for the 310 K–350 K layer. He found relatively uniform high PV values on the polar, stratospheric side of the jet stream, and uniform low values on the equatorward side, with a near discontinuity at the jet stream. He then proposed this as a suitable basic state to consider in a barotropic model.

Kleinschmidt (1950a, b; 1951, 1955, 1957) used the notion of PV anomalies in the upper troposphere to explain observed cyclogenesis events, noting among other things the likely importance of quasi-horizontal advection along isentropes from the stratospheric 'reservoir' of high-PV polar air found by Platzman (1949). Seen in the light of more recent developments, Kleinschmidt's work must be classed as an extremely important piece of pioneering, if only because for the first time he placed full emphasis on what we

shall call the 'invertibility principle' for potential vorticity, namely the idea that not only can one use the PV as a Lagrangian tracer but that one can also deduce, diagnostically, the complete flow structure *from* the spatial distribution of PV. His ideas were, however, too far ahead of theoretical developments and were largely displaced, at the time, by the ideas associated with the new baroclinic instability theory with its emphasis on lower tropospheric baroclinicity. Ironically, we shall see that a more complete picture of what goes on in the atmosphere actually involves *both* ideas, and in particular that the invertibility principle gives the key to one of the simplest ways of thinking physically about the baroclinic instability process itself.

The next significant line of development was initiated in the important papers by Reed and Sanders (1953) and Reed (1955), which began the process of mapping out in detail the upper air PV structures whose gross features Platzman and Kleinschmidt had seen earlier. Useful historical reviews are given by Shapiro (1976) and by Keyser and Shapiro (1985). Reed and Sanders used PV values to determine which air was of stratospheric origin in a detailed cross-section of an upper tropospheric front. In a further study of upper tropospheric frontogenesis Reed (1955) produced IPV maps like those of Starr and Neiburger, but locally with far better resolution, for the 300 K isentropic surface over North America daily from 13 to 15 December 1953. Both papers strongly suggested a remarkable fact, to be amply confirmed by later studies, namely that stratospheric air could be advected down sloping isentropic surfaces to altitudes usually considered to be well within the troposphere. The suggested picture was confirmed in greater detail by Reed and Danielsen (1959), who gave the first convincing local picture of the associated 'tropopause folding' process in vertical cross-sections. Further confirmation, including exquisitely detailed pictures using special aircraft data, has been recorded in papers by, for instance, Mahlman (1965), Reiter and Mahlman (1965), Danielsen (1968), Danielsen *et al.* (1970) and Shapiro (1974, 1976, 1978), to name only a few. Staley (1960) sounded a note of caution by finding examples in which PV was apparently not well enough conserved to use as a tracer. However, Danielsen (1968), using simultaneous measurements of ozone and nuclear test radioactivity as well as calculations of potential vorticity, was able to make a strong case that the PV, even in cases where it was changed by sub-synoptic-scale mixing or convection, generally provided an extremely good indication of air of recent stratospheric origin, and that a definition of the 'tropopause' in terms of PV is more useful than a lapse-rate definition. Implicit in this and other practical uses of the PV is a presumption that a certain amount of fine-grain structure in its horizontal and vertical distribution can be ignored for practical purposes. As we shall indicate shortly, some theoretical justification can be found for this. Observational evidence for such structure, especially that due to fine structure in static stability profiles, and for its local generation by, for instance, clear air turbulence, is plentiful (e.g. Danielsen 1959; Shapiro 1976, 1978; Vaughan and Tuck 1985).

Another historical landmark was the first publication of computer-generated IPV maps giving a global-scale, coarse-grain view, by Obukhov (1964) and Danielsen (1967, 1968) for the 300, 305 and 310 K surfaces. In subsequent years, IPV maps have been increasingly used in the diagnosis of observed atmospheric behaviour (e.g. Danielsen *et al.* 1970; Bleck 1973; Holopainen and Rontu 1981; McIntyre and Palmer 1983, 1984; Bleck and Mattocks 1984; Uccellini *et al.* 1985; Al-Ajmi *et al.* 1985; Clough *et al.* 1985; Shutts 1985; Young *et al.* 1985), in the diagnosis of atmospheric model simulations (e.g. Hsu 1980; Dunkerton *et al.* 1981), and in studies of oceanic circulations (McDowell *et al.* 1982; Sarmiento *et al.* 1982; Holland *et al.* 1984; Woods *et al.* 1985). For a recent review of related developments in the Soviet Union, the reader may consult Obukhov (1984).

(*d*) *The invertibility principle for potential vorticity*

As Rossby, Starr and Neiburger, and Kleinschmidt had evidently appreciated, the significance of potential vorticity does not end with its importance as an air-mass tracer. It is also the key to a very powerful and succinct view of the *dynamics*. IPV maps, in particular, are a natural diagnostic tool well suited to making dynamical processes directly visible to the human eye and to making meaningful comparisons between atmospheric models and reality. We shall see that IPV maps can play a role in improving our understanding of dynamical processes closely analogous to the role played by maps of the absolute vorticity ζ_a in understanding the behaviour of barotropic models.

Two properties of P and ζ_a underlie these statements. First, there are the Lagrangian conservation properties already noted and, where relevant, their frictional, diabatic generalizations. Second, there is the invertibility principle. For the barotropic model, this is the familiar fact that, given the vorticity distribution, one can deduce the streamfunction, and thence the wind field, the calculation involving the inversion of a Laplacian operator as mentioned earlier. Thus one can, if one wishes, *think* entirely in terms of the (barotropic) vorticity field, since this contains all the relevant information; and indeed the simplifying power and general usefulness of this particular mode of thinking about rotational fluid motion has long been recognized, and made routine use of, in another field, that of classical aerodynamics (e.g. Prandtl and Tietjens 1931; Goldstein 1938; Lighthill 1963; Batchelor 1967; Saffman 1981).

As Kleinschmidt seems to have realized, at least intuitively, a similar invertibility principle must hold for the potential vorticity field in baroclinic flow: given the PV distribution, one can deduce the wind, pressure and temperature fields. Kleinschmidt's early attempts to express this idea mathematically were successful as far as they went, but they were tied to particular examples and did not establish procedures suitable for general use either computationally or conceptually. Seeing in a general way that such procedures exist was a far less trivial problem than in the barotropic case, and had to await later theoretical developments, one of which was quasi-geostrophic theory in the form developed by Charney and Stern (1962). More refined and powerful versions of the theory are still being developed today; they will be discussed in section 5.

It might be asked (and this may have been part of why Kleinschmidt's ideas were not widely accepted) how knowledge of the potential vorticity, i.e. of the *product* of absolute vorticity and static stability as expressed by (14), can possibly determine the two factors separately—for that is what the invertibility principle asserts. The answer lies in the fact that we must of course specify some more information—a fact so familiar to theoretical specialists today that it is not always spelt out explicitly. Regardless of which particular theoretical device is used for the inversion, one must

(i) specify some kind of balance condition, the simplest albeit least accurate option being ordinary geostrophic balance;
(ii) specify some kind of reference state, expressing the mass distribution of θ, in essentially the same way as in Lorenz's theory of available potential energy; and
(iii) solve the inversion problem *globally*, with proper attention to boundary conditions.

The last point comes as no surprise, of course, because it is equally true of the barotropic case. The first and second points imply that if we wish to be precise we must speak of inversion under this or that balance condition, and relative to this or that reference state. It is only when all these constraints operate in the problem that the arbitrariness is removed. In other words, while it is perfectly true that a purely local knowledge of P cannot determine the local absolute vorticity and static stability separately, there is only one pair of values of them which will fit into a given *global* distribution

of P in which thermal wind balance (or some more accurate balance) is satisfied everywhere.

There are, of course, two well-known provisos to all this. First, the balance referred to in condition (i) must be a physically realizable and therefore stable balance in the sense that static instability, inertial instability, and the related parcel instabilities usually studied under the heading 'symmetric baroclinic instability', are all assumed to be absent. Second, the space and time scales of the motion must be compatible with the assumed balance; in practice this will often mean that the invertibility principle will apply most accurately to motions of synoptic scale upwards, although it may also apply to certain mesoscale motions.

It might well be asked whether it is possible to talk about any scale separation at all, in this connection, since there are many observational and theoretical reasons for supposing that actual IPV distributions may have significant fine structure on scales even smaller than the mesoscale. Indeed it can be argued that there will be no lower limit to the fineness of the structures that may occur, all the way down to the length scales on which molecular diffusion acts. It is here that integral relations like relation (10) between PV and absolute circulation are of fundamental importance. The relation suggests that, for the purpose of applying the invertibility principle, it should be meaningful to integrate over the fine structure on each isentropic surface and think in terms of 'coarse-grain' IPV distributions expressing the absolute circulation around all isentropic circuits of resolvable size, with reference to a vertically smoothed static stability. Indeed, if some such coarse-grain approximation to actual IPV distributions were *not* dynamically meaningful, "numerical model simulations of the large-scale behaviour of the atmosphere would hardly be practicable" (McIntyre and Palmer 1983). The same could be said of any other attempt at understanding and predicting the evolution of weather patterns on the basis of what can be resolved observationally. The integral constraint noted in section 7 is similarly relevant to devising self-consistent ways of ignoring fine structure in the PV field.

