BALANCE, POTENTIAL-VORTICITY INVERSION, LIGHTHILL RADIATION, AND THE SLOW QUASIMANIFOLD

M. E. McINTYRE
Centre for Atmospheric Science at the
Department of Applied Mathematics and Theoretical Physics,
Silver Street, Cambridge CB3 9EW, England.
http://www.atm.damtp.cam.ac.uk/people/mem/[†]

Abstract

Practically our entire understanding of large-scale atmosphere–ocean dynamics depends on the notions of balance and potential-vorticity inversion. These are essential, for instance, for a clear understanding of the basic Rossby-wave propagation mechanism, or quasi-elasticity, that underlies almost every large-scale fluid-dynamical phenomenon of meteorological and oceanographical interest, from the global-scale transport of terrestrial greenhouse gases (and similar problems in the solar interior) to Rossby-wave-mediated global teleconnection, baroclinic and barotropic shear instability, vortex coherence, and vortex-core isolation. The ideas involved in understanding balance and inversion continue to hold special fascination because of their central importance both for theory and for applications, such as data assimilation, and the fact that complete mathematical understanding is still elusive. The importance for applications was adumbrated by Richardson in his pioneering study of numerical weather prediction. The importance for theory — and the exquisite subtlety involved — was adumbrated by Poincaré in his discovery of the homoclinic tangle, and by Lighthill in his discovery of the quadrupole nature of acoustic radiation by unsteady vortical motion.

1. Introduction

I'll begin by quoting Lewis Fry Richardson's classic vignette — which *isn't* a limerick — following Jonathan Swift and Augustus De Morgan and summarizing the nature of three-dimensional turbulence:

[†]In press for Proc. IUTAM/IUGG/Royal Irish Academy Symposium on *Advances in Mathematical Modelling of Atmosphere and Ocean Dynamics* held at the Univ. of Limerick, Ireland, 2–7 July 2000, ed. P. F. Hodnett, pp. 45–68.

Big whirls have little whirls
That feed on their velocity;
Little whirls have lesser whirls,
And so on to viscosity.

This is what buffets us, and brings us Atlantic clear air, on a walk in the beautiful Irish hills; and Richardson was well aware of what can be learnt by casual observation of one's surroundings together with acute thinking. I think many of us who find fluid dynamics fascinating share a delight in such learning by observation, of anything from the power of a great storm seen from space or felt on the ground, to the progression of waves on an otherwise calm pond to the wavelike and vortical phenomena hints of which we see in stirred tea or coffee. Something that thrills me personally is that one can learn about fluid phenomena on a grand scale, up to global scales on our home planet and even vaster scales in the Sun's interior, from theory together with simple kitchen-sink experiments that even a theoretician like me can easily perform. It's all about what J. E. Littlewood called the "impudence" of creative thinking, or Erasmus Darwin the "damn fool experiment", which often fails but sometimes comes off brilliantly.

Jonathan Swift added, to the original well-known lines that inspired De Morgan and Richardson (about fleas, as you'll recall) the aphorism "Thus every poet, in his kind, is bit by him that comes behind." One might say that the same thing happens to scientists, even though for the sake of the scientific ethic we try to be more polite about it. I don't think Richardson's vignette of turbulence would be disputed as far as it goes, despite all the sophistication of fractal modelling and data analysis that refines the picture today. But the nature of large-scale atmosphere–ocean turbulence is in some ways at an opposite extreme:

Big whirls meet bigger whirls,
And so it tends to go on:
By merging they grow bigger yet,
And bigger yet, and so on.

And that's not all because, co-existing and competing with the tendencies toward vortex merging and upscale energy cascade, there are yet more, very different processes. One of these, as Rhines (1975) may have been the first to point out in a turbulence-related context, is Rossby-wave propagation. The eddying flows conspicuous in today's satellite images of the atmosphere and ocean do indeed, very often, look turbulent in a perfectly real sense. The flow is manifestly chaotic, and it stirs and mixes tracers, properties often quite reasonably taken to be characteristic of turbulence. But as soon as you have an important wave propagation mechanism — and there are several such mechanisms in stratified rotating fluid systems — you have a very profound difference by comparison with any classical turbulence picture, even if lots of tracer mixing is going on at the same time and in the same locations.

That difference is central to understanding global-scale atmospheric circulations, and hence to understanding, for instance, the whole business of chlorofluorocarbon (CFC) lifetimes and other problems concerning the ozone layer, the distribution of greenhouse gases, and their role in the entire Earth System.

Why should the presence of wave propagation mechanisms make such a "profound difference", and what's wrong with classical pictures of turbulence? The answer is fundamental: wave propagation changes everything qualitatively. (This is changing our ideas about the Sun's differential rotation as well; see Gough and McIntyre (1997).) As soon as you have a wave propagation mechanism you are liable to have, by that very fact, systematic correlations between fluctuating fields, of a kind that are completely neglected in classical turbulence theories, and in most modern turbulence theories as well. Such correlations profoundly alter the nature of the fluctuation-induced momentum transport, and with it the global circulation dynamics. Instead of momentum transport over the relatively short ranges characterizing eddy-induced material displacements, of the order of classical "mixing lengths", you suddenly have a long-range momentum-transport mechanism, limited only by the distances over which waves can propagate.

You can then have phenomena like anti-frictional behaviour. This means fluctuations that on average drive the system not toward, but away from, solid rotation: upgradient momentum transport, if you will. The most conspicuous illustrations are the quasi-biennial oscillation (QBO) of the zonal winds in the equatorial stratosphere, and the celebrated laboratory experiment of Plumb and McEwan (1978), which produces a similar phenomenon in a very simple way.

The Plumb–McEwan experiment illustrates anti-frictional behaviour in the most clearcut and unequivocal way imaginable: stratified flow in an annulus with symmetry axis vertical is driven away from solid — in this case zero — rotation, entirely by fluctuations of relatively high frequency, though still low enough to feel the internal gravity wave propagation mechanism. The fluctuations are excited simply by standing oscillations of the top or bottom boundary. The imposed conditions are mirror-symmetric (to reflection in a vertical mirror), but that symmetry is spontaneously broken and a strongly sheared differential rotation arises, first one way and then the other, reversing again and again on a much longer timescale and displaying a characteristic spacetime signature in which shear zones move systematically upward or downward, toward the fluctuating boundary that is the sole means of driving the system.[1]

Such anti-frictional behaviour is impossible in a classical turbulence scenario. But it is perfectly easy to understand in terms of the long-range momentum-transport — reaching far beyond mixing-length range — associated with the generation of waves in one place and their dissipation in another. Further discussion would be too much of a digression here; let me just say that the tendency toward anti-frictional behaviour, and the relevance to it of wave breaking, wave dissipation, critical layers and so on, have been discussed systematically in a recent review

[1]The experiment has recently been repeated at Kyoto University, with results available in the form of a movie at http://www.gfd-dennou.org/library/gfd_exp/ where, in addition, "technical tips" are given on how to run the experiment successfully.

