Reprinted with pagination changes from *Phil. Trans. Roy. Soc. London* **352**, 227-240 (1995). *Corrigendum: See footnote to section 2, page 4.*

The stratospheric polar vortex and sub-vortex: fluid dynamics and midlatitude ozone loss

M. E. McIntyre

Centre for Atmospheric Science* at the
Department of Applied Mathematics and Theoretical Physics,
Silver Street, Cambridge CB3 9EW, England

Abstract

It has been suggested on the basis of certain chemical observations that the wintertime stratospheric polar vortex might act as a chemical processor, or flow reactor, through which large amounts of air — of the order of one vortex mass per month or three vortex masses per winter — flow downwards and then outwards to middle latitudes in the lower stratosphere. If such a flow were to exist, then most of the air involved would become chemically "activated", or primed for ozone destruction, while passing through the low temperatures of the vortex where fast heterogeneous reactions can take place on polar-stratospheric-cloud particles. There could be serious implications for our understanding of ozone-hole chemistry and for midlatitude ozone loss, both in the Northern and in the Southern hemisphere. This paper will briefly assess current fluid-dynamical thinking about flow through the vortex. It is concluded that the vortex typically cannot sustain an average throughput much greater than about a sixth of a vortex mass per month, or half a vortex mass per winter, unless a large and hitherto unknown mean circumferential force acts persistently on the vortex in an eastward or "spin-up" sense, prograde with the Earth's rotation. By contrast, the "sub-vortex" below pressure-altitudes of about 70 hPa (more precisely, on isentropic surfaces below potential temperatures of about 400 K) is capable of relatively large mass throughput depending, however, on tropospheric weather beneath, concerning which observational data are sparse.

Key words: stratosphere, ozone, halocarbons, chlorofluorocarbons, chemical activation, flowing processor, mean circulation, gyroscopic pumping.

1. Introduction

There is controversy over the causes of the observed midlatitude decline in stratospheric ozone (e.g. Pyle, this volume). The main questions are (*a*) whether the midlatitude decline represents polar ozone loss that is spreading, in some sense, to middle latitudes, (*b*) whether such spreading involves further ozone loss in middle latitudes, or mere dilution by ozone-depleted polar air, and (*c*) how far the total midlatitude decline in stratospheric ozone might proceed in future. These problems concern the Northern as well as the Southern Hemisphere. Although the chemical ozone-loss mechanisms are generally strongest in the Southern Hemisphere, the fluid-dynamical transport mechanisms are generally, as it happens, strongest in the Northern Hemisphere.

What is our most secure piece of knowledge about these problems? It is that the most conspicuous ozone loss seen so far, which takes place in the springtime Antarctic stratosphere, is mostly

* The Centre for Atmospheric Science is a joint initiative of the Department of Chemistry and the Department of Applied Mathematics and Theoretical Physics. Correspondence on this paper should go to the Department of Applied Mathematics and Theoretical Physics, or to the email address M.E.McIntyre@amtp.cam.ac.uk.

caused by man-made halocarbons. These stable chemicals are observed to be well mixed in the troposphere, with Northern and Southern Hemispheric mixing ratios generally within about 10% of each other. This is possible because of the halocarbons' exceptional chemical stability, and low solubility in water. The halocarbons are carried up into the tropical stratosphere, and taken to high stratospheric altitudes in a slow but inexorable large-scale upwelling motion; they are then photolysed by the Sun's hard ultraviolet radiation to which they become exposed above about 20 km altitude. This global-scale pattern of tropospheric mixing and large-scale, tropical-stratospheric upwelling is part of the reason why Northern Hemispheric pollution can cause Southern Hemispheric ozone loss.

The mechanism of the large-scale tropical upwelling is well understood, being describable to first approximation as a global-scale 'gyroscopic pumping'. The extratropical stratosphere and mesosphere, up to altitudes of about 80 km, act on the tropical stratosphere like a gigantic, seasonally and interannually variable suction pump, whose action depends on the Earth's rotation together with certain wavelike and turbulent eddy motions. Further explanation and references are given in the Appendix. The effect is to pull air gently but persistently upward and poleward out of the tropical troposphere and lower stratosphere, then push it back down toward the extratropical troposphere, most of it through the winter stratosphere via complicated, chaotic pathways. Some of the poleward and downward moving air, carrying photolysed halocarbon fragments, gets into the polar lower stratosphere where, in the case of the Antarctic, the photolysed fragments participate in the formation of a conspicuous 'ozone hole' via reactions during winter and spring whose final stages depend on the springtime return of sunlight. Meanwhile, more tropospheric air is being pulled up into the tropical stratosphere, importing more halocarbons to be photolysed in their turn.

Typical large-scale upwelling velocities in the tropical lower stratosphere (altitudes 15-20 km) are seasonally variable roughly from $0.2\,\mathrm{mm\,s^{-1}}$ in northern summer to $0.4\,\mathrm{mm\,s^{-1}}$ in northern winter, or roughly 6-13 km per year with the largest values confined mainly to the northern winter. Such velocities are of course not directly measurable but there are now several independent ways of estimating them, including new and relatively precise tracer information (Mote *et al.* 1996; Holton *et al.* 1995). The gyroscopic pumping rate sets the e-folding timescale, of the order of 100 years, for removal of the halocarbons from the troposphere, because rates of land and ocean uptake of halocarbons are at least a decimal order of magnitude slower (e.g. Junge 1976).