Section 3 below gives some simple examples of the static stability and wind fields associated with—or, as an aerodynamicist would put it, 'induced by' or, as Kleinschmidt originally said, "produced by"—a given potential vorticity distribution. The examples convey an intuitive idea of how the PV partitions itself between static stability and absolute vorticity. As in Kleinschmidt's work, invertibility is expressed, in section 3, in terms of an elliptic operator not unlike that occurring in the familiar omega equation (e.g. Eliassen 1984), and having the qualitative character of a three-dimensional Laplacian. The well-known fact that such operators are *smoothing* operators provides another way of seeing the likely dynamical relevance of coarse-grain approximations to IPV distributions. Even if it were practicable to retain all fine-grain structure, the inversion operation itself would generally speaking tend to be insensitive to the details of that structure.

Once one has an idea of the typical wind, temperature and pressure fields associated with a given IPV distribution, one can begin to use IPV maps not only as a means of studying tracer-transport processes, but also as a means of gaining a very direct insight into the concomitant dynamical processes, in much the same way as aerodynamicists have long used 'vorticity thinking', as already mentioned, to study the dynamical processes with which they are concerned. One such process, which is of meteorological as well as aerodynamical interest, is that of barotropic shear instability. The analogous use of 'IPV thinking' to gain insight into *baroclinic* shear instability will be noted in section 6. Another aerodynamical phenomenon in whose baroclinic counterpart we shall be interested, for reasons to emerge in section 8, is that of 'vortex rollup'.

The quantitative precision of IPV maps as a complete representation of the dynamics

is limited, in principle, by one thing only, namely the limits of accuracy inherent in condition (i) above. Related questions arise in the context of initialization procedures for numerical weather forecasting; for some interesting discussion bearing on these questions, the reader may consult, for instance, the paper by Leith (1980) and the comprehensive review by Daley (1981). As far as large-scale motions are concerned, the main limitation of IPV maps is their inability to represent equatorial Kelvin modes in the tropics. On smaller scales, other limitations are likely to be encountered in fast processes like the 'collapse' stage of surface frontogenesis, in which the motion may develop a spontaneous imbalance leading to a strong local coupling to gravity modes (Ley and Peltier 1978; Blumen 1980; Ogura and Portis 1982; Cullen and Purser 1984), and similarly in other highly transient situations especially if very rapid advection is involved (e.g. Uccellini *et al.* 1984). It is to be hoped that current research may soon lead to a better understanding of these limitations.

2. Isentropic potential vorticity maps from routine analyses

(*a*) *Preliminaries*

Daily northern hemispheric IPV maps have been constructed for numerous isentropic surfaces for the 42 days from 20 September to 31 October 1982. The data source is the routine analysed and initialized product available four times a day from the European Centre for Medium Range Weather Forecasts (ECMWF). The actual data set employed in this study comprises the 12 z analyses on a 3·75° horizontal grid and at 11 pressure levels (1000, 850, 700, 500, 400, 300, 250, 200, 150, 100, 50 mb). Simple centred finite differences were used to construct P objectively on isobaric surfaces from Eq. (13) and then linear interpolation to give $\mathbf{v}$ and P on θ surfaces. Some results will be presented for the first 18 days analysed. It should be kept in mind that the maps can at best be coarse-grain approximations to the real IPV distributions and indeed may well fall short of the best possible such coarse-grain estimates because of the fact that the raw data were analysed isobarically rather than isentropically. Superadiabatic lapse rates were encountered in the data in small regions in the subtropics where the boundary layer was extremely warm, but these boundary layer values were ignored and the PV set to zero in what was presumed (for large-scale dynamical purposes) to be effectively an adiabatic layer.

To get a feel for the numerical values, it is useful to consider first the standard potential vorticity

$$P_s = -fg\,\partial\theta/\partial p = \rho^{-1} f\,\partial\theta/\partial z \tag{15}$$

associated with idealized, standard-atmosphere profiles at various latitudes, for a fictitious atmosphere at rest. Figure 1 shows P_s for standard-atmosphere profiles at 15°N, 45°N and 75°N. Apart from the polar boundary layer and the top of the trade wind layer, the tropospheric values on the left show a latitudinal variation which is essentially that of f, and a vertical variation attributable mainly to the factor ρ^{-1} on the right of (15). Note that for $f = 10^{-4}$ s^{-1}, a unit of 10^{-6} m^2 s^{-1} K kg^{-1} corresponds approximately to a 10 K change in θ per 100 mb. The unit of 10^{-6} m^2 s^{-1} K kg^{-1} is convenient in practice and we shall call it a 'potential vorticity unit', 'PV unit', or simply 'unit' for short. We see from Fig. 1 that tropospheric values are generally below about 1·5 units. At the tropopause P_s jumps to stratospheric values typically in excess of four units, and then rises steeply with height. The crosses in Fig. 1 mark points at which $\theta = 350$ K, and show how the 350 K isentropic surface, though varying little in height, generally samples stratospheric PV values in middle and high latitudes but tropospheric values in the subtropics.

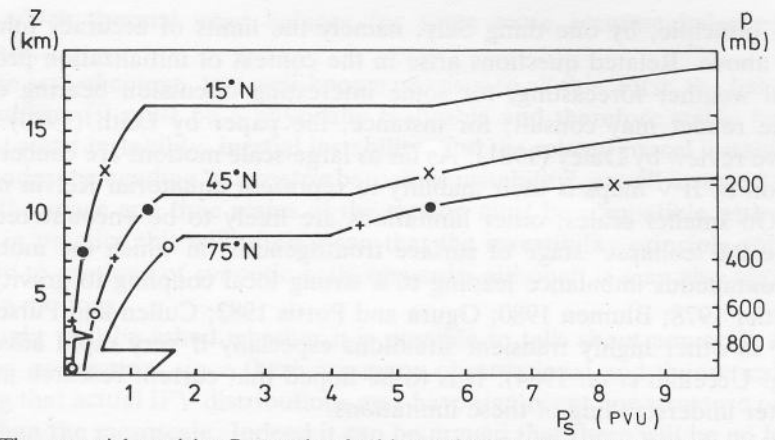


Figure 1. The potential vorticity P_s associated with standard atmospheric temperature profiles and zero relative vorticity as defined in (15). The profiles are the U.S. Standard Atmosphere 15°N annual, mid-latitude spring/autumn (labelled 45°N) and 75°N January (cold). The unit of the abscissa (1 PVU) is 10^{-6}m^{2}s^{-1}K kg^{-1}. At each latitude, P_s is discontinuous at the tropopause. The positions of various isentropic surfaces are indicated by × (350 K), ● (330 K), + (315 K) and ○ (300 K).

The 330 K isentropic surface (dots) and the 315 K surface (pluses) also sample both stratospheric and tropospheric PV values, intersecting the tropopause at different latitudes but in a manner broadly consistent with the overall picture for the 310–350 K layer originally found by Platzman (1949). The 300 K surface (open circles) is generally in the upper troposphere even in polar regions, descending to sea level in the tropics.