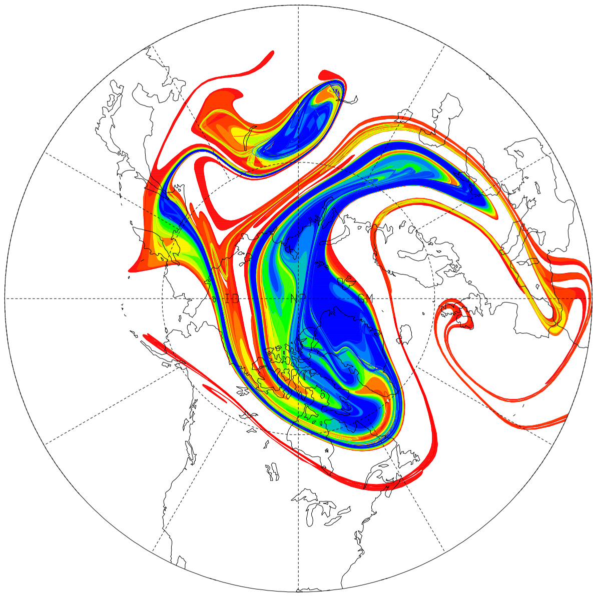

Figure 1. Contour-advective plot of polar-vortex air in the real lower stratosphere on 28 January 1992, from the work of Plumb *et al.* (1994) and D. W. Waugh (personal communication). Similar plots emphasizing midlatitude instead of polar air show strong mixing along the isentropic (stratification) surfaces, very like that seen in Figure 2.

(McIntyre 2000) in which special care is taken to distinguish, among other things, critical-layer myths from critical-layer realities.

And, as already hinted, there is more to the whole story than just the QBO and similar phenomena. Wave-induced momentum transport is now recognized as central to the fluid dynamics of greenhouse-gas-transporting atmospheric circulations, a crucial part of the Earth system, or climate system if you will. On a rapidly rotating planet such as the Earth you can, and do, have an almost ubiquitous, global-scale "gyroscopic pumping" remotely driven by wave-induced momentum transport. This is a persistent mechanical pumping action that drives, for instance, the stratospheric Brewer–Dobson circulation. Wave-induced mean forces push stratospheric air persistently westward, and Coriolis effects persistently turn the air poleward. Air is pulled up in the tropics and pushed back down in higher latitudes, against thermal radiative relaxation.

The Brewer–Dobson circulation is basic to understanding the behaviour of the ozone layer, and its part in the Earth system and in shaping the detectable and attributable patterns of climate change. The strength and persistence of the circulation accounts for many aspects of greenhouse-gas behaviour, for instance explaining why CFC lifetimes are of the order of 10^2 years. CFCs, which would otherwise take millennia or more to be removed from the atmosphere, are transported by the Brewer–Dobson circulation to altitudes above about 25 km, where they are destroyed relatively rapidly by hard solar ultraviolet.

And in the case of the stratospheric Brewer–Dobson circulation, where our understanding and its observational underpinning are now very secure, it is clear that the wave propagation mechanism chiefly responsible is the Rossby-wave mechanism.

2. The Rossby-wave jigsaw

In view of the above, I was astonished to hear someone say earlier this week, at this Symposium, that some proposal or paper had been rejected on the grounds that "Rossby waves do not exist in the real atmosphere or ocean". We actually know quite a lot about the real Rossby waves that propagate or diffract up from the real troposphere into the real stratosphere, where they dissipate through breaking and infrared radiative damping and drive the real Brewer–Dobson circulation — whose reality is itself consistent with many lines of chemical tracer evidence, all the way from Brewer's historic water-vapour measurements of 1949 (Norton *et al.* 2000) to today's beautiful visualizations of the tropical upwelling branch through the "tape recorder effect", in which the annual water-vapour cycle is imprinted on the air as it rises at about $0.2\,\mathrm{mm\,s^{-1}}$ (Mote *et al.* 1996, 1998). The Rossby mechanism is also central to our understanding of such phenomena as vortex coherence, barotropic and baroclinic shear instabilities and their nonlinear saturation, dynamical teleconnections, and indeed nearly all large-scale fluid motions of oceanographic and meteorological interest (*e.g.* Hoskins *et al.* 1985).

Much of today's knowledge of real, finite-amplitude Rossby waves began with the theoretical work of Rossby (1936), Ertel (1942), Charney (1948) and Kleinschmidt (1950–1) on the concept of potential vorticity (PV) and its inversion, together with observational studies of isentropic distributions of PV in the real stratosphere, beginning with a "damn fool experiment", in Erasmus Darwin's sense, that was done in the early 1980s at the UK Meteorological Office.

Having been working on some relevant theory at the time (McIntyre 1982, Killworth and McIntyre 1985), I became closely involved in interpreting that experiment together with one of the Met Office's young luminaries, Dr Tim Palmer (McIntyre and Palmer 1983–5). The experiment was foolish in Erasmus' sense because it attempted to use satellite data to visualize isentropic distributions of Rossby–Ertel PV in the middle stratosphere. In those days atmospheric researchers were already well aware of the problems with satellite data retrieval, and there was one version of the folklore saying that anyone with the temerity — the impudence — to try to compute, from satellite data, such a highly differentiated quantity as the Rossby–Ertel PV was, not to put too fine a point on it, a fool.

Nevertheless, the experiment defied the folklore and worked. It gave us our first glimpses of the PV distributions at altitudes around 30 km, in the midwinter extratropical stratosphere. We likened what we saw to "a blurred view of reality seen through... knobbly glass." Idealized theoretical models, especially the Stewartson–Warn–Warn (SWW) model of a Rossby-wave critical layer, helped us to make sense of that blurred view; and we were able to do enough cross-checks, for instance from Lagrangian trajectory computations — I remember the labour of doing some of