The gyroscopic pumping mechanism has often been called 'wave driving' and the resulting global-scale circulation the 'wave-driven circulation', because of the nature of the wavelike and turbulent eddy motions involved (see the Appendix), giving rise for instance to the so-called 'Rossby-wave surf zone' in middle latitudes. These same wavelike and turbulent eddy motions give rise to another remarkable and chemically important phenomenon, the approximate chemical isolation of the winter stratospheric polar vortex, now widely believed to be important for ozone-hole chemistry. There is strong evidence, both observational and theoretical — for historical background see McIntyre (1989) and for recent observational evidence see, e.g., among many other references, Dahlberg & Bowman (1994), Lahoz *et al.* (1995), Manney *et al.* (1994b), Waugh *et al.* (1993), Norton and Chipperfield (1995) — there is strong evidence that the edge of the stratospheric vortex acts as a flexible 'eddy-transport barrier' bounding the surf zone and inhibiting the eddy transport of lagrangian tracers along the stratification or isentropic surfaces and across the edge of the vortex.

Fast eddy transport along isentropic surfaces would be expected to take place freely on the basis of the usual intuitions about quasi-horizontal or layerwise-two-dimensional turbulence in strongly stratified fluids. Such fast eddy transport does indeed take place, but only in the surf zone and similar regions. The spatial inhomogeneity vitiates standard turbulence-theoretic assumptions, but it is an almost inevitable consequence of a relevant fluid-dynamical theorem, the conservation of

Rossby–Ertel potential vorticity (PV). Such inhomogeneity is practically certain to occur in real flows of the kind in question, for simple but robust theoretical reasons (McIntyre 1994). A typical example was shown at the Discussion Meeting using a videotape made from a high-resolution numerical model stratospheric simulation (Norton 1994), and available as a VHS PAL or NTSC cassette on request from Dr W. A. Norton, Dept. of Atmospheric, Oceanic and Planetary Physics, Oxford (wan@atm.ox.ac.uk).

This videotape demonstrates all the effects in question, in particular the dramatic way in which the eddy-transport barrier marking the edge of the polar vortex can strongly inhibit eddy tracer transport, even in the presence of a violently distorted vortex and strong layerwise-two-dimensional turbulence in the midlatitude Rossby-wave surf zone. The barrier effect depends on two things, first the 'Rossby-wave quasi-elasticity' associated with isentropic gradients of PV, which gradients tend to be concentrated in the vortex edge, and second the horizontal shear near the edge of the vortex (Juckes and McIntyre 1987).

The likely fine structure of the vortex edge is shown schematically in figure 1. The effect of the surrounding 'surf zone' is to erode the edge and steepen the isentropic gradients of PV and chemical tracers. There is usually a multiple-edge structure in these gradients, here suggested only schematically. This has been demonstrated especially clearly by the recent work of Waugh *et al.* (1993) using high-resolution lagrangian advection techniques and meteorological estimates of the wind field. Because of the combined effects of surf-zone turbulence and vertical shear, the edge is likely to show a 'screwthread' structure, left-handed in the Antarctic and right-handed in the Arctic, with slopes of the order of Prandtl's ratio (Coriolis parameter over stratification buoyancy frequency, of the order of 10^{-2}). The screwthread is probably not as regular as is shown schematically here for the sake of visual clarity. Manney *et al.* (1994a) show, in a relatively low-resolution numerical model experiment, the initial stages of formation of such a structure. Figure 2 shows a lidar slice through the Arctic vortex and surrounding aerosol,, suggesting a somewhat coarser structure than in figure 1, whose bottom left hand corner may be compared to the lowest part of the white region in figure 2.

2. The controversy over polar-vortex isolation: the 'flowing processor' hypothesis

Despite the foregoing considerations, the effectiveness of the transport-inhibition mechanism in the real stratosphere has recently been called into question (Tuck *et al.* 1993 & refs.). Tuck *et al.* and others have tentatively hypothesised that, even in the Antarctic, there is a large flow of air through the real polar vortex, downwards and then outwards to middle latitudes in the lower stratosphere. Such a flow would take large amounts of air and halocarbon photolysis fragments through the vortex, where they would be chemically 'activated' by fast heterogeneous reactions within the colder parts of the vortex, then exported to sunlit middle latitudes in the lower stratosphere, the densest part of the ozone layer, with possibly serious consequences in the form of ozone depletion over populated areas. The hypothesised picture has been summarised in the phrase 'the vortex as a flowing processor', or 'flow reactor'.

The flow through the vortex is hypothesised to persist throughout winter and early spring, to be a major contributor to the observed midlatitude ozone decline, both in the northern and in the southern hemisphere, and to explain other observed phenomena as well — particularly the dryness of air in the extratropical southern-hemispheric lower stratosphere that has been observed on some occasions. This dryness, in other words, is hypothesised to be due mainly to the dehydration of air transported through the very cold Antarctic vortex (Tuck *et al.* 1993 & refs.), rather than through the coldest parts of the tropical lower stratosphere, or through the polar 'sub-vortex' region below the 400K isentropic surface, roughly corresponding to pressure-altitudes below 70 hPa.