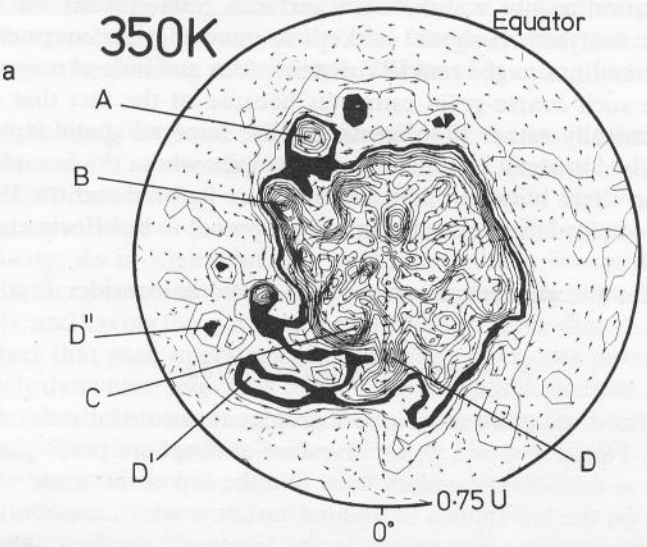
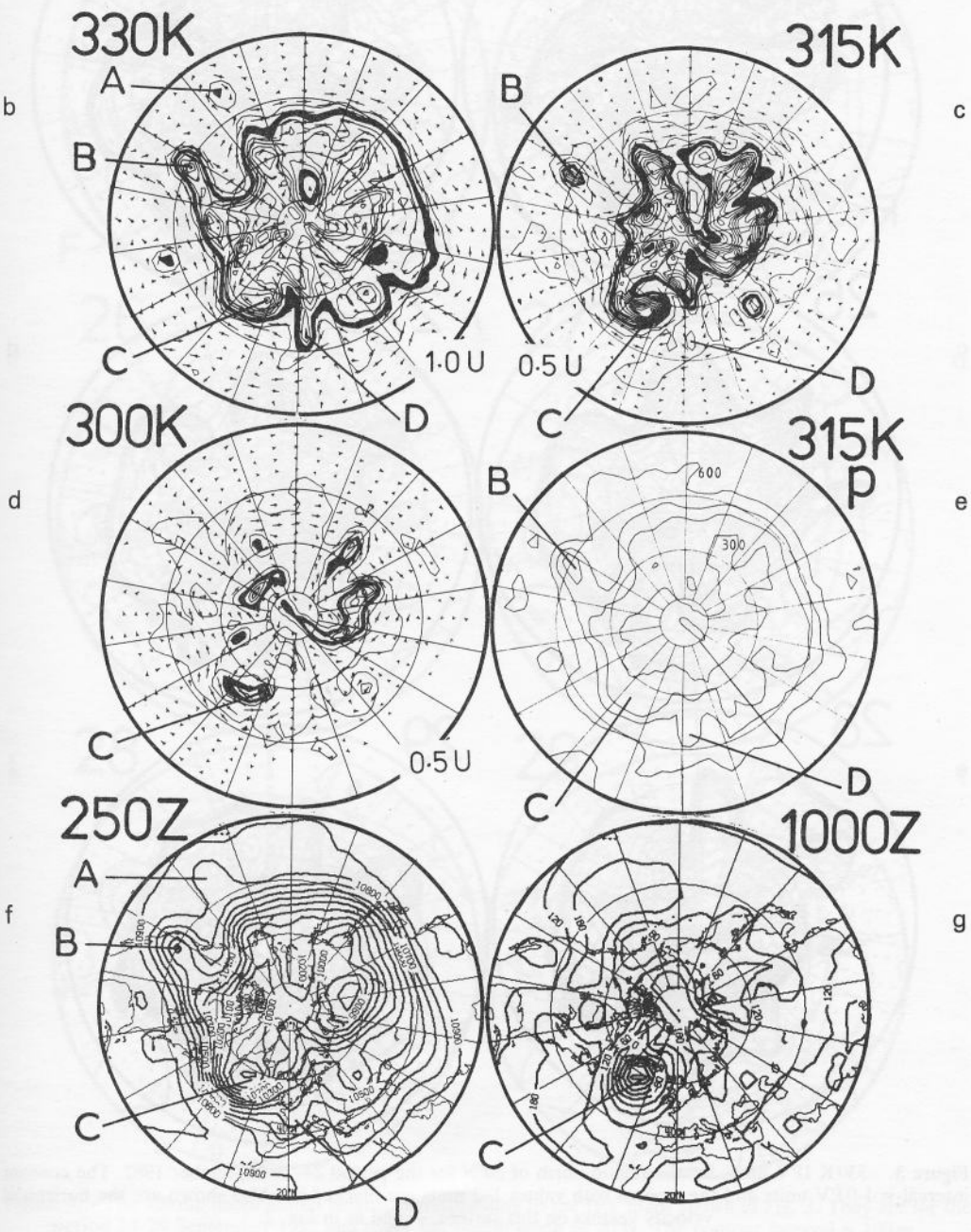


Figure 2. Northern hemispheric synoptic fields for the 30 September 1982. Shown in (a) above, (b) opposite top left, (c) opposite top right and (d) are the IPV maps for the 350 K, 330 K, 315 K and 300 K isentropic surfaces, respectively. As indicated, the contour intervals are 0·75, 1·0, 0·5 and 0·5 PV units ('U') respectively. The regions with values 2·25–3 units in (a), 2–3 units in (b) and 1·5–2 units in (c) and (d) have been blacked in. Also shown in these figures are arrows indicating the horizontal component of the velocity vector on these isentropic surfaces. The points of the arrowheads are plotted at the mid-points of the arrows which also mark the grid-points to which the vectors refer. An arrow length from the 40°–60° latitudinal circles would indicate a speed of 100 m s^{-1}. The boundary circle in (a) is the equator, but it is 20°N in all the other maps. The contours in (e) are of the pressure p on the 315 K surface, the contour interval being 100 mb. The 250 mb and 1000 mb geopotential height fields are shown in (f) and (g), with contour intervals of 100 m and 60 m respectively. The lettering denotes features referred to in the text.

(b) Vertical structure and time-variability

To give an idea of how coarse-grain IPV distributions vary from one isentropic surface to another in the real atmosphere, IPV maps for the 350 K, 330 K, 315 K and 300 K surfaces for 30 September 1982 are presented in Figs. 2(a)–(d). Also shown are the 315 K pressure contours (Fig. 2(e)) and the 250 mb and 1000 mb geopotential height fields (Figs. 2(f), (g)). It should be noted that for compactness of presentation all the fields *except* the 350 K IPV maps are restricted to the domain north of 20°N. In the PV units used above, the contour intervals for Figs. 2(a), (b), (c), (d) are respectively 0·75,



1·0, 0·5 and 0·5 units. The contours representing values from 1·5 to 4 units tend to bunch together conspicuously at the edge of the stratospheric air, indicating that the data analysis, while not capturing the almost discontinuous change at the tropopause seen in

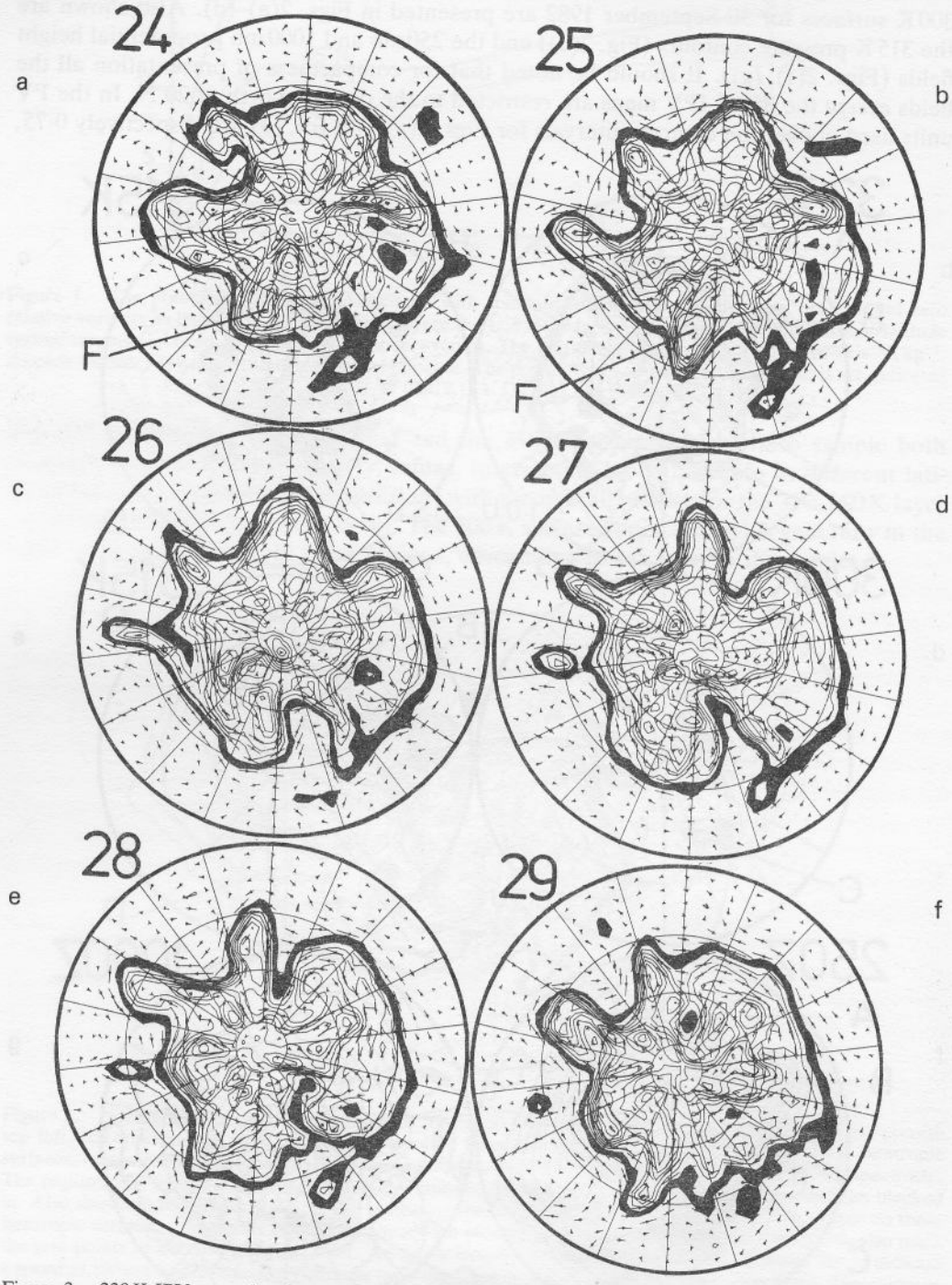


Figure 3. 330 K IPV maps for the region north of 20°N for the period 24–29 September 1982. The contour interval is 1·0 PV units and the regions with values 1–2 units are blacked in. Also shown are the horizontal velocity vectors on this surface, scaled as in Fig. 2.