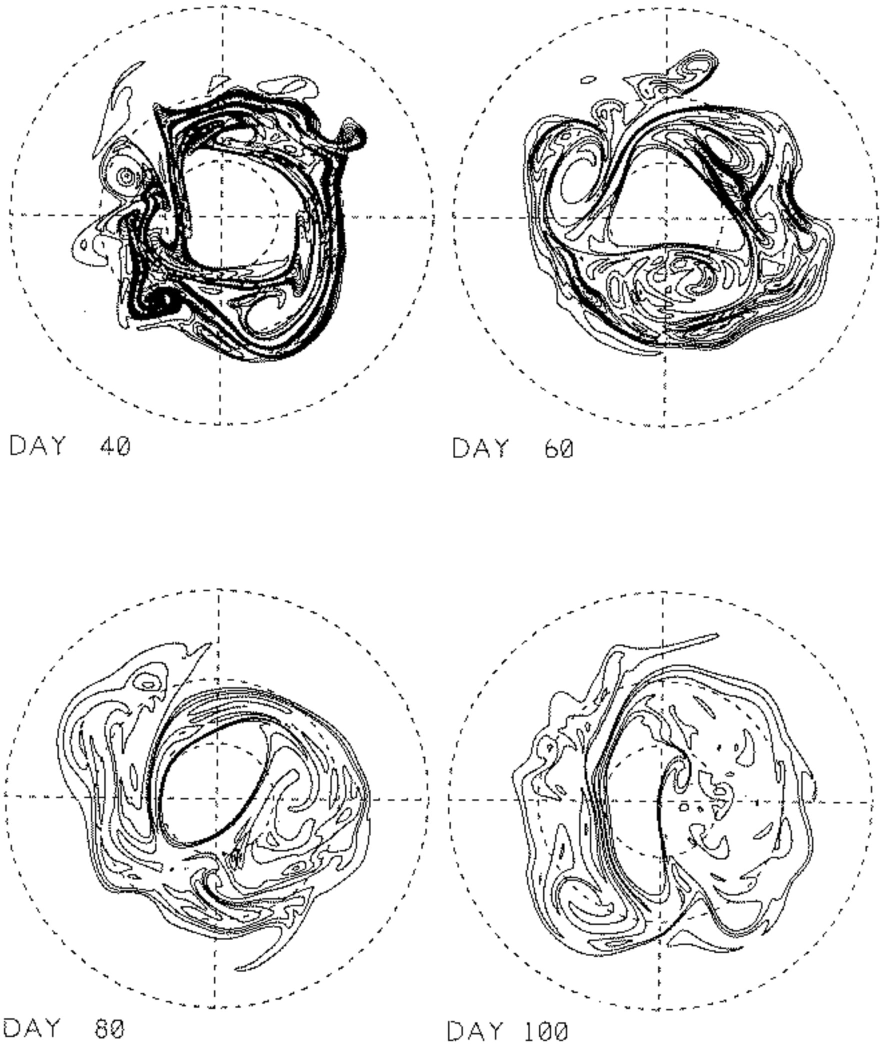

Figure 2. Shallow-water model flow on the sphere, an animation of which was shown at the Symposium, closely resembling flow in the real wintertime stratosphere at altitudes around 25 or 30 km. The map projection is conformal (polar stereographic), with the equator and the 30°N and 60°N latitude circles shown dashed. The flow is visualized by passive tracer released as a compact blob into the midlatitude stratospheric surf zone, clearly showing the fast two-dimensional turbulent mixing in that region, despite which the stratospheric polar vortex remains almost completely isolated from its surroundings, and likewise, to a lesser extent, the tropics (cf. Mote *et al.* 1998). The isolation of the (core of the) polar vortex recalls classic smoke rings and is of great importance to stratospheric polar chemistry, including the Antarctic ozone hole and its (so far less severe) Arctic counterpart. The isolation is due to the combined effects of the Rossby-wave restoring mechanism and the strong shear just outside the edge. Courtesy of Dr W. A. Norton, from whom an animated video of the model run is available (Dept of Atmospheric, Oceanic, and Planetary Physics, Clarendon Laboratory, Parks Road, Oxford, OX1 3PU, UK). Details of the model and the model run are given in Norton (1994). The mean depth is $h_0 = 4$ km, giving behaviour qualitatively like that implied by Equations (20) and (25) below, with the Rossby length $L_0 \sim 2000$ km in middle latitudes, and quantitatively very close to that implied by Equations (10)–(19) below.

them by hand — to add up to a convincing picture. The rest is history. The original experiment was enough to to make conspicuous the highly inhomogeneous "wave–turbulence jigsaw puzzle" that is typical of real, finite-amplitude Rossby-wave fields. The midlatitude stratosphere at altitudes around 30 km revealed itself as a gigantic "surf zone" driven by the "world's largest breaking waves." The surf zone corresponds to the theoretician's critical layer, whose supposed narrowness was thus revealed as one example of a critical-layer myth.

The same gigantic surf zone and spatial inhomogeneity is now seen routinely, and much more clearly, through clever combinations of data analysis and adaptive Lagrangian computational techniques such as "contour advection" (Norton 1994, Waugh and Plumb 1994), or "domain-filling trajectories" (Manney *et al.* 1994, Lahoz *et al.* 1996), applied to state-of-the-art meteorological wind analyses, themselves the result of operational data assimilation into large numerical models. In addition, the picture has been confirmed again and again by global-scale images of chemical tracer distributions remotely sensed from space, an outstanding recent example being some remarkably detailed CFC-11, methane, and nitrous oxide images from the helium-cooled CRISTA spectrometer flown on the Space Shuttle (Riese *et al.* 2001).

Figure 1 shows a contour-advection example from the work of Plumb *et al.* (1994), for the lower stratosphere at around 18 km altitude. This is a passive-tracer picture emphasizing polar-vortex air, constructed by contour advection over 4 days (see figure caption) on the 450 K isentropic surface. Some of the fine detail in pictures like these has been directly verified by *in situ* airborne chemical measurements (Waugh *et al.* 1994). Figure 2 shows another example, this time from a shallow-water model (Norton 1994), a full animation of which was shown as a video at the Symposium. This model behaves in a manner astonishingly like the real wintertime stratosphere at altitudes anywhere between about 20 and 40 km, as revealed for instance by the CRISTA images, which cover a similar range and an example of which is shown in Figure 3; note the different map projection emphasizing polar regions.

In the model example of Figure 2, whose bottom-right panel may be qualitatively compared to Figure 3, a small blob of passive tracer released in middle latitudes quickly fills the well developed surf zone, a layerwise-two-dimensional turbulent region sandwiched between relatively isolated polar-vortex and tropical regions. The video animation of Figure 2 makes especially vivid the wavelike, quasi-elastic behaviour of the polar-vortex edge under the Rossby-wave mechanism.[2] It is a clear example of what is now called an "eddy transport barrier", almost completely preventing surf-zone air from penetrating into the polar vortex region, a matter of some chemical importance.

From a dynamical perspective the wavelike and turbulent regions — the quasi-elastic vortex edge and the adjacent surf zone — are closely interdependent, as the term "jigsaw" is meant to suggest. That interdependence, a manifestation of nonlinear dynamics, is illustrated most plainly and explicitly by the SWW critical-

[2] Copies of the video are obtainable from Dr W. A. Norton at the Department of Atmospheric Physics, Clarendon Laboratory, Parks Road, Oxford, OX1 3PU, UK; see also Norton (1994).

layer model, of which our understanding is comprehensive (Killworth and McIntyre 1985, Haynes 1989). The model applies in a different parameter regime — in which the surf zone or critical layer is, in fact, narrow — but has the virtue that one can precisely quantify the dynamical interdependence of the surf zone and its more wavelike surroundings through matched asymptotic expansions.

The upshot of all these observational and theoretical studies, then, is that we may think of the main mass of polar-vortex air as being held together quasi-elastically by the Rossby mechanism,[3] with help from differential advection by the shear just outside — a good example of what is also called "vortex coherence" — even though some of the polar airmass is being eroded and mixed into the surrounding surf zone. The erosion and mixing involve the irreversible deformation of "otherwise-wavy" material contours, *i.e.* those contours, also PV contours in this case, that would undulate reversibly under the conditions implicitly assumed by linearized wave theory.