Tuck *et al.* (*op. cit.*)* state that the vortex is "flushed several times" during a single winter. This presumably means that the flushing time, $\tau_{\rm F}$ say, defined by

$$\tau_{\rm F} = M/\dot{M} \ , \tag{1}$$

is of the order of a month or less. Here M is the mass of air within the vortex in a lower-stratospheric layer, say a density scale height or about 7 km deep, and $\dot{M}$ the average mass-flow rate at which the air within the vortex in the same layer is exported to middle latitudes, and replaced by downward flow within the vortex. Average rate for this purpose means averaged over the winter season.

Although the flowing processor hypothesis seems to be at variance with the results of fluid dynamical studies done so far, the possibly serious implications compel a careful re-examination of the basis for our present understanding of the fluid dynamics. One concern is that computer limitations preclude any direct check from a three-dimensional model having adequate spatial and temporal resolution. It is conceivable, perhaps, that small-scale, numerically unresolved motions might change the picture; so we need to bring to bear all available fluid-dynamical insight as well as modelling. The remainder of this paper presents a summary of such a re-examination; it is hoped to publish the results in more detail during the coming year. The conclusion will be that some vortex air is indeed exported, but at a rate that is both highly variable, and far less than one vortex mass per month.

3. The possibility of mean outflow

Any large export of vortex air in the lower stratosphere can be assumed to take place mainly along the isentropic surfaces of the strong stable stratification. This is because the most typical ways in which air can cross isentropic surfaces outside the vortex are quasi-diffusive, with small diffusivities K_{zz} of order $0.2\,{\rm m^2\,s^{-1}}$ giving diffusion or dispersion height-scales $(K_{zz}t)^{1/2}$ of only a kilometre or so for $t = 2$ months. The two main mechanisms (both giving small-step random walks in the cross-isentropic motion) are, first, local vertical mixing by clear-air turbulent layers (e.g. Dewan 1981) and, second, quasi-horizontal, stratification-constrained (layerwise-two-dimensional) turbulence ('surf zone dynamics') giving a *diabatic* random walk with, by accident, as it happens, diffusivities of roughly the same order of magnitude, $0.2\,{\rm m^2\,s^{-1}}$. Such values are corroborated by lidar observations of volcanic aerosol layers (P. H. Haynes, personal communication). Superposed on these vertical random walks can be weak Lagrangian-mean descent rates, not well known but almost certainly of order perhaps 0.1 to $0.2\,{\rm mm\,s^{-1}}$or $\frac{1}{2}$ to 1 km in 2 months. So any air rapidly exported from the vortex in the lower stratosphere would have to stay fairly close to one isentropic surface.

Transport along isentropic surfaces can be regarded as due either to mean outflow or to outward eddy transport, or to both. First consider the possibility of mean inflow or outflow along isentropic surfaces, as seen in vortex-following coordinates. Wave–mean interaction theory, together with observational knowledge of the relevant Rossby, gravity and inertia–gravity waves, points not to mean outflow but to weak mean inflow (Mo *et al.* 1995). This is part of the wave-driven gyroscopic pumping action already mentioned.

Additional weak inflow is needed to create the vortex on the seasonal timescale. Total poleward parcel displacements of a few degrees latitude are enough to create a vortex of the typical strength observed in the lower stratosphere, if drag on the vortex can be neglected; these displacements are robustly of the same order as would be given by a simplistic angular momentum budget for

* *Note:* This citation was incorrectly printed in the published version, as "(Tuck *et al.* 1993 & references)", through a copy-editing accident.

a frictionless, exactly circular vortex. Thus a hypothetical ring of air, conserving its angular momentum and initially at rest relative to the Earth, will be spun up to $30\,\mathrm{ms}^{-1}$eastwards on arriving at latitude 60° if displaced polewards by only 2.1° latitude. These are typical lower-stratospheric vortex-edge values.

Essentially the same calculation run backwards implies that a mean outflow along isentropes strong enough to conform to the flowing processor hypothesis in the form quoted below (1), i.e., an outflow strong enough to export one vortex mass per month would, in the absence of an eastward force, obliterate the vortex in 4 days or so. The area enclosed by a ring moving 2.1° equatorward *from* latitude 60° expands by a factor

$$\int_{57.9^\circ}^{60^\circ} \cos\phi d\phi \Bigg/ \int_{60^\circ}^{90^\circ} \cos\phi d\phi \ ,$$

which is very close to $\frac{2}{15} = 4$ day/1 month. Similar estimates can be derived in several other, more sophisticated ways, for instance from wave–mean interaction theory (Mo *et al.*, *op. cit.*) and from much more general considerations of the 'dilution of PV-substance' (Haynes and McIntyre 1990), which for this purpose apply to a disturbed as well as to an undisturbed vortex. But the orders of magnitude are little affected: 'strong eastward force of unknown origin' means very strong indeed: strong enough to spin up a new vortex every 4 days or so.

There is no known mechanism that could come anywhere near exerting an eastward (anti-drag) force as strong as this. The known forces, Rossby and gravity wave drag, are far weaker and in any case tend to be westward rather than eastward, being part of the gyroscopic pumping action.

In summary, weak mean inflow is required to create the vortex on the seasonal timescale. Additional mean inflow, of a similar order of magnitude, is generated by gyroscopic pumping action associated with Rossby and gravity wave drag (Appendix & refs.). This inflow, induced by pumping, can also be looked on as maintaining the vortex against Rossby and gravity wave drag. Conversely, if a strong mean outflow were to exist in the real wintertime polar lower stratosphere, it would imply an exceedingly strong *reversed* pumping action, driven by an unknown agency in the form of an exceedingly strong eastward force in the vortex edge.