higher resolution analyses using special data, such as those of Danielsen (1968) and Shapiro (1976), is representing the transition from tropospheric to stratospheric air in a reasonable manner, in much the same way as the objective 300–310 K maps of Obukhov

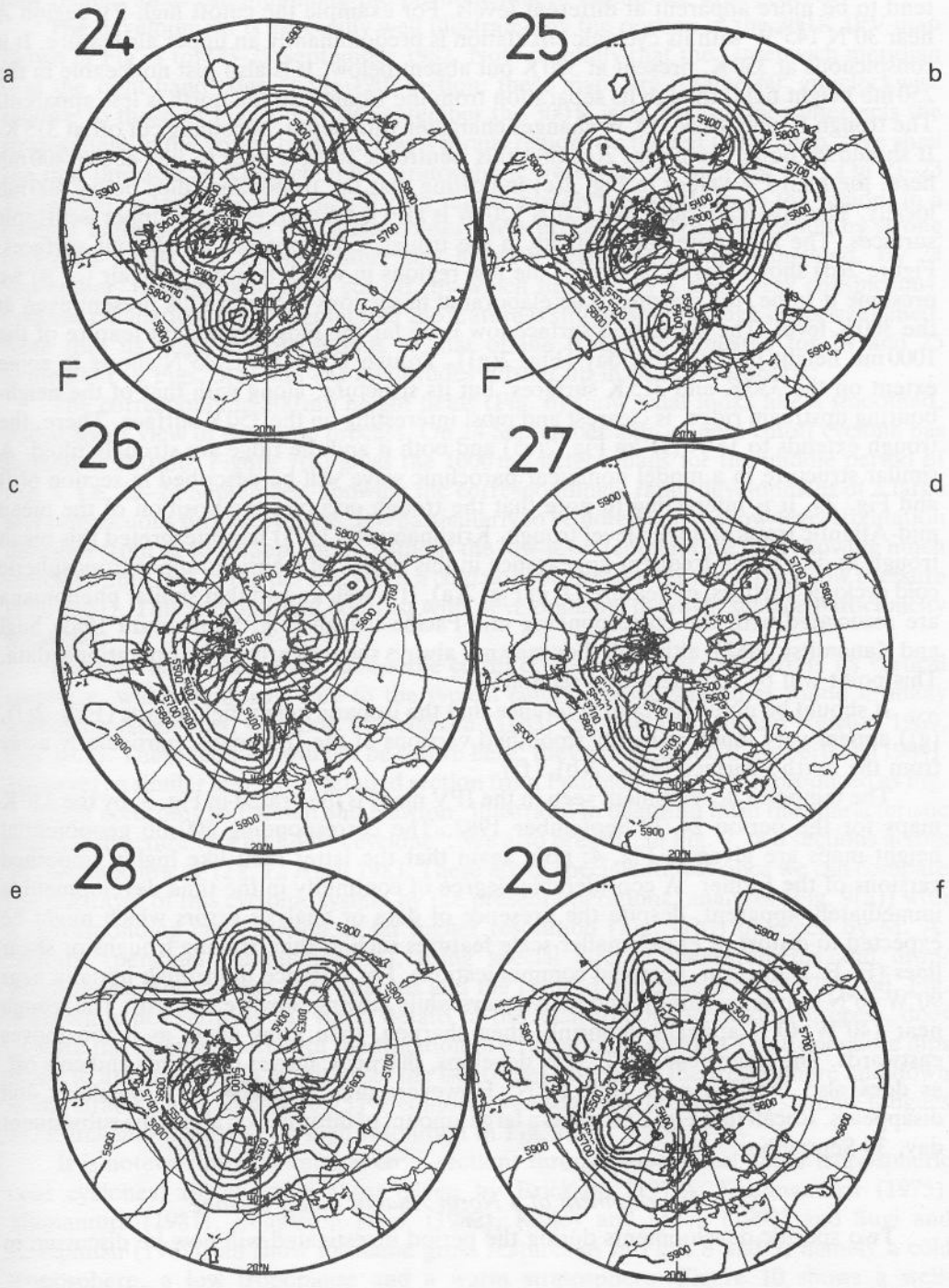


Figure 4. The 500 mb geopotential heights corresponding to the IPV maps shown in Fig. 3. They are for the period 24–29 September 1982 and cover the region north of 20°N. The contour interval is 100 m.

(1964), Danielsen (1967, 1968), and others. To aid visibility in Figs. 2(a)–(d), the region from 2·25 to 3 PV units is darkened in (a), and similarly 2 to 3 units in (b) and 1·5 to 2 units in (c) and (d).

On inspection of Figs. 2(a)–(d), vertical continuity is clear although different features tend to be more apparent at different levels. For example the cutoff high-PV region A near 30°N 145°W with its cyclonic circulation is predominantly an upper air feature. It is conspicuous at 350 K, present at 330 K but absent below. It is also just noticeable in the 250 mb height field, though its separation from the trough* to the north is less apparent. The trough B near 40°N 115°W changes character with depth, becoming cut off at 315 K. It should be noted from Fig. 2(e) that this isentropic surface rises to just above 400 mb here, the high PV values in Fig. 2(c) indicating that the tropopause dips below 400 mb locally. The upstream ridge near 60°N 120°W is also most marked in the upper isentropic surfaces. The high potential vorticity in the trough C is apparent on all four surfaces. Figure 2(d) shows this to be one of the few regions in which stratospheric air (or so we presume it to be, for reasons to be elaborated in sections 2(*c*) and 8) is present even at the 300 K level. The associated surface low is by far the most prominent feature of the 1000 mb height field on this day (Fig. 2(g)). Trough D near 5°E 45°N shows to some extent on the 330 K and 315 K surfaces, but its structure, along with that of the neighbouring upstream ridge, is clearest and most interesting on the 350 K surface. There, the trough extends to 15°N (D′ in Fig. 2(a)) and both it and the ridge are strongly tilted. A similar structure in a model nonlinear baroclinic wave will be described in section 6(*d*) and Fig. 20. It is interesting to note that the trough occupies the position of the mean mid-Atlantic tropical upper-level trough. Krishnamurti (1975) has interpreted this mean trough as due to the frequent appearance in this region of transient upper tropospheric cold cyclonic vortices; cf. feature D″ in Fig. 2(a). It is well known that similar phenomena are associated with the corresponding mid-Pacific trough (e.g. Shimamura 1981; Sugi and Kanamitsu 1982), although they may not always show clearly in the operational data. This point will be returned to in section 9.

It should be noted for later reference that the isobaric height field charts (Figs. 2(f), (g)) appear very much as highly smoothed versions of the IPV maps, particularly away from the earth's surface (Figs. 2(b), (f)).

The temporal development seen in the IPV maps is illustrated in Fig. 3, by the 330 K maps for the period 24–29 September 1982. The corresponding 500 mb geopotential height maps are given in Fig. 4; note again that the latter look like highly smoothed versions of the former. A considerable degree of continuity in the time development is immediately apparent, despite the presence of data or analysis errors which might be expected to distort or erase smaller-scale features such as thin, trailing troughs or shear lines (E. F. Danielsen, personal communication). For instance the trough initially near 90°W 40°N sharpens, cuts off and then decays while moving slowly eastwards. The trough near 140°W 40°N appears to shrink, then sharpen, then grow again as it too moves eastwards. The ridge near 140°E 50°N develops, sharpens, moves eastwards and cuts off, as does also the one near 10°E 50°N. However, that at 120°W 60°N weakens and disappears. Recall that Fig. 2 provides a large amount of information about the subsequent day, 30 September.

(*c*) *Development of a North Atlantic cutoff cyclone*

Two specific developments during the period investigated will now be discussed in

* The word 'trough' is used for an equatorward extension of high PV values since this corresponds to a trough in the associated height field structure. Similarly, 'ridge' is used for the poleward extension of low PV values.

more detail. The first is the formation of a cutoff cyclone in middle and high latitudes, as opposed to the lower-latitude cyclones already mentioned. The associated surface development is rapid and substantial, and the whole process closely resembles one of the scenarios envisaged by Kleinschmidt. Recently, Bleck and Mattocks (1984) have shown how related ideas can illuminate the phenomenon of Alpine lee cyclogenesis.

The development is especially well illustrated by the portion of the 300 K IPV map shown in Fig. 5 for 20–25 September 1982. Note carefully that the Greenwich meridian is at the right-hand edge of the 120° section: these maps are rotated 60° anticlockwise relative to those of Fig. 3. On 20 September the 300 K isentropic surface intersects the stratosphere in the vortex over the Davis Straits region, on the left of the map. (We shall see that this acts as Kleinschmidt's stratospheric 'reservoir' of high-PV air.) On 21–23 September, a large piece of high-PV air moves rapidly eastwards from the region, in a manner highly suggestive of a quasi-conservative process involving advection by strong upper air winds. By 23 September this high-PV region is greatly elongated. By 24 September, it appears to have cut off, by which time it has it own, very conspicuous, cyclonic circulation. On 25 September the feature is still present, though it has weakened. Note that its position corresponds to that of the trough F indicated for 24 and 25 September in Figs. 3(a) and (b) at altitudes further up in the stratosphere.