Irreversible deformation of such otherwise-wavy contours violates the nonacceleration theorem of wave–mean interaction theory, a consequence of Kelvin's circulation theorem applied to such contours. This makes the contour-deformation process fundamental to understanding the circumstances under which irreversible wave-induced momentum transport occurs. When the contour deformation is irreversible the process may therefore, very reasonably in this context, be designated as "Rossby-wave breaking", and a comparison drawn with the more familiar breaking of ocean-beach waves, which leads to the convergence of wave-induced momentum transport and its well-known consequence near ocean beaches, the generation of longshore currents. The big long tongue of polar air in Figure 1, curving away from the main polar airmass toward Mediterranean Europe, is quite like what Palmer and I originally saw "through knobbly glass" and is one of the typical large-scale Rossby wave-breaking patterns seen in practice. By contrast, the quasi-elastic region near the edge of the vortex core — showing up especially clearly in the case of Figure 2 — is marked by a set of material and PV contours that to good approximation undulate reversibly, as Norton (1994) showed in detail through high precision contour-advection calculations. Such reversible undulations are in stark contrast with the irreversible behaviour of the contours just outside.[4]

It will have been noticed that the extreme spatial inhomogeneity that shows up so vividly in all these cases, including the SWW model, represents another big departure from classical turbulence-theoretic scenarios. Those scenarios assume statistical homogeneity or near-homogeneity. The real spatial inhomogeneity is a very robust feature, showing up not only again and again in observations such

[3]See Figure 4. I am aware that "Rossby waves" might more aptly be called "Kelvin waves", especially in the context of vortex coherence. Though no historian of science, I have the impression that if you see someone's name on something there is a more than even chance that someone else thought of it first. There are always good reasons. "Kelvin waves" are usually understood to refer to Coriolis-trapped gravity waves, a different animal entirely. So in speaking of "Rossby waves" I am only using what has become the standard terminology, tending to displace the more logical alternatives "vorticity waves" and "potential-vorticity waves".

[4]McIntyre and Palmer (1984) discuss the time-reversal "paradox" involved here in the use of the term "irreversible".

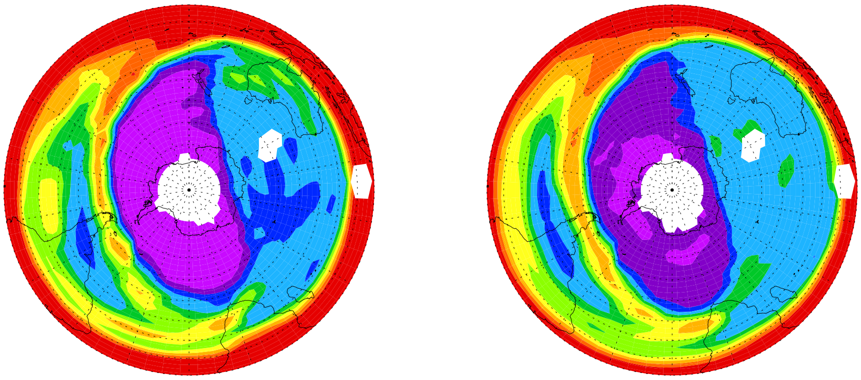

Figure 3. N_2O (nitrous oxide) mixing ratios observed at two stratospheric altitudes on 11 August 1997 by the CRISTA instrument, from Riese *et al.* (2001), for the southern hemisphere but mirror-flipped, in this display, for qualitative comparison to Figure 2 bottom right, which has positive Coriolis parameter $f > 0$ as appropriate to the northern hemisphere. Notice the different map projections, with polar regions emphasized here but tropical regions in Figure 2. Irregular white areas are data gaps. On Rossby-wave timescales of days and weeks N_2O is an accurate passive tracer, though destroyed photochemically on Brewer–Dobson timescales of years. In the right half of each picture the N_2O mixing ratios increase nearly monotonically or stepwise monotonically outward (being nearly constant over large areas in the surf zone). They increase from polar-vortex values close to zero to large tropical values imported from the troposphere by the Brewer–Dobson upwelling. **At left and right respectively,** pressure-altitudes are 4.64 hPa and 10 hPa, roughly 37 km and 31 km; ranges of mixing ratios in parts per billion by volume are 0–90+ and 0–150+ with contour intervals 10 and 16.67, where "+" signifies that values may slightly overshoot the plotted range. The lightest band at the subtropical edge of the surf zone highlights the ranges 60–70 and 100–116.67 ppbv. CRISTA (Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere) detects a number of chemical species through their infrared spectral signatures and is a large (1350 kg) helium-cooled instrument flown from the Space Shuttle.

as Figure 3 but also in model simulations, all the way from idealized shallow-water models to the big three-dimensional models now widely used in studies of stratospheric chemistry and ozone depletion.

3. PV inversion

V. I. Arnol'd once stated that "Hamiltonian mechanics cannot be understood without differential forms" (Arnol'd 1978). In the same spirit I would say that Rossby waves cannot be understood without PV inversion. An understanding worth having will include, of course, an understanding of the nonlinear effects, such as the "jigsaw dynamics", the dynamical interdependence of vortex-edge undulation and surf-zone turbulence.

Rossby waves are often discussed without mentioning the idea of PV inversion at all — and if you are content with an idealized background state and infinitesimal wave amplitude then it's easy enough to calculate wavelike solutions and say that's all there is to it — but the idea of PV inversion is always there implicitly, right from the moment one tries to go further and understand even the linearized wave dynamics intuitively. I think it clarifies one's thinking to make the inversion

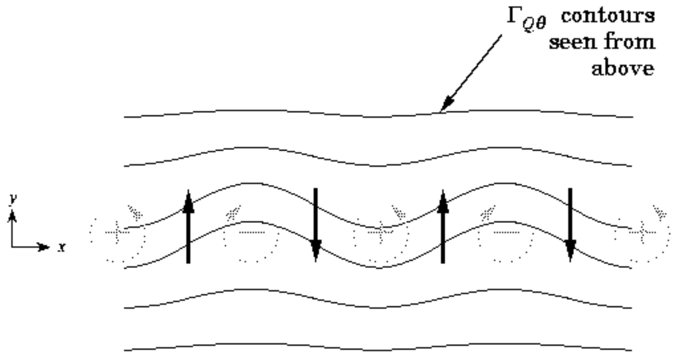

Figure 4. Visualization of the Rossby-wave propagation mechanism, redrawn from Hoskins *et al.* (1985) and McIntyre (2000).

idea explicit. Moreover, the resulting insight does help to bridge the gulf between idealized models and complex, nonlinear reality. I feel it's part of one's job as a theoretician to try to discern what is robust about an idealized model, and hence what aspects are likely to carry over to the real world and what aspects are not.

Thus when, for instance, the edge of the polar vortex is distorted, as illustrated in the video corresponding to Figure 2, this means that the PV, which has steep gradients across the vortex edge and is approximately a materially conserved, or advected, quantity, develops a pattern of anomalies relative to the undistorted circular shape of the edge. This pattern of PV anomalies is qualitatively as shown by the curved arrows and $\pm$ signs in Figure 4, if we read the coordinate x as azimuthal distance around the edge, periodically repeating as appropriate. The figure assumes that the background PV gradient is positive in the positive y-direction. The sideways displacement of material and PV contours on an isentropic surface — the edge distortion itself — is like the displacement of the wavy contours marked $\Gamma_{Q\theta}$, in the central part of Figure 4.