4. Outward eddy transport by vortex erosion

If the foregoing is accepted, it leaves outward eddy transport as the only mechanism that could act against the presumed mean inflow. But outward eddy transport is rate-limited by the rate at which the vortex edge can be eroded, or material 'stripped' or 'peeled off', by disturbances to the edge and associated midlatitude stirring. The aforementioned model studies — which use either eulerian (fixed-grid) techniques or high-resolution lagrangian advection techniques, either contour advection or many-particle, on model-generated wind fields or meteorologically analysed wind fields (e.g. Pierce and Fairlie 1993; Manney et al 1994b; Norton 1994; Rood et al, personal communication; Fisher et al 1993; Waugh and Plumb 1994; Waugh et al 1993) — all give weak erosion rates, in the sense that the mass transported is, conservatively, no more than about a third of a vortex mass per month, regardless of the ambiguity in defining the vortex edge (due to its filamentary fine structure) and regardless of the very wide range of model resolutions and the consequent values of artificial model eddy diffusivities.

5. Outward eddy transport by inertia–gravity wave motions

The single effect neglected in these model studies is the possible parcel dispersion by the inertia–gravity wave field in the real lower stratosphere (e.g. Pierce *et al.* 1994). This should be roughly equivalent to a quasi-horizontal Fickian diffusivity (McIntyre and Pinhey 1995). The reason is that the intermittent breaking of these waves by Kelvin–Helmholtz instability will give rise to a small-step quasi-horizontal random walking of the typical molecules of any chemical tracer along, as well as across (Dewan 1981), isentropic surfaces. The conclusion is unaltered even if the vortex edge has a fine-scale screwthread structure, as was suggested schematically in figure 1, making vertical mixing by breaking inertia–gravity waves (Dewan 1981) possibly significant as well as quasi-horizontal mixing along isentropes by the same waves. Estimates of the effective diffusivities in the vicinity of the vortex edge put it well within the range of artificial diffusivities used in the numerical model studies of vortex dynamics, and strongly reinforces the conclusions about limited vortex erosion rates already drawn from those studies. McIntyre and Pinhey (1995) look at this question systematically.

It can be added that, in order to sustain large transport rates down through, as well as out from, the vortex, of the order of one vortex mass per month, diabatic descent rates within the lower-stratospheric vortex would need to be several times greater than seems compatible with observed temperatures and with very extensive studies in atmospheric radiation, whose physics is fundamentally well understood, (e.g. Valero et al 1993 & refs.). The predictions of such radiative studies receive independent support from direct, balloon-borne observations of descent of nitrous oxide isopleths within the Arctic polar vortex (Bauer *et al.* 1994). Nitrous oxide is chemically inert in the polar vortex and provides an excellent passive tracer of air motion. In the Antarctic diabatic descent can be expected to be, if anything, weaker than in the Arctic.

6. The stratospheric sub-vortex as a possible flowing processor

There is a transition altitude, usually at about 400 K, that appears to mark a transition between the stratospheric vortex and what will be called the stratospheric *sub-vortex* below. Like the vortex itself, the sub-vortex is also a region of low temperatures and dehydration, during late winter in the southern hemisphere and intermittently during winter in the northern hemisphere. But it is very different from the viewpoint of transport. It is also more massive because of the lower range of altitudes, below around 70 hPa. The distinction between the vortex and sub-vortex has been observationally clear for some time, from their different chemical signatures (e.g. Tuck 1989) and is consistent, moreover, with robust fluid-dynamical expectations based on the concept of 'PV inversion' (Hoskins *et al.* 1985).

The key point is that extratropical, synoptic-scale weather systems linked to strong, synoptic-scale PV anomalies on isentropic surfaces intersecting the tropopause, associated with distortions in the shape of the tropopause itself, have upward extensions into the stratosphere that tend to evanesce with altitude over height scales of a few kilometres. This upward evanescence is a robust property of the zonally asymmetric contributions to the wind, temperature and associated potential fields resulting from the inversion of synoptic-scale PV anomalies localised near the tropopause, mathematically similar to calculating the electrostatic potential due to localised charges. Such near-tropopause PV anomalies are very often part of the typical 'tropospheric' weather systems resulting from synoptic-scale cyclogenesis and anticyclogenesis in the extratropics; the PV anomalies near the tropopause tend to be strong, hence dynamically important, simply because isentropic gradients of PV tend to be strong near the tropopause. (Indeed it has long been found convenient, for reasons discussed in the review by Holton *et al.* (1995), to define the extratropical tropopause in terms of maximal isentropic gradients of PV; to this extent it is somewhat like the polar vortex edge.) The upward evanescence implies, then, that it is the lowermost isentropic surfaces in the lower stratosphere that are the most strongly stirred by the layerwise-two-dimensional motion induced by synoptic-scale weather systems. Reinforcing this effect is the cyclostrophically dictated vertical shear of the vortex itself, making it weaker and more stirrable with decreasing altitude.