The corresponding 500 mb height fields (Fig. 6) show what looks like a smoothed, out-of-focus view of the same process. On 24 September a prominent cutoff cyclone has formed at 500 mb. Figure 7(a) gives the 1000 mb height maps for the middle period of Fig. 5, i.e. 22–24 September, showing the correspondingly rapid development of a large and conspicuous surface cyclone. It is particularly to be noted that the low-level circulation develops from 22 to 24 September without the low-level temperature field showing much structure until late in the period, when a cold pool of air sits over the surface low pressure (Fig. 7(b)). This is in marked contrast with developments in which surface baroclinicity is crucial.

Sections have been made across the cutoff cyclone in order to illustrate its vertical structure, which conforms well to the typical cutoff cyclone structures found in many earlier investigations, e.g. that of Peltonen (1963; see also Palmén and Newton 1969, Fig. 10.8). These sections will not be shown here; they can be found in Robertson (1984). However, a similar and more detailed section from Peltonen's paper is reproduced as Fig. 8, which according to Palmén and Newton "illustrates in beautiful form the characteristic thermal structure of a high-level cyclone". We also present, in Fig. 9, two sections across an Icelandic low at 12 z, 12 April 1983. These are of special interest since we can compare the portrayal of this cyclonic system by the present operational analyses (Fig. 9(a)) with a corresponding cross-section at far higher resolution (Fig. 9(b)) obtained by M. A. Shapiro (personal communication), using special aircraft measurements and drop-windsondes. The routine data analysis and the centred differences on a 3·75° grid, Fig. 9(a), while not showing the finest details of the local tropopause-folding process, does capture the essential nature of the situation to a remarkable degree. The tropopause dips down to at least 500 mb in this case, the isentropes bow up in the troposphere and down in the stratosphere, and the 0·75 PV contour even suggests, crudely, the shallow tongues of stratospheric air seen at high resolution in Fig. 9(b).

It is noteworthy that vertical cross-sections through *subtropical* upper tropospheric cold cyclones, for instance those given by Erickson (1971), Krishnamurti (1975), Shimamura (1981), Bengtsson *et al.* (1982), Kelley and Mock (1982) and Sugi and Kanamitsu (1982), all show the same gross features as in Figs. 8 and 9, namely a cold troposphere, a low tropopause and a warm stratosphere. Figure 10 shows a well-documented example taken from Erickson (1971). The same structure was also seen in

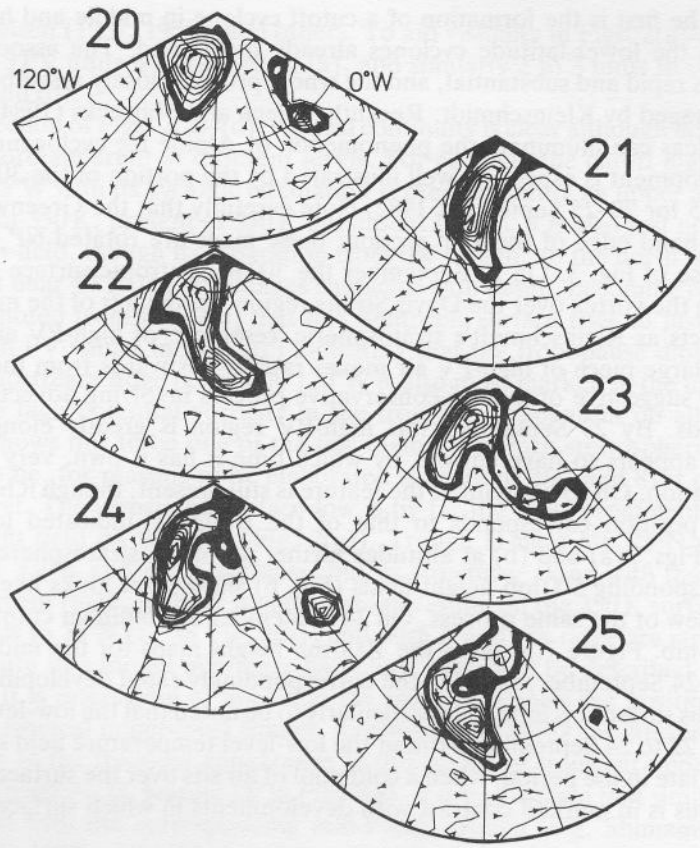


Figure 5. Sectors of the 300 K IPV maps for the period 20–25 September 1982. The region covered is from 40°N to the north pole and from 120°W to 0°W with the 60°W meridian central. The contour interval is 0·5 PV units and the region 1·5–2 units is blacked in. Also shown are the horizontal velocity vectors on this surface, scaled as in Fig. 2.

the present data in association with the feature D″ in Fig. 2(a). We shall see in section 3 that such a structure, which all the above-mentioned cases share qualitatively, is precisely that expected theoretically for the flow structure associated with an upper air high-PV anomaly like that originally postulated by Kleinschmidt. Moreover such gross differences as there are between tropical and high latitude cases are qualitatively consistent with those predicted theoretically from the different values of the Coriolis parameter (Eqs. (33) below).

Finer details such as the shallow tongues seen in Fig. 9(b), whose presence or absence may be difficult to determine observationally, are probably explicable, when they occur, in terms of the detailed spatial structure of the PV anomaly and its formation by competing processes such as different amounts of advection along different isentropic surfaces, and non-conservative effects such as clear air turbulence (e.g. Staley 1960; Shapiro 1978, 1980; Holopainen and Rontu 1981), which may tend to modify such tongues locally (see section 7).

(*d*) *A minor blocking episode*

The other development to be singled out for discussion is that of a blocking anticyclone, again in the North Atlantic. A portion of each 330 K IPV map for the period 30 September–7 October is shown in Fig. 11. Note that this period follows on from that

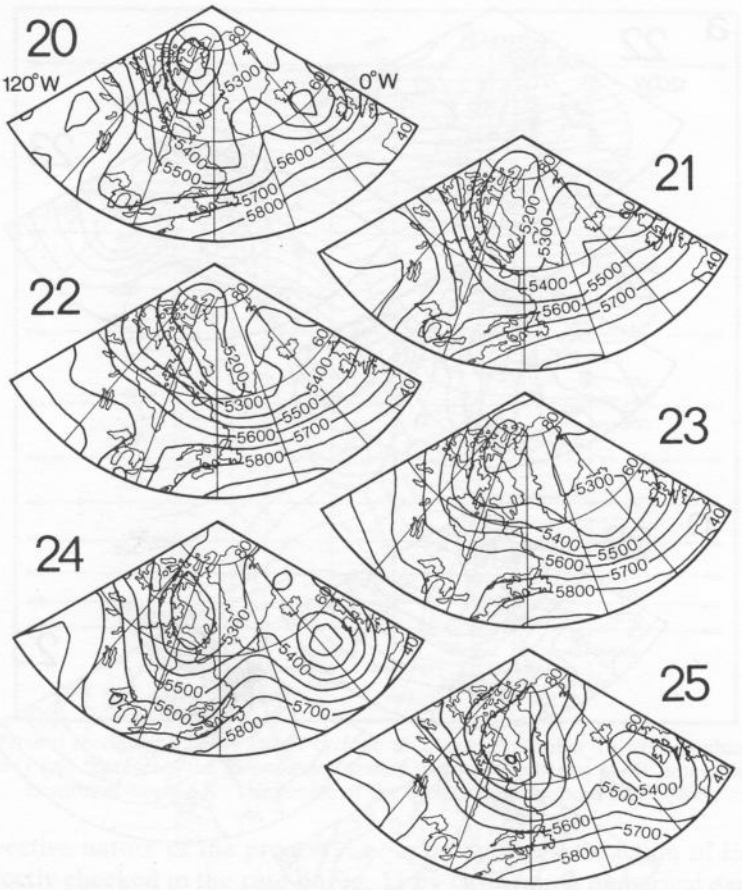


Figure 6. The 500 mb geopotential heights corresponding to the IPV maps shown in Fig. 5. They are for the period 20–25 September 1982 and the region 40°N–90°N and 120°W–0°W. The contour interval is 100 m.

illustrated in Figs. 2 and 3 with an overlap of one day (Figs. 2(b) and 11(a)), and that unlike Figs. 5–7 these maps have the *same* orientation as Figs. 2 and 3, with the Greenwich meridian central. The trough near 30°W in Fig. 11(a) corresponds to the broad trough C of Fig. 2(b).