PV invertibility says that PV inversion is possible. This means that, under reasonable assumptions about boundary conditions, you can *deduce* the velocity field from the PV field. Equations (1)–(2) below provide the simplest illustration of this. In particular, any x-periodic pattern of PV anomalies like that in Figure 4 *implies* an x-periodic pattern of velocity anomalies a quarter wavelength out of phase with the displacement pattern, as suggested by the straight arrows. But if you have velocities a quarter wavelength out of phase with displacements, then you can directly see, by making a mental movie of the way in which the $\Gamma_{Q\theta}$ contours are advected, that the velocity field *makes* the undulations propagate, in this case in the $-x$ direction, and in all cases to the left of the background isentropic gradient of PV.

In summary, displacements produce PV anomalies, which (via PV inversion) imply velocities a quarter wavelength out of phase with displacements, which produce propagation in whichever direction has high PV values on the right. This is the basic Rossby-wave mechanism. It is basic to any stratified fluid motion that has gradients of PV on stratification surfaces (or buoyancy-acceleration gradients acting as concentrated PV gradients at upper or lower boundaries), and for which invertibility holds (Hoskins *et al.* 1985, & refs.). It is therefore basic to most large-scale fluid motions of oceanographic and meteorological interest; for instance, if you put opposite-signed PV gradients next to each other you will often get barotropic or baroclinic shear instabilities, in which counterpropagating Rossby waves phase-lock and make each other grow.

I think everyone is familar with the world's simplest example of the Rossby-wave mechanism, in the context of two-dimensional, nondivergent barotropic vortex dynamics in the original Rossby beta-plane or "approximately flat Earth" model. In this model, invertibility reduces simply to solvability of a Poisson equation $\nabla^2\psi = (Q - f)$, under boundary conditions such as evanescence at infinity. Symbolically,

$$\psi \;=\; \nabla^{-2}(Q - f)\ . \tag{1}$$

Here f is the Coriolis parameter, and $Q(x, y, t)$ is the PV, which in this example is the same as the two-dimensional absolute vorticity. The stream function $\psi(x, y, t)$ describes a strictly nondivergent velocity field $\mathbf{u}(x, y, t)$,

$$\mathbf{u} \;=\; (u, v)\ , \qquad u \;=\; -\,\partial\psi/\partial y\ , \qquad v \;=\; \partial\psi/\partial x\ , \tag{2}$$

where (x, y) are eastward and northward Cartesian coordinates and (u, v) the corresponding components of the velocity vector $\mathbf{u}(x, y, t)$. The inversion is a diagnostic process; no time derivatives appear in (1)–(2), and so one can talk about inversion at a single instant t, independently of neighbouring t values.

In this dynamical system there is just one evolution equation,

$$DQ/Dt \;=\; 0\ , \tag{3}$$

where D/Dt is the two-dimensional material derivative,

$$D/Dt \;:=\; \partial/\partial t + \mathbf{u}\cdot\nabla \;=\; \partial/\partial t + u\partial/\partial x + v\partial/\partial y\ . \tag{4}$$

We can think of the dynamical system as a limiting case of real stratified flow in which friction goes to zero and the stratification becomes infinitely strong, constraining the motion to be exactly horizontal. This picks out the Coriolis parameter f as precisely the vertical component of twice the Earth's angular velocity, implying that f increases monotonically with northward distance y like the sine of the latitude. In taking such a limit, one must assume that vertical scales of motion stay finite. Then the dependence on the vertical coordinate z becomes a mere parametric dependence, with no vertical derivatives $\partial/\partial z$ appearing in the problem. The equations at each altitude z then reduce to Equations (1)–(4) above.

The standard Rossby-wave theory for this system — linearize equation (3) about rest, set $\beta = df/dy =$ constant, look for solutions $\psi = \Re\hat{\psi}\exp(ikx{+}ily{-}i\omega t)$ ($\hat{\psi}$ and ω constant, k and l real, constant) so that $\nabla^{-2} \;=\; -(k^2 + l^2)^{-1}$ and

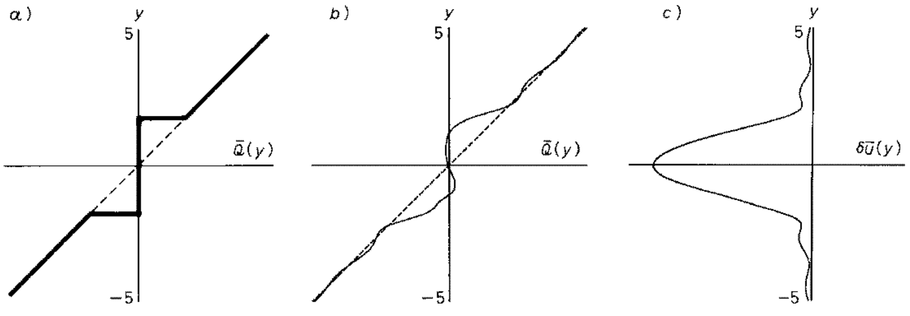

Figure 5. The relation between mean wave-induced force and Q-rearrangement by a breaking Rossby wave, in the simplest relevant model system, the dynamical system (1)–(4). Courtesy P. H. Haynes; for mathematical details see Killworth and McIntyre (1985) and Haynes (1989). Plot (a) shows idealized Q distributions before and after mixing Q in some y-interval or latitude band; (b) shows the x-averaged Q distribution in an actual model simulation using equations (1)–(4); (c) shows the resulting mean momentum deficit, equation (5), whose profile would take on a simple parabolic shape in the idealized case corresponding to (a).

$\omega = -\beta k/(k^2+l^2)$ (real) — confirms the qualitative picture suggested by Figure 4. The y-velocity $\partial\psi/\partial x$, with complex amplitude $ik\hat{\psi}$, is a quarter wavelength out of phase with the displacement, with complex amplitude $(-i\omega)^{-1}ik\hat{\psi}$. The chirality — the one-way character of the propagation, tied to the sense of the Earth's rotation — is expressed by the single power of ω in the dispersion relation $\omega = -\beta k/(k^2+l^2)$.

The association between Rossby-wave breaking and irreversible wave-induced angular momentum transport, and the dynamical interdependence of vortex-edge undulation and surf-zone turbulence in general, are likewise well illustrated by the dynamical system (1)–(4). The PV inversion operator (1) is non-local, so it's obvious from the outset that there has to be some such dynamical interdependence. And for angular momentum transport we may consider, for instance, an idealized thought experiment in which Rossby waves arrive and form a surf zone by perfectly mixing the PV in some band of y-values, as shown on the left of Figure 5. The mean change $\delta\bar{Q}(y)$ in $\bar{Q}$ is a linear function within that band, integrating to zero across the band. Inversion according to (1) with reasonable boundary conditions gives an x-averaged mean flow change

$$\delta\bar{u}(y) \;=\; \int_y^\infty \delta\bar{Q}(\tilde{y})\,d\tilde{y} \;\leqslant\; 0\,. \qquad (5)$$

This is a parabolic profile qualitatively like the $\delta\bar{u}(y)$ profile shown on the right of Figure 5, which is from a less idealized, fully detailed chaotic critical-layer calculation due to Haynes (1989), with the $\delta\bar{Q}(y)$ profile shown in the middle. In both cases there is a robustly negative net momentum change. This is a phenomenon in which wavelike and turbulent motion are both inextricably involved.[5]

It is also an example of anti-frictional behaviour, because $\delta\bar{u}(y)$ has to be added to a background flow $\bar{u}(y) \propto y$ with positive shear $\partial\bar{u}(y)/\partial y > 0$ (not shown in the figure), and the wave source is located at positive y, outside the domain of the figure. Thus the momentum has been transported against its own mean gradient.