This is the overwhelmingly likely fluid-dynamical explanation for the existence of the transition altitude at about 400K. According to this explanation (to which no alternative has been proposed, to my knowledge) the transition altitude is simply the altitude below which the layerwise-two-dimensional stirring is strong enough to overcome the Rossby quasi-elasticity of the vortex edge that would otherwise form. Any vortex edge that tries to form below the transition altitude, by diabatic or any other processes, is more or less stirred out of existence, depending on tropospheric weather activity. The same stirring is free to transport air parcels between the sub-vortex and the midlatitude regions on the same isentropic surfaces below the transition altitude. Such large-scale eddy transport can freely exchange polar and midlatitude air parcels without any requirement for a mean force field like that described in section 3; there is no single spinning air mass or coherent vortex to which angular-momentum and related arguments can be applied.

In this respect, the Arctic and Antarctic seem quite similar. Figure 2 clearly illustrates how such free exchange affects aerosols. The transition near 400 K is conspicuous.

An important corollary is that the sub-vortex could act as a dehydration and chlorine-activation site through which large masses of air could be efficiently transported, to and from middle latitudes at any time during winter. Such transport is indeed likely, since it would be brought about by the same large-scale eddy exchange processes that account for the existence of the transition altitude. This picture is strongly supported by further analyses of chemical tracer data (Jones & Kilbane-Dawe 1995).

Figure 3 shows an example of how vortex isolation increases with altitude, from Chen (1994), in an Antarctic case. In the real atmosphere the variation with altitude is probably stronger, since sparse observations in the Antarctic (and, in model studies, low model resolution) tend to result in an underestimation of the intensity of synoptic-scale tropospheric weather, hence of the stirring (Tuck 1994). Insight into the fundamental dynamics of Antarctic synoptic-scale tropospheric weather, and how it arises as baroclinic instabilities on a basic state driven, or strongly influenced by, the Antarctic katabatic surface drainage flow, is provided in an important recent paper by Juckes *et al.* (1994).

In summary, robust fluid-dynamical arguments based on the theory of potential-vorticity inversion, and on the cyclostrophically dictated downward weakening of vortex horizontal shear and quasi-elasticity, predict the existence of a transition altitude — more precisely a transition isentrope — which the observational evidence, exemplified by figure 2, and by chemical measurements (e.g. Tuck 1989, Jones & Kilbane-Dawe 1995), puts at potential temperatures around 400 K. The corresponding pressure altitudes are near 60 hPa at the centre of the vortex and near 80 hPa at its edge. The sub-vortex region below the transition isentrope could export activated air at a variable but potentially much larger rate than the vortex itself, especially in the Antarctic where temperatures are generally lower.

7. Postlude: the stratospheric overworld

By a strange accident, 400 K is also highly significant in the tropical stratosphere. Indeed, to a first approximation the 400 K isentropic surface divides the whole stratosphere rather simply into two very different transport regimes, an extratropical 'lowermost stratosphere', below about 400 K, and a global 'overworld' above about 400 K, the 'overworld' being further divided into tropical and extratropical parts by a subtropical eddy-transport barrier, weaker than the polar vortex edge but effective year round (Mote *et al.* 1996 & refs.). This division is another factor in the chemical differences seen above and below about 400 K, and shows more clearly the relevance of the extratropically pumped global-scale circulation to the problem of chemical 'stratosphere–troposphere exchange' (Holton *et al.* 1995).

I thank Dr Martin Wirth and Dr Ping Chen for kindly allowing me to reproduce figures 2 and 3, and Dr P H Haynes and Professor R. A. Plumb for important comments. This work received support from the Natural Environment Research Council through the UK Universities' Global Atmospheric Modelling Programme, from the Engineering and Physical Sciences Research Council through the award of a Senior Research Fellowship, and from the Isaac Newton Trust. The Centre for Atmospheric Science is a joint initiative of the Department of Chemistry and the Department of Applied Mathematics and Theoretical Physics.

Appendix. Gyroscopic pumping versus global-scale heating

The fluid dynamics of the extratropical pumping action, and its interplay with infrared radiative relaxation, has been the subject of many careful analytical and numerical modelling studies and has been thoroughly reviewed by Holton *et al.* (1995). It depends on the abovementioned wavelike and turbulent eddy motions, which for this purpose can usefully be thought of in terms of breaking Rossby and gravity waves, with 'breaking' understood in a suitably generalised sense (McIntyre and Palmer 1985; McIntyre 1992, 1993a). These eddy effects give rise to a persistently one-signed, or ratchet-like, irreversible global-scale transport of angular momentum, which produces the global-scale pumping via an interaction with strong Coriolis forces that can be described as 'quasi-gyroscopic'. It is quasi-gyroscopic in the sense that pushing air, say, westwards tends to make it move polewards. For brevity's sake let us simply call it 'gyroscopic'.

This gyroscopic pumping involves, of course, the kind of non-local influence found in most fluid-dynamical problems. As with an indrawn breath, or the air sucked in by ordinary domestic vacuum cleaner, the essential points are that constraints like mass conservation are crucial, and that associated influences may propagate via fast wave motions. For the indrawn breath, the relevant waves are ordinary acoustic waves. For the stratosphere, large-scale internal gravity and inertia–gravity waves are also relevant. The timescales of interest for global-scale chemical transport, say seasonal to decadal, are far longer than the wave propagation times in question. For practical purposes the concomitant non-local influence can be regarded as acting instantaneously. Further discussion of these points and of the modelling studies that demonstrate them is given in the review by Holton *et al.* (1995).