Figures 2(a)–(d) and 2(g) indicate that the trough C represents an IPV anomaly of great strength, depth and breadth with which is associated a strong cyclonic circulation at all levels. Note in particular the suggestion of a cyclonic spiral pattern in Fig. 2(c). Figures 11(a) to (c) indicate that the strong south-westerly winds ahead of this system are very effective in advecting low-PV air polewards, the more so once they are reinforced by the anticyclonic circulation developing around the encroaching low-PV air. By 2 October, this low-PV air has reached 70°N and is starting to cut off. On 3 October the cutting-off process is complete with a large pool of low-PV air of subtropical origin stranded at 60°N with its own strong anticyclonic circulation. Over the next few days this low-PV air moves eastward and southward, spinning as it goes, with little change in magnitude and, on 7 October, appears to have linked back with the subtropical air. Upstream, starting on 4 and 5 October, the same sequence of events recurs so that by 8 October (not shown) a new cutoff has formed to the west of the previous one. Again the 250 mb height field sequence in Fig. 12 shows what looks like a smoothed version of these maps, with a blocking anticyclone forming and moving south-eastwards and a new

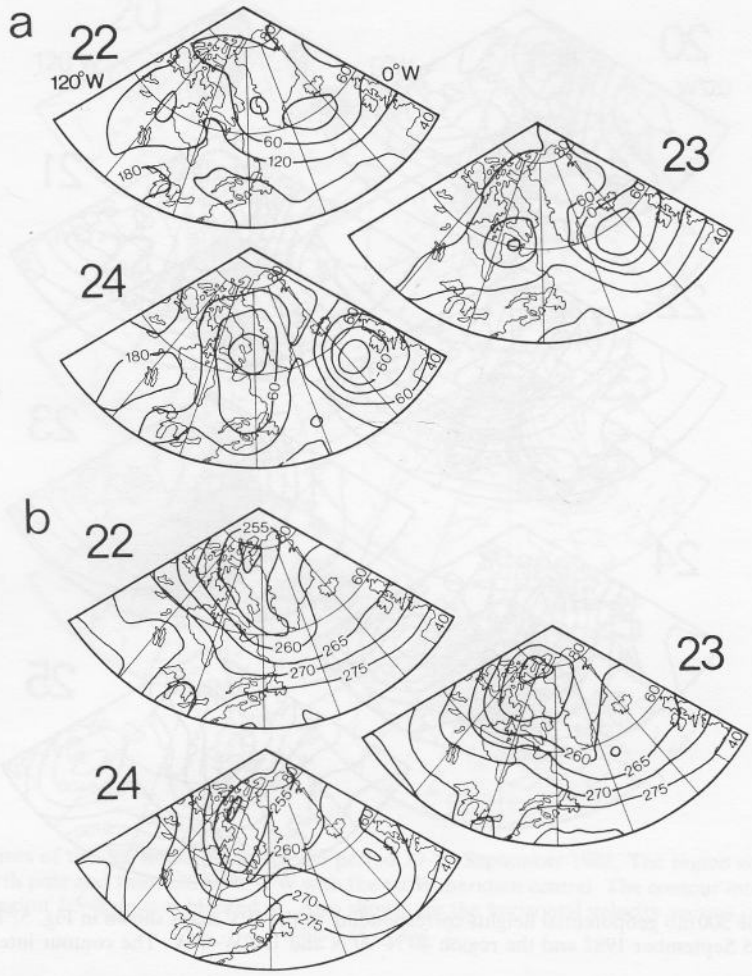


Figure 7. The 1000 mb geopotential (a) and 700 mb temperature (b) for the period 22–24 September 1982 and the region 40°N–90°N and 120°W–0°W corresponding to that shown in Figs. 5 and 6. The contour intervals are 60 m and 5 K, respectively.

one beginning to appear upstream. The 1000 mb height and 700 mb temperature for 2 October (Fig. 13) emphasize the deep vertical scale of the poleward flow of subtropical air. The low-level temperature reflects the invasion of the subtropical air mass seen so clearly in Fig. 11(c).

Having studied this one event we should note that Fig. 3 shows a very similar incipient blocking sequence occurring from 25 to 29 September (although less high-PV air is finally cut off) so that near 20°E the entire two-week period, 25 September to 8 October, is dominated by blocking associated with two separate poleward incursion episodes and the subsequent cutting off of low-PV subtropical air. As already mentioned, a third incursion occurred after 8 October, albeit further west. Such repeated incursions seem typical (e.g. Berggren *et al.* 1949; Palmén and Newton 1969, § 10.1), and as Mahlman (1979) points out they may be the physical reality behind the recently developed 'eddy-mean' view of the maintenance of time-averaged or low-pass-filtered blocking events (Green 1977; Mahlman 1979; Shutts 1983, 1985; Hoskins *et al.* 1983; Illari and Marshall 1983; Illari 1984).

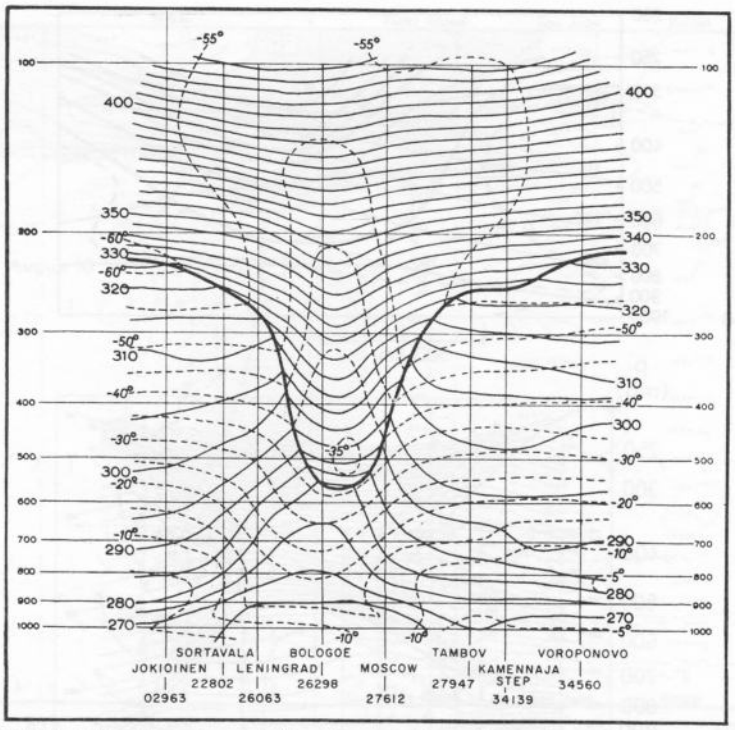


Figure 8. A vertical section through a cutoff cyclone at 12 GCT November 16 1959 produced by Peltonen (1963). The heavy line represents the tropopause; dashed lines are isotherms at 5°C intervals and solid lines isentropes every 5 K. The centre of the cyclone was at about 35°E, 58°N.

The advective nature of the process, i.e. approximate satisfaction of Eqs. (11) and (12), was directly checked in the case of Fig. 11 by performing numerical experiments in which a passive tracer was advected starting with the initial conditions of the 330 K IPV map of 2 October, and using the 330 K horizontal velocity field interpolated between 2 and 3 October (Robertson 1984). The advection equation was integrated using simple centred finite differences in space and time with a half-hour timestep. With no dissipation included, the essential nature of the cutoff at 3 October is reproduced, but there is a large amount of small-scale structure. The inclusion of a nonlinear scale-selective dissipation, acting on a time scale of order one day for features of length scale $L \sim 200$ km, reduces the noise and produces a picture which resembles the observed picture very well in the features to which attention has been drawn including, again, the cutoff at 3 October.

Similar phenomena are observed in ocean currents. The best-documented cases are probably those of 'Gulf Stream rings' (e.g. Rhines 1979, Fig. 1; Robinson 1983; Kennelly *et al.* 1985).

(*e*) *The conceptual duality between cutoff cyclones and blocking anticyclones*

The individual cutoff cyclones and blocking anticyclones which we have studied appear to be formed in a basically similar manner, by advection leading to the cutting-off of upper IPV anomalies of the appropriate sign. Indeed the cutoff cyclone and blocking anticyclone appear from this viewpoint to be fundamentally the same phenomenon, the only difference (apart from differences of spatial scale) being the sign of the IPV anomaly. Confidence in this conceptual duality is further strengthened by a comparison of vertical cross-sections. Figure 14 shows a cross-section through the blocking anticyclone of 4 October along the line of longitude indicated in Fig. 11(e), from the operational data.