The point of all this is that more realistic models have just the same generic structure, and qualitative properties, provided only that we have PV invertibility. In particular, if we go back to realistic finite stratification, then the dynamical system still has the single evolution equation (3) with its single time derivative, expressing chirality, together with an inversion operator that behaves qualitatively like that in (1). It implies the same qualitative description of Rossby-wave propagation as before, including the one-way character associated with the single time derivative and the single power of ω in the dispersion relation. It implies the same qualitative association between surf-zone formation and irreversible wave-induced momentum transport, robustly one-signed. This has been verified explicitly by the work of Robinson (1988), on stratified flow in spherical geometry.

PV invertibility is helpful, moreover, in coping with the central difficulty in fluid dynamics. This is the advective nonlinearity, as expressed by the $\mathbf{u}\cdot\nabla$ terms in the Euler equations. If the main effects of $\mathbf{u}\cdot\nabla$ are captured by the D/Dt in a single evolution equation of the simple form (3), then we have the simplest conceivable way of visualizing and understanding the effects of the advective nonlinearity. It is in all these ways that the qualitative insight from the world's simplest model does, indeed, carry over to more sophisticated models of reality.

The discussion up to this point is one way of showing why the ideas of PV inversion and PV invertibility are important, and therefore, by implication, why the ideas of balance and "slow manifold" are important. As is well known, the same ideas are crucial to an understanding of how to initialize weather-prediction models (*e.g.* Lynch and McGrath, this Proceedings), a matter whose importance was first revealed by Richardson's pioneering attempt to integrate the equations numerically.

4. Accurate balance and PV inversion

As is well known, PV invertibility, and weather-prediction initialization, both depend on imposing some balance condition. Imposing such a condition amounts to artificially constraining the dynamical system to some prescribed "slow manifold" within its full phase space. In the system (1)–(4), the balance condition is simply the incompressibility condition (2). It implies an absence of sound waves. In more

[5]Results like (5) have sometimes been thought to show that the PV mixing illustrated at left and middle of Figure 5 is impossible (*e.g.* Stewart and Thomson 1977) since, the argument goes, momentum conservation would be violated. This underlines the danger of neglecting the wavelike aspects of the jigsaw puzzle, *i.e.* of thinking solely in terms of classical turbulence paradigms and hence missing the possibility of momentum being imported from elswhere via wave-induced momentum transport. It is just this possibility, or rather actuality, that is so cogently illustrated by the SWW and related critical-layer solutions such as those of Haynes (1989).

realistic systems, from shallow-water to fully-stratified, condition (2) is replaced by a functional relation of the form

$$\mathbf{u}(\mathbf{x}) \;=\; \mathbf{u}^{\mathrm{B}}\,(\mathbf{x}; h(\cdot)) \;, \tag{6}$$

where explicit reference to time t is suppressed for the moment, and where $h(\cdot)$ symbolically represents the mass configuration of the fluid system at a given instant t. In a shallow-water system, for instance, $h(\cdot)$ is shorthand for the height $h(\mathbf{x}) = h(x, y)$ of the free surface above some horizontal reference level. The balance condition (6) says that the velocity field is completely determined, at each instant t, by a knowledge of the mass configuration — the spatial distribution of fluid mass throughout the physical domain at that instant. It says that the velocity field no longer represents independent degrees of freedom in phase space. The condition (6) confines the system to a submanifold within phase space. A simple albeit crude illustration is provided by the familiar geostrophic relation, defined for the shallow-water case by

$$\mathbf{u}(\mathbf{x}) \;=\; \mathbf{u}^{\mathrm{B}}\,(\mathbf{x}; h(\cdot)) \;=\; \mathbf{u}^{\mathrm{G}}\,(\mathbf{x}; h(\cdot)) \;:=\; \frac{g}{f}\left(-\frac{\partial h(\mathbf{x})}{\partial y}\,,\;\frac{\partial h(\mathbf{x})}{\partial x}\right)\;, \tag{7}$$

where g is the gravitational acceleration. The associated PV inversion operator is defined by (7) together with a definition of the PV and suitable boundary conditions, such as evanescence at infinity. Once again, no time derivatives appear: we have a diagnostic process, defined at each instant t. For the definition of PV we may use for instance the shallow-water PV discovered by Rossby (1936), who showed it to be an exact material invariant for frictionless flow (*op. cit.*, Eq. (75)). In the simplest case of a flat bottom identified with the horizontal reference level, the Q of §3 is then replaced by the exact Q defined by Rossby's formula[6]

$$Q \;=\; \frac{1}{h}\left(f + \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) \;=\; \frac{1}{h}(f+\zeta)\;, \quad \text{say.} \tag{8}$$

The resulting boundary-value problem has variable coefficients but is elliptic and robustly invertible, like (1) above, at least when Froude and Rossby numbers are sufficiently small:

$$F \;=\; \sup(|\mathbf{u}|/c) \;\ll\; 1\;; \qquad R \;=\; \sup(|\zeta|/f) \;\ll\; 1\;. \tag{9}$$

Here $c = \surd(gh)$, the shallow-water gravity-wave speed. Replacing Q by a linearized counterpart based on writing $h' = h - h_0$ and neglecting squares $(h'/h_0)^2$ and products $(h'/h_0)(\zeta/f)$, with h_0 constant, gives the standard "quasi-geostrophic" version of the theory in which the PV inversion problem has constant coefficients.

Now balance conditions far more accurate than (7) are known. For instance we may replace (7) by the following set of diagnostic equations, to be solved (at

[6]Rossby's formula is still valid for a sloping bottom, of course, provided that h is temporarily redefined as layer depth instead of surface elevation, in which case (7) is modified appropriately.

each instant t) for $\mathbf{u}^{\mathrm{B}}$ when a suitable $h'(\mathbf{x})$ field is given. It is convenient to introduce the Helmholtz decomposition of $\mathbf{u}^{\mathrm{B}}$ into its solenoidal and irrotational contributions, under suitable boundary conditions,

$$\mathbf{u}^{\mathrm{B}} \;=\; \mathrm{curl}^{-1}\zeta^{\mathrm{B}} + \mathrm{div}^{-1}\delta^{\mathrm{B}} \tag{10}$$

where ζ^{B} and δ^{B} are the corresponding vorticity and divergence defined respectively as $\partial v^{\mathrm{B}}/\partial x - \partial u^{\mathrm{B}}/\partial y$ and $\partial u^{\mathrm{B}}/\partial x + \partial v^{\mathrm{B}}/\partial y$. In the remaining equations, $c_0 := \sqrt{(gh_0)}$; the quantities with subscripts 1 and 2 are auxiliary fields to be explained shortly:

$$\begin{aligned}
f\zeta^{\mathrm{B}} &= g\nabla^2 h' + \nabla\cdot(\mathbf{u}^{\mathrm{B}}\cdot\nabla\mathbf{u}^{\mathrm{B}}) + \delta_1^{\mathrm{B}}\ , && (11)\\
\delta^{\mathrm{B}} &= \mathcal{L}^{-1}\nabla\cdot\Big\{f\zeta^{\mathrm{B}}\mathbf{u}^{\mathrm{B}} + \mathbf{u}_{\mathbf{1}}^{\mathrm{B}}\cdot\nabla\mathbf{u}^{\mathrm{B}} + \mathbf{u}^{\mathrm{B}}\cdot\nabla\mathbf{u}_{\mathbf{1}}^{\mathrm{B}} - g\nabla^2(h'\mathbf{u}^{\mathrm{B}})\Big\}\,, && (12)\\
\delta_1^{\mathrm{B}} &= \mathcal{L}^{-1}\nabla\cdot\Big\{f\zeta_1^{\mathrm{B}}\mathbf{u}^{\mathrm{B}} + f\zeta^{\mathrm{B}}\mathbf{u}_{\mathbf{1}}^{\mathrm{B}} + \mathbf{u}_{\mathbf{2}}^{\mathrm{B}}\cdot\nabla\mathbf{u}^{\mathrm{B}} + 2\mathbf{u}_{\mathbf{1}}^{\mathrm{B}}\cdot\nabla\mathbf{u}_{\mathbf{1}}^{\mathrm{B}} \\
&\qquad + \mathbf{u}^{\mathrm{B}}\cdot\nabla\mathbf{u}_{\mathbf{2}}^{\mathrm{B}} - g\nabla^2(h_1^{\mathrm{B}}\mathbf{u}^{\mathrm{B}} + h'\mathbf{u}_{\mathbf{1}}^{\mathrm{B}})\Big\}\ , && (13)\\
\mathbf{u}_{\mathbf{1}}^{\mathrm{B}} &= \mathrm{curl}^{-1}\zeta_1^{\mathrm{B}} + \mathrm{div}^{-1}\delta_1^{\mathrm{B}}\ , && (14)\\
\mathbf{u}_{\mathbf{2}}^{\mathrm{B}} &= \mathrm{curl}^{-1}\zeta_2^{\mathrm{B}}\ , && (15)\\
\zeta_1^{\mathrm{B}} + f\delta^{\mathrm{B}} &= -\nabla\cdot(\mathbf{u}^{\mathrm{B}}\zeta^{\mathrm{B}})\ , && (16)\\
\zeta_2^{\mathrm{B}} + f\delta_1^{\mathrm{B}} &= -\nabla\cdot(\mathbf{u}_{\mathbf{1}}^{\mathrm{B}}\zeta^{\mathrm{B}} + \mathbf{u}^{\mathrm{B}}\zeta_1^{\mathrm{B}})\ , && (17)\\
h_1^{\mathrm{B}} + h_0\delta^{\mathrm{B}} &= -\nabla\cdot(\mathbf{u}^{\mathrm{B}}h')\ . && (18)
\end{aligned}$$

Here

$$\mathcal{L} \;:=\; c_0^2\left(\nabla^2 - {L_0}^{-2}\right)\ , \tag{19}$$

a modified Helmholtz operator depending on the natural length scale

$$L_0 \;:=\; c_0/f\ , \tag{20}$$

the Rossby length or "radius" based on c_0. No time derivatives appear anywhere in (10)–(18). When $h'(\mathbf{x})$ is given, these are nine equations (six scalar and three vector) to determine, diagnostically, the nine unknown functions $\zeta_2^{\mathrm{B}}(\mathbf{x})$, $\zeta_1^{\mathrm{B}}(\mathbf{x})$, $\zeta^{\mathrm{B}}(\mathbf{x})$, $\delta_1^{\mathrm{B}}(\mathbf{x})$, $\delta^{\mathrm{B}}(\mathbf{x})$, $h_1^{\mathrm{B}}(\mathbf{x})$, $\mathbf{u}_{\mathbf{2}}^{\mathrm{B}}(\mathbf{x})$, $\mathbf{u}_{\mathbf{1}}^{\mathrm{B}}(\mathbf{x})$, and $\mathbf{u}^{\mathrm{B}}(\mathbf{x})$. The quantities with subscripts 1 and 2 are diagnostic estimates of first and second partial time derivatives. Being diagnostic quantities, they must be sharply distinguished from actual rates of change in a model integration. The point is disussed more fully in my paper with Norton (2000, hereafter MN00), along with related issues concerning local mass conservation and Galilean invariance. If $\delta_1^{\mathrm{B}}(\mathbf{x})$ were to be replaced by $\partial\delta/\partial t$ and the superscripts B deleted from the other variables, then (11) would become the divergence equation of the exact shallow-water equations; (16) similarly corresponds to the exact vorticity equation and (18) to the exact mass-conservation equation. Note also that if every term on the right of (11) were to be deleted except the first, then (11) would reduce to the curl of the geostrophic relation (7). Equations (12) and (13) are derived from the first and second time derivatives of the divergence equation with the leading time derivatives deleted. Standard scale-analytic considerations would argue that these deleted terms are relatively small if F and R are small.

When (8) is appended to (10)–(19), one obtains the PV inversion operator that MN00 called a "third order direct" inversion operator. It is exquisitely accurate, over an astonishingly wide range of values of F and R, as MN00 demonstrate for complicated shallow-water vortical flows on a hemisphere. (See Appendix A of MN00 for notes on the numerical procedures and for the counterpart of (10)–(19) in spherical coordinates, taking account of variable f.) Not only the vorticity but also the divergence field are reconstructed in considerable detail from a knowledge of the PV alone. Almost incredibly, this accuracy is obtained despite R being infinite at the equator, and F reaching values in excess of 0.7 in some cases. One can hardly say that ∞ and 0.7 are small. Trying to carry out a PV inversion for such parameter values certainly counts as another "damn fool experiment" — and most of the credit for it is due to Dr Warwick Norton, who was my research student at the time, and who showed marvellous intellectual courage as well as considerable computational ingenuity. The exquisite accuracy comes at a price, of course: it depends on subtle, weakly nonlinear corrections that vitiate the superposition principle and demand elaborate iterative numerical methods.

Even more accurate inversion operators can be defined, based on normal mode expansions. For further discussion see MN00 and Mohebalhojeh and Dritschel (2001). Figure 6 shows an example taken from MN00 — which still astonishes me even though it was first obtained a decade ago — again in a hemispherical domain and with F again exceeding 0.7 . This is a cumulative accuracy test, using the PV-conserving balanced model defined by a normal-mode-based PV inversion operator together with the single prognostic equation $DQ/Dt = 0$, for shallow-water flow on the hemisphere, slightly modified with a ∇^6 hyperdiffusion to control numerical noise at discretization scales. The top half of Figure 6 shows two PV fields from a 10-day run of the balanced model, and the bottom half the corresponding fields from a carefully initialized run of the exact shallow-water equations, serving as the benchmark of accuracy. This is a complicated, highly unsteady vortical flow exhibiting hyperbolicity or phase-space sensitivity. See MN00 for evidence of that sensitivity and for the precise specification of the inversion operator. Even after 10 days or several eddy times, and some nontrivial vortex interactions including merging, the two PV distributions track each other almost perfectly.