It is perhaps worth adding one further remark to that discussion. There is a widespread misconception that the rising branch of the global-scale stratospheric circulation is locally 'caused' by solar heating, as in a certain sense is true of the tropospheric Hadley circulation. But this is one of the ways in which the stratosphere differs from the troposphere. What is important for this purpose, in the stratosphere, is the far larger scale of the circulation and the far larger scale of the radiative heating and cooling, together with the relaxational character of the infrared contribution already mentioned. What this turns out to mean is that, under the conditions characteristic of the stratosphere, the main effect of solar heating is to raise temperatures rather than to cause any persistent global-scale circulation.

One reason for the misconception might be a tacit assumption that an equation of the form $A = B$, where B represents heating and A involves the circulation, implies a causal relation 'B causes A' notwithstanding that the equation could just as well be written $B = A$. (It is a similar misconception to conclude, from Newton's law 'fluid acceleration equals pressure-gradient force', that air moves toward the suction nozzle of a domestic vacuum cleaner 'because it is pushed by the pressure-gradient force'.) The similar misconception that solar heating 'causes' the global-scale stratospheric circulation tends, unfortunately, to be strengthened by the observed fact that the rising branch of the circulation has a slight tendency to follow the sun during the seasonal cycle. The rising branch leans slightly toward the summer side of the tropical upper stratosphere at solstice.

But the location of the rising branch, more or less within the tropical stratosphere, is enough in itself to demonstrate the incorrectness of the idea that the global-scale stratospheric circulation is caused by solar heating. This is because, throughout the greater part of the summer, the strongest diurnally averaged insolation and hence the strongest solar heating is nowhere near the tropics. Rather, it is at the summer pole. At solstice the insolation at the summer pole is, for instance, 39.5% higher than the average insolation over the tropical latitude band $\pm 20°$. A stratospheric circulation caused by, or dominated by, the pattern of solar heating would have to have its rising branch entirely over the summer pole at solstice, and not over the tropics at all.

It is easy to verify these points about insolation. For solar declination α relative to a spherical, rotating earth, the fractional length of day $\lambda(\phi)$ at latitude ϕ is

$$\lambda(\phi) = \pi^{-1} \arccos[\max\{-1, \min(1, -\tan\alpha \tan\phi)\}] ,$$

and the diurnally averaged vertical component of solar irradiance is the full solar irradiance multiplied by

$$S(\phi) = \lambda(\phi) \sin\alpha \sin\phi + \pi^{-1} \sin\{\pi\lambda(\phi)\} \cos\alpha \cos\phi .$$

It is easy to check first that the function $S(\phi)$ has an absolute maximum $S_{\max} = \sin\alpha$ at the north pole, $\phi = 90°$, whenever α is within $2.8°$ of its maximum solstitial value $\alpha_{\max} = 23.6°$, and second that, when $\alpha = \alpha_{\max}$, $S_{\max} = \sin\alpha_{\max}$ is 1.395 times the area average of $S(\phi)$, i.e. the average weighted by a factor $\cos\phi$, over the tropical latitude band $\pm 20°$.

Now visualise the declination α as $90°$minus the angle between two vectors one of which gives the direction of the Earth's axis, i.e. points toward the pole star, and the other of which points from the Earth toward the Sun at the time of interest. Projecting this picture on the Earth's orbital plane, and approximating the Earth's orbit as circular, we get $\sin\alpha = \sin\alpha_{\max} \cos\Delta t$ where Δt is the time after the northern summer solstice measured in units of (1 year)/(2π). The time during which insolation is maximal at the summer pole is then, in units of one year,

$$\pi^{-1} \arccos\{\sin(\alpha)/\sin(\alpha_{\max})\}$$

with $\alpha = 23.6° - 2.8°$, which is just over three-fifths of the summer season.

In summary, then, the relatively slight diversion of the rising branch of the stratospheric circulation toward the summer hemisphere shows just how weakly the stratospheric circulation pattern is influenced by solar heating, and how strongly by the gyroscopic pumping effects associated with the Earth's rapid rotation. It is precisely such gyroscopic pumping effects that single out the tropics as special rather than the summer pole. The seasonal cycle of the stratospheric circulation pattern is mainly due, therefore, to seasonality in the extratropical pumping.

Seen in this light, it is a strange coincidence — a perverse trick of Nature — that the circulation at higher, mesospheric altitudes should have a rising branch over the summer pole. That mesospheric circulation is in no way an exception to what has just been said. It is driven and controlled by gyroscopic pumping, due mainly to the violently breaking gravity waves in the lower thermosphere, above about 80 km. So strong is this pumping that the resulting refrigeration of the summer polar mesopause near 80–90 km makes it the coldest place on earth, as well as the sunniest place on earth, with temperature minima sometimes as low as 110 K. For further discussion the reader is referred to my (1992) review.

REFERENCES

Bauer, R., Engel, A., Franken, H., Klein, E., Kulessa, G., Schiller, C., Schmidt, U., Borchers, R., Lee, J., 1994 Monitoring the vertical structure of the Arctic polar vortex over northern Scandinavia during EASOE: regular N_2O profile observations. Geophys. Res. Lett., 21, 1211–1214.

Charney, J. G. & Drazin, P. G. 1961. Propagaatiion of planetary-scale disturbances from the lower into the upper atmosphere. J. Geophys. Res., 66, 83-109.

Chen, P., 1994 The permeability of the Antarctic vortex edge. J. Geophys. Res., 99, 20563–20571.

Dahlberg, S. P. & Bowman, K. P., 1994 Climatology of large-scale isentropic mixing in the Arctic winter stratosphere from analysed. J. Geophys. Res., 99, 20585-20599.