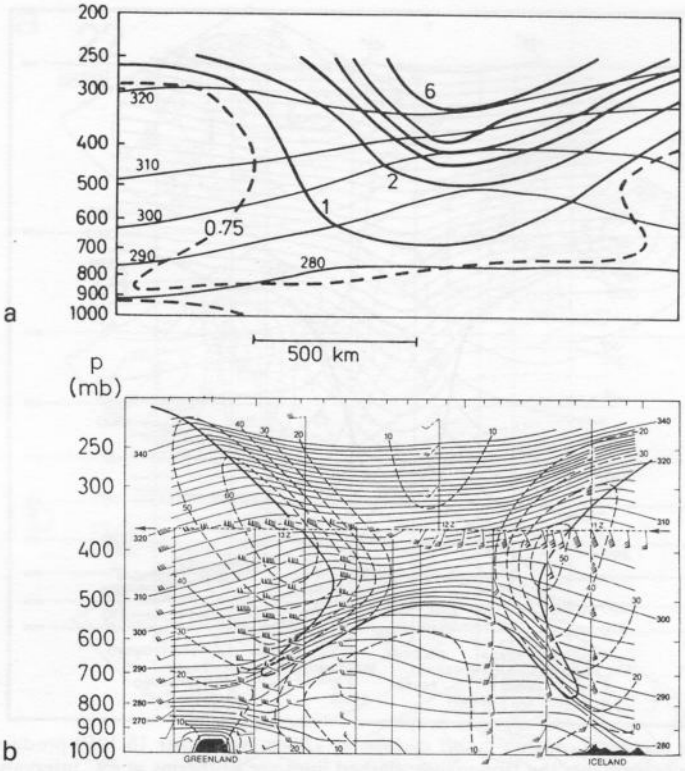


Figure 9. Vertical cross-sections along a SW–NE line through the Icelandic low on 12 April 1983. The first section (a) shows the results obtained by simple finite differences applied to the routine ECMWF analysis at 12 z on that day. The thin lines indicate isentropes at 10 K intervals and the thick lines PV at 1 unit intervals. The 0·75 PV unit contour is shown by a heavy dashed line. The second section (b), from M. A. Shapiro (personal communication), is the result of detailed measurements using aircraft and dropwindsondes. Isentropes every 2 K are indicated by light solid lines, isotachs every 10 m s^{-1} by dashed lines and the estimated tropopause position by a heavy continuous line. The scales and orientation of the two sections are approximately the same.

As far as gross features are concerned, everything is the opposite way round from Figs. 8–10. The tropopause is high, the isentropes in the troposphere bow downwards, and those in the stratosphere bow upwards. The very low potential vorticity just under the tropopause, less than 0·5 PV units, is particularly striking. Once again we shall see that this structure is just that expected theoretically for the fields induced by an isolated, upper air IPV anomaly of the appropriate sign.

The question of diabatic and frictional effects will be postponed until sections 6(*e*) and 7. As we shall argue in section 7, the duality ends there: diabatic effects, in particular, are likely to be quite different for the two signs of anomaly.

3. Some simple examples, following Kleinschmidt

We next present and discuss some theoretical examples which will illustrate, in idealized form, the typical flow structures induced by isolated IPV anomalies of either sign in a given, statically stable reference state. These examples will prove useful as basic building blocks for IPV thinking.

The examples assume a horizontally uniform reference state, with a constant Coriolis parameter f. The mass distribution of potential temperature θ is specified by prescribing the associated hydrostatic pressure field p as a function of θ:

$$p = p_{\mathrm{ref}}(\theta). \tag{16}$$

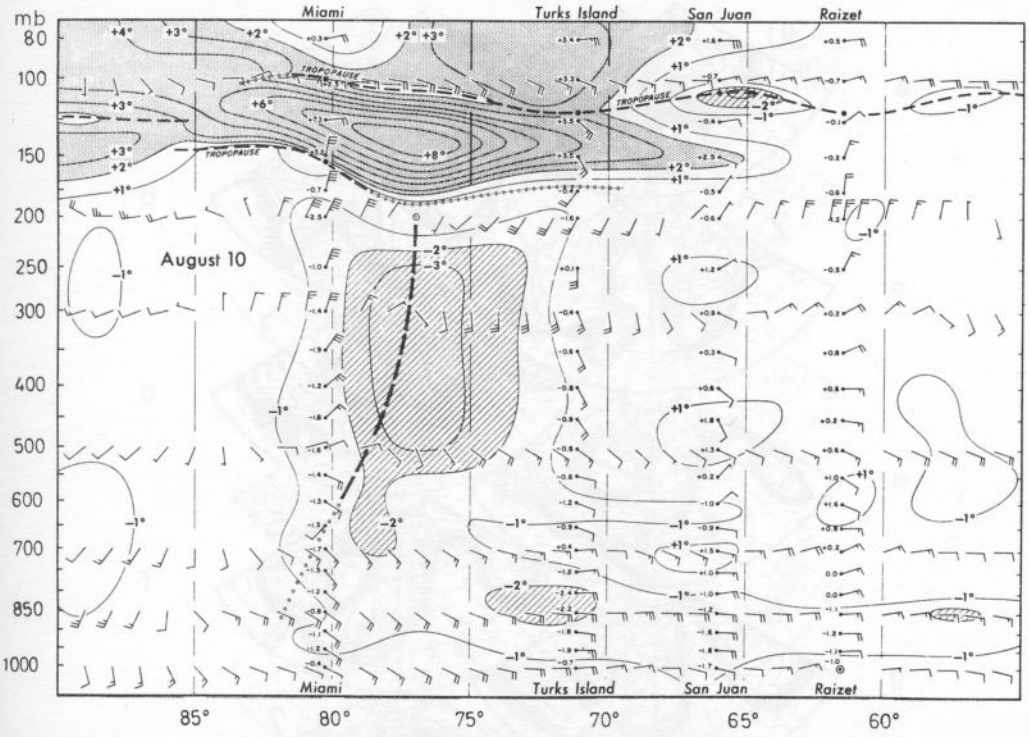


Figure 10. A vertical section along a NW–SE line just north of Cuba through a subtropical oceanic upper-level cold vortex at 1200 GMT 10 August 1966, produced by Erickson (1971). Isolines are temperature deviations (°C) from a mean Caribbean atmosphere for August. Deviations exceeding 2°C are shaded. Data are also plotted for four individual stations; interpolated winds (full barb = 10 kt ≃ 5 m s^{-1}) are from streamline–isotach analysis. The heavy dashed line is the axis of the upper-level cyclonic curvature.

The function $p_{\mathrm{ref}}(\theta)$ is assumed to be monotonically decreasing, so that the reference state is statically stable. Being horizontally uniform, the reference state has constant PV on each isentropic surface, $P_{\mathrm{ref}}(\theta)$ say.

We express condition (ii) of section 1(d) by requiring that the actual state have the same mass distribution as the reference state (16), i.e. that the actual state can be obtained from the reference state by an adiabatic rearrangement of mass. Thus the actual hydrostatic pressure field $p(x, y, \theta)$, expressed in isentropic coordinates, will be required to satisfy, for each value of θ, the relation

$$\iint \frac{\partial p(x, y, \theta)}{\partial \theta}\, dx\, dy = \frac{dp_{\mathrm{ref}}(\theta)}{d\theta} \iint dx\, dy \tag{17a}$$

where the derivative $\partial/\partial\theta$ is taken at constant horizontal position (x, y). The area integral is understood to be taken over a horizontal domain much larger than the size of the IPV anomaly. The condition (17a) is equivalent to requiring that the mass lying between the isentropic surfaces θ and $\theta + d\theta$ be the same as in the reference state. It should be noted that together with (14) this implies, in the same way as in (10), a relationship between the distribution of IPV anomalies $P' = P(x, y, \theta) - P_{\mathrm{ref}}(\theta)$ on each isentropic surface, and the relative circulation $C_{\mathrm{b}}(\theta)$ around the intersection of that surface with the domain boundary, assuming smooth stratification. The relationship may be written

$$-\iint g^{-1}(\partial p/\partial\theta) P'\, dx\, dy = C_{\mathrm{b}}(\theta). \tag{17b}$$

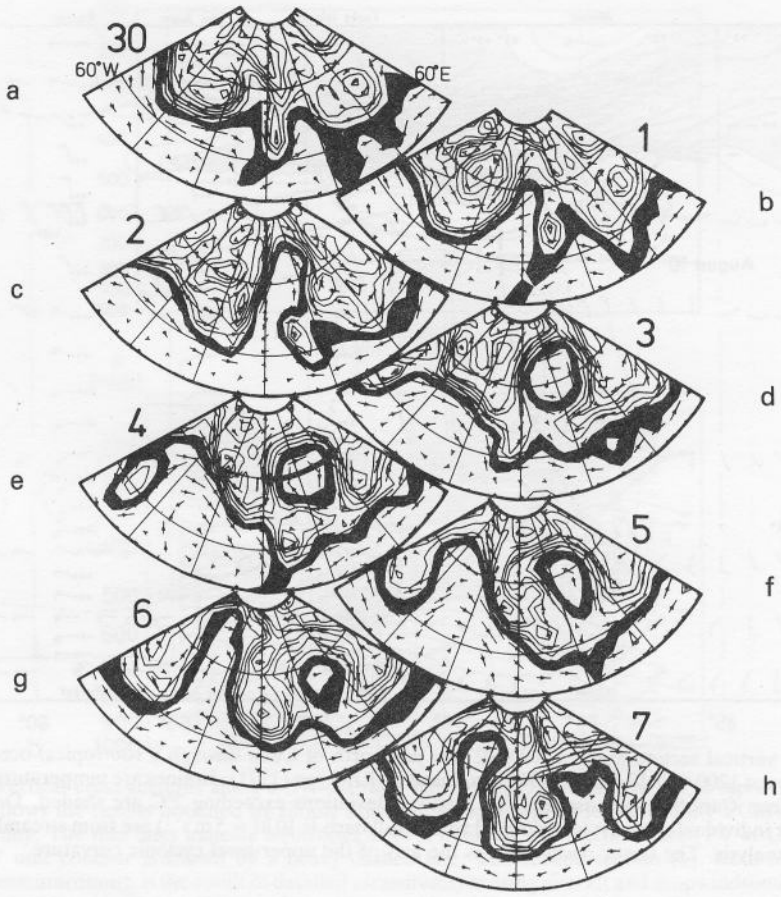


Figure 11. Sectors of the 330 K IPV maps for the period 30 September–7 October 1982. The region covered is from 30°N to 80°N and from 60°W to 60°E, with the Greenwich meridian central. The contour interval is 1 PV unit with the regions 1–2 units blacked in. The line at 60°N on 4 October indicates the location of the section shown below in Fig. 14. Also shown are the horizontal velocity vectors on this surface, scaled as in Fig. 2. Again, the points of the arrow heads are plotted at the mid-points of the arrows which also mark the grid-points to which the vectors refer.