Dameris, M., Wirth, M., Renger, W., Grewe, V., 1995 Definition of the polar vortex edge by LIDAR data of the stratospheric aerosol: a comparison with values of potential vorticity. Beitr. Phys. Atmos., in press.

Dewan, E. M., 1981 Turbulent vertical transport due to thin intermittent mixing layers in the stratosphere and other stable fluids. Science, 211, 1041–1042.

Dritschel, D. G., 1989: Contour dynamics and contour surgery: numerical algorithms for extended, high-resolution modelling of vortex dynamics in two-dimensional, inviscid, incompressible flows. Computer Phys. Rep., 10, 78–146.

Fisher, M., O'Neill, A., Sutton, R., 1993 Rapid descent of mesospheric air into the stratospheric polar vortex. Geophys. Res. Lett., 20, 1267–1270.

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

Holton, J. R., Haynes, P. H., McIntyre, M. E., Douglass, A. R., Rood, R. B., Pfister, L., 1995 Stratosphere–troposphere exchange. Revs. Geophys. Space Phys., 33, 403–439.

Hoskins, B. J., McIntyre, M. E., Robertson, A. W., 1985 On the use and significance of isentropic potential-vorticity maps. Q. J. Roy. Meteorol. Soc., 111, 877–946. Corrigendum, etc., 113, 402–404.

Jones, R. L., Kilbane-Dawe, I., 1995 Observational evidence for the role of the stratospheric sub-vortex region in midlatitude ozone loss. Geophys. Res. Lett., submitted.

Juckes, M. N., McIntyre, M. E., 1987 A high resolution, one-layer model of breaking planetary waves in the stratosphere. Nature, 328, 590–596.

Juckes, M. N., James, I. N., Blackburn, M., 1994 The influence of Antarctica on the momentum budget of the southern extratropics. Q. J. Roy. Meteorol. Soc., 120, 1017–1044.

Junge, C., 1976 The role of the oceans as a sink for chlorofluoromethanes and similar compounds. Z. Naturforsch., 31a, 482–487.

Lahoz, W. A., O'Neill, A., Heaps, A., Pope, V., Swinbank, R., Harwood, R. S., Froidevaux, L., Read, W. G., Waters, J. W., Peckham, G. E., 1995: Vortex dynamics and the evolution of water vapour in the stratosphere of the Southern Hemisphere. Q. J. Roy. Meteorol. Soc., in the press.

Manney, G. L., Farrara, J. D., Mechoso, C. R., 1994a Simulations of the February 1979 stratospheric sudden warming: model comparisons and three-dimensional evolution. Mon. Wea. Rev., 122, 1115–1140.

Manney, G. L., Zurek, R. W., O'Neill, A., Swinbank, R., 1994b On the motion of air through the stratospheric polar vortex. J. Atmos. Sci., 51, 2973–2994.

McIntyre, M. E., 1989 On the Antarctic ozone hole. J. Atmos. Terrest. Phys., 51, 29–43.

McIntyre, M.E., 1992 Atmospheric dynamics: some fundamentals, with observational implications. Proc. Internat. School Phys. 'Enrico Fermi' CXV Course, *The Use of EOS for Studies of Atmospheric Physics,* edited by J. C. Gille and G. Visconti. North-Holland Amsterdam, pp. 313–386 (1992). ISBN 0444898964.

McIntyre, M. E., 1993a On the role of wave propagation and wave breaking in atmosphere-ocean dynamics, Sectional Lecture, Proc. XVIII Int. Congr. Theor. Appl. Mech., Haifa, ed. S. Bodner, J. Singer, A. Solan, and Z. Hashin. Amsterdam, Elsevier, 281–304.

McIntyre, M. E., 1993b Isentropic distributions of potential vorticity and their relevance to tropical cyclone dynamics. In: Proc. ICSU/WMO Internat. Symp. on Tropical Cyclone Disasters, Beijing, ed. J. Lighthill, Z. Zheng, G. Holland, K. Emanuel; Beijing, Peking University Press, 143–156.

McIntyre, M. E., 1994 The quasi-biennial oscillation (QBO): some points about the terrestrial QBO and the possibility of related phenomena in the solar interior. In: The Solar Engine and its Influence on the Terrestrial Atmosphere and Climate (Vol. 25 of NATO ASI Subseries I, Global Environmental Change), ed. E. Nesme-Ribes; Heidelberg, Springer-Verlag, 293–320.

McIntyre, M. E., Palmer, T. N., 1985: A note on the general concept of wave breaking for Rossby and gravity waves. Pure Appl. Geophys., 123, 964–975.

McIntyre, M. E., Pinhey, N. J. G., 1995 On the possible effects of inertia–gravity waves on the permeability of the stratospheric polar-vortex edge. J. Geophys. Res., in preparation.

Mo, R-P., McIntyre, M.E., and Bühler, O., 1995 On the mean mass flux across the edge of a stratospheric polar vortex due to thermally dissipating, steady nonbreaking Rossby waves. Q. J. R. Meteorol. Soc., to be submitted.

Mote, P. W., Rosenlof, K. H., McIntyre, M. E., Carr, E. S., Gille, J. C., Holton, J. R., Kinnersley, J. S., Pumphrey, H. C., Russell, J. M., Waters, J. W., 1996 An atmospheric tape recorder: the imprint of tropical tropopause temperatures on stratospheric water vapor. J. Geophys. Res., 101, 3989–4006.

Norton, W. A., 1994 Breaking Rossby waves in a model stratosphere diagnosed by a vortex-following coordinate system and a technique for advecting material contours. J. Atmos. Sci., 51, 654–673.

Norton, W. A., Chipperfield, M. P., 1995 Quantification of the transport of chemically activated air from the northern hemisphere polar vortex. J. Geophys. Res., in the press.

Pierce, R. B., Fairlie, T. D., 1993 Chaotic advection in the stratosphere: implications for the dispersal of chemically perturbed air from the polar vortex. J. Geophys. Res., 98, 18589–18595.

Pierce, R. B., Fairlie, T. D., Grose, W. L., Swinbank, R., O'Neill, A., 1994 Mixing processes within the polar night jet. J. Atmos. Sci., 51, 2957–2972.

Pyle, J. A., 1995 Ozone loss in middle latitudes and the role of the Arctic polar vortex. Phil. Trans. R. Soc. Lond. A (1995), 352, 241-245.

Tuck, A. F., 1989: Synoptic and chemical evolution of the Antarctic vortex in late winter and early spring, 1987. J. Geophys. Res., 94, 11687–11737. (Airborne Antarctic Ozone Experiment Special Issue.) Corrigendum, 94 (D14), 16855.

Tuck, A. F., 1994 The use of ECMWF products in stratospheric measurement campaigns. Proc. Workshop on the Stratosphere and Numerical Weather Prediction, pp. 73-105. Obtainable from the Librarian, European Centre for Medium-range Weather Forecasts, Shinfield Park, Reading RG2 9AX.

Tuck, A. F., Russell, J. M., Harries, J. E., 1993 Stratospheric dryness: antiphased dessiccation over Micronesia and Antarctica. Geophys. Res. Lett., 20, 1227–1230.

Valero, F. P. J., Platnick, S., Kinne, S., Pilewskie, P., Bucholtz, A., 1993 Airborne brightness temperatures of the polar winter troposphere as part of the AASE II and the effect of brightness temperature variations on the diabatic heating in the lower stratosphere. Geophys. Res. Lett., 22, 2575–2578.

Waugh, D. W., Plumb, R. A., Atkinson, R. J., Schoeberl, M. R., Lait, L. R., Newman, P. A., Loewenstein, M., Toohey, D. W., Avallone, L. M., Webster, C. R., May, R. D., 1993 Transport of material out of the stratospheric Arctic vortex by Rossby wave breaking. J. Geophys. Res., 99, 1071–1078.

Waugh, D. W., Plumb, R. A., 1994 Contour advection with surgery: a technique for investigating fine scale structure in tracer transport. J. Atmos. Sci., 51, 530–540.

FIGURE CAPTIONS

[Figures are in separate Z-compressed .ps and/or .gif files. *Warning:* the file for Fig 3 is large; it contains much fine detail.]

Figure 1. Rough schematic cross-section through the lower-stratospheric part of a polar vortex, updating figure 5 of McIntyre (1989) where it was pointed out (*a*) that the jagged edge is to be expected from the general properties of layerwise-two-dimensional turbulence in the surrounding 'Rossby-wave surf zone', and (*b*) that such a structure could explain some or all of the layering seen in balloon soundings from McMurdo Sound, Antarctica. McMurdo Sound is often near the vortex edge in the Antarctic winter.

Figure 2. Airborne lidar cross-section through the side and bottom of the Arctic polar vortex; compare bottom part of the left hand edge in figure 1. The lower part of the vortex, and the eddy-transport barrier constituting its edge, is made visible by the presence of the stratospheric aerosol layer or Junge layer, extending to its usual midlatitude altitudes of about 550 K, but largely excluded from the vortex interior (in right half of figure. Note that potential temperature is used as the vertical coordinate; kilometre altitudes shown on the right are only approximate. Grayscale values give the so-called backscatter ratio — a measure of aerosol mixing ratio, or mass of aerosol per unit mass of air — at a visible wavelength strongly scattered by typical stratospheric aerosol particles. From unpublished research extending the work of Dameris *et al.* (1995); courtesy of Dr Martin Wirth.

Figure 3. High-resolution tracer-advection pictures on different isentropic or stratification surfaces ((a) 350 K, (b) 375 K, (c) 400 K, (d) 425 Kon 31 August 1993) illustrating the difference between Antarctic stratospheric vortex and sub-vortex behaviour, see text, as computed from meteorologically analysed winds over 40 days from 21 July to 31 August 1993 (Chen 1994). Maps are polar-stereographic out to 30° latitude, with the 60° latitude circle shown. A high-resolution adaptive Lagrangian 'contour advection' technique (Dritschel 1989, Norton 1994, Waugh and Plumb 1994) is used to trace the material contour that lengthens least hence best represents the vortex edge. The sub-vortex may well be more strongly ventilated than suggested by (a), in some winters at least, both because of the tendency of meteorological analyses to underestimate synoptic-scale, weather-related disturbances in data-sparse regions, and also because of the interannual variability of tropospheric weather-related disturbances beneath. In the Arctic, the sub-vortex may be even more strongly ventilated, consistent with its slightly higher vertical penetration (415 K as opposed to the Antarctic's 400 K, Proffitt *et al.* 1990), and consistent with the aerosol distribution illustrated in figure 2.