For horizontally large domains it may be useful to restrict attention to cases in which $C_b(\theta)$ is, say, zero or constant, and the restrictions then imposed by (17b) on possible configurations of IPV anomalies should be noted. In particular, an 'isolated' anomaly must then be considered to be embedded in a surround of weak P' values inversely proportional to the area of the domain. This is to be understood, where necessary, in what follows.

Suppose now that the actual state has a given, circularly symmetric PV anomaly on some of the isentropic surfaces. Considerations of symmetry imply that the associated wind, pressure and temperature fields will also be circularly symmetric. If friction and diabatic heating are neglected, the flow must also be steady, nondivergent and purely azimuthal. This will now be assumed to be the case, and the azimuthal wind speed written as a function $v(r, \theta)$ of θ and the horizontal distance r from the centre of the anomaly. The *balance condition*, condition (i) of section 1(d), is here expressed by the gradient wind and hydrostatic relations, which are exact under the present assumptions. In terms of the Montgomery potential $M(r, \theta) = \phi + c_p T$, ϕ being the geopotential, the gradient

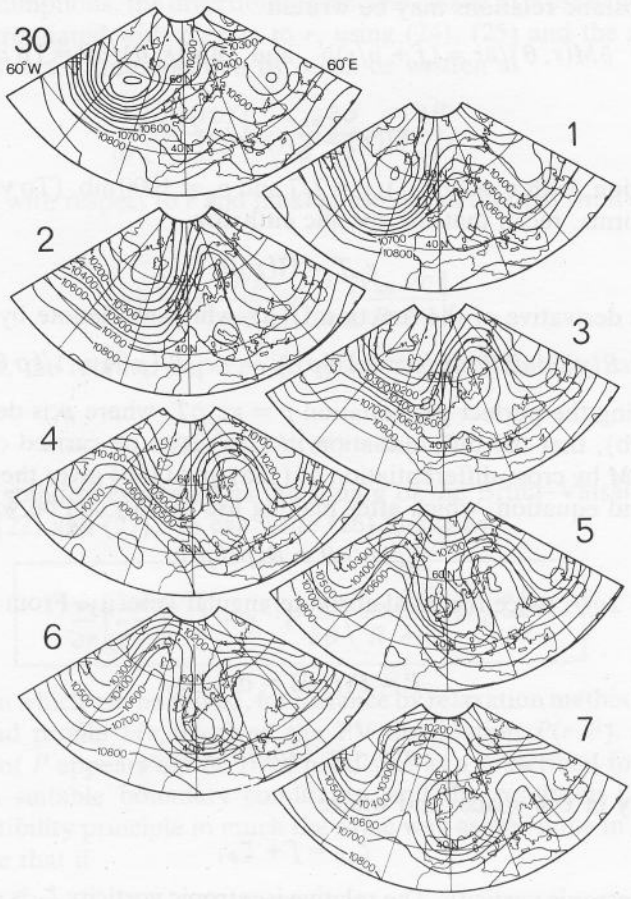


Figure 12. The 250 mb geopotential heights corresponding to the IPV maps shown in Fig. 11. They are for the period 30 September–7 October 1982 and the region 30°N–80°N and 60°W–60°E. The contour interval is 100 m.

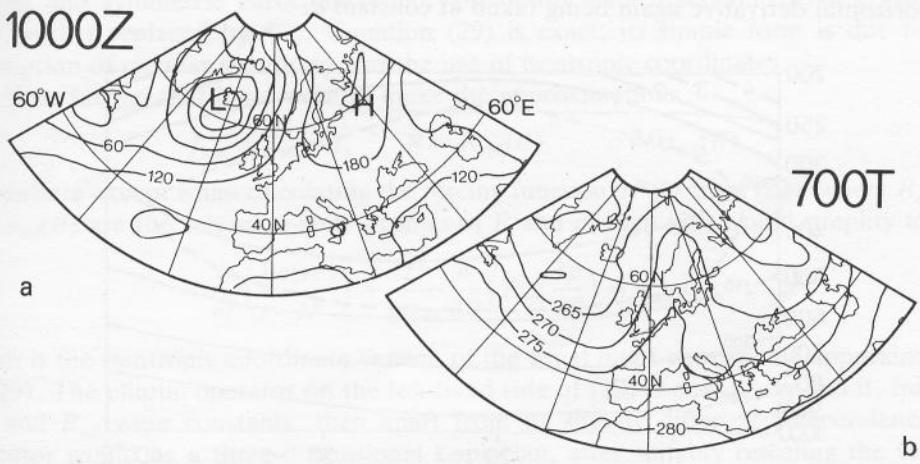


Figure 13. The 1000 mb geopotential height (a), and 700 mb temperature (b), for 2 October 1982 and the region 30°N–80°N and 60°W–60°E corresponding to that shown in Figs. 11 and 12. The contour intervals are 60 m and 5 K respectively.

wind and hydrostatic relations may be written

$$\partial M(r,\theta)/\partial r = (f + v/r)v \quad \text{and} \quad \partial M(r,\theta)/\partial\theta = \Pi(p) \tag{18a, b}$$

where

$$\Pi(p) = c_p(p/p_0)^\kappa, \tag{19a}$$

the Exner function, with $\kappa = (\gamma - 1)/\gamma \simeq 2/7$ and $p_0 = 1000\,\text{mb}$. (To verify (18a, b) from their isobaric forms, recall that the specific enthalpy

$$c_p T = \Pi(p)\theta, \tag{19b}$$

so that the first derivative of the function $\Pi(p)$, which we denote by $R(p)$, satisfies

$$R(p) = d\Pi(p)/dp = \kappa\Pi(p)/p = \kappa c_p T/(p\theta) = 1/(\rho\,\theta), \tag{20}$$

the last step using the perfect gas equation $p = \kappa c_p \rho T$, where ρ is density. Recall also, in applying (19b), that the differentiation in (18a) is to be carried out at constant θ.) Elimination of M by cross-differentiation of (18a) and (18b) gives the isentropic form of the thermal wind equation, which after making use of (20) can be written

$$f_{\text{loc}}\partial v/\partial\theta = R(p)\partial p/\partial r \tag{21}$$

where $f_{\text{loc}} = f + 2v/r$, twice the local absolute angular velocity. From (14), the potential vorticity is

$$P = P(r,\theta) = \sigma^{-1}\zeta_{a\theta}, \tag{22}$$

where

$$\sigma = -g^{-1}\partial p/\partial\theta > 0, \tag{23}$$

the mass density in $xy\theta$ space, and where

$$\zeta_{a\theta} = f + \zeta_\theta, \tag{24}$$

the absolute isentropic vorticity. The relative isentropic vorticity ζ_θ is given by (9), which under the assumption of circular symmetry is equivalent to

$$\zeta_\theta = r^{-1}\partial(rv)/\partial r, \tag{25}$$

the horizontal derivative again being taken at constant θ.

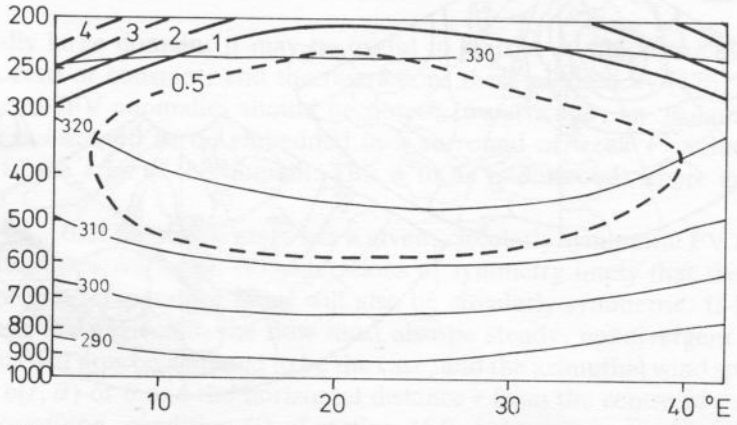


Figure 14. A vertical section on 4 October 1982 along the line at 60°N shown in the corresponding panel of Fig. 11. As in Fig. 9(a), the light continuous lines are isentropes drawn every 10 K and the heavy continuous lines are PV contours every 1 unit. The heavy dashed contour indicates a PV value of 0·5 units.