Now there is indeed something truly mysterious about such accuracy. Standard order-of-magnitude arguments say that we have no right to expect the concepts of balance and inversion, and the resulting balanced models, to be accurate unless F and R are small. Accurate balance and inversion involve nonlocal functional relations, through operators such as curl^{-1}, div^{-1}, and $\mathcal{L}^{-1}$, as the notation $\mathbf{u}^{\mathrm{B}}(\mathbf{x}; h(\cdot))$ in (6) was meant to suggest. Just as with (1), such relations imply, so to speak, action-at-a-distance. A change in the Q value here influences the velocity over there; and it does so instantaneously. In the simple system (1)–(4) this is reasonable, because the balance condition (2) makes the speed of sound infinite. But $F > 0.7$ means that even the fastest inertia–gravity waves are barely faster, in terms of group velocity, than relative fluid velocities. PV inversion is indeed

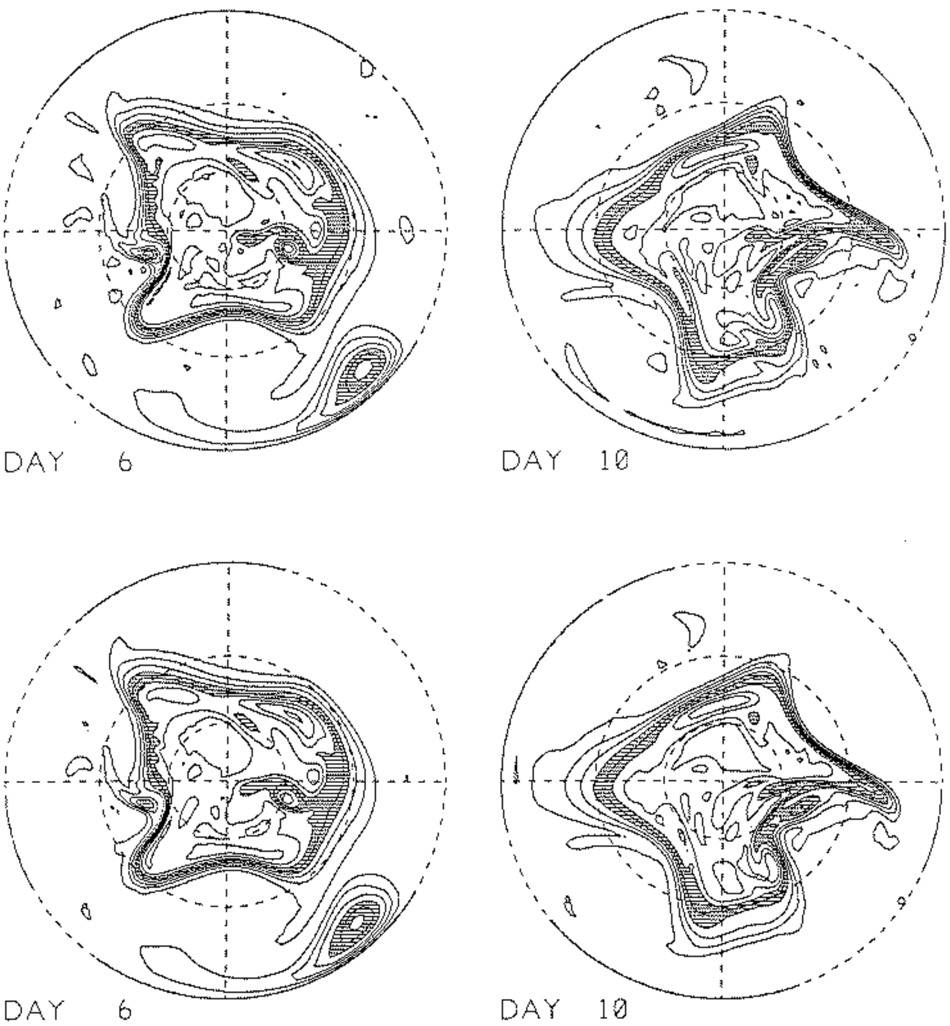

Figure 6. PV fields showing cumulative accuracy over 10 days (several eddy times) of a sophisticated "nonlinear normal mode" PV inverter for shallow-water flow on a hemisphere. The top pair are from a 10-day run of a PV-conserving balanced model based on this PV inverter; the bottom pair are from the corresponding run of the full shallow-water equations. Parameter regime is realistic for upper-tropospheric flow, with velocities in the tens of $\mathrm{m\,s^{-1}}$ and gravity-wave speed $c_0 = \sqrt{(gh_0)} = 100\,\mathrm{m\,s^{-1}}$, with h_0 defined as area-averaged layer depth. The PV is defined as in (8). The contour interval is $3\times10^{-8}\,\mathrm{m^{-1}\,s^{-1}}$; the shading marks values lying between 4 and 6 of these units. The two sets of PV fields are astonishingly close to each other, almost down to the discretization scales that feel the model ∇^6 hyperdiffusion, despite the infinite Rossby number R at the equator and the value, just over 0.7, reached by the Froude number F as defined by (9).

"foolish" under such conditions; and the experiment of Figure 6, daring to test cumulative accuracy, is more "foolish" still.

5. Poincaré's homoclinic tangle and the slow quasimanifold

So let us draw a deep breath and stand back a moment. As has often been pointed out, balanced motion has its counterparts in simpler dynamical systems. One of these is the "springy pendulum" consisting of a mass suspended from a pivot by a stiff elastic spring. Such a pendulum has slow, rotating or swinging, modes of motion in which the relatively fast, compressional, modes of the mass and spring are hardly excited. A first approximation in describing such slow motions is simply to set the length of the spring to be constant. This might be compared to the incompressibility condition (2). More accurate approximations would allow the spring to change its length in a quasi-static way. The fast modes are then, so to speak, "slaved" to the slow modes. In the fluid system this is another way of viewing the effect of a balance condition $\mathbf{u}(\mathbf{x}) = \mathbf{u}^{\mathrm{B}}\,(\mathbf{x}; h(\cdot))$, be it the geostrophic relation (7) or any of its more accurate counterparts such as (10)–(19). Such approximations and their ultimate limitations can be studied mathematically via techniques ranging all the way from formal two-timing (multiple scales) to bounded-derivative theory and, in the pendulum case at least, KAM (Kolmogorov–Arnol'd–Moser) theory and other dynamical-systems techniques for finite phase spaces. There is an enormous literature (*e.g.* Bokhove and Shepherd 1996 & refs.).

Basic to understanding what is involved is Henri Poincaré's picture of the "homoclinic tangle". This is a dynamical-systems classic that is now textbook material; it applies unequivocally, for instance, to the pendulum started slowly from its upside-down equilibrium position. In the phase space of a simple, idealized *rigid* pendulum, there is a homoclinic orbit representing one complete circuit of the pendulum, taking a logarithmically infinite time. Poincaré's homoclinic tangle (between perturbed stable and unstable manifolds, whose formidable fine-grain complexity Poincaré himself did not even attempt to draw, and which there is no room to discuss adequately here) tells us, in effect, that if we nonlinearly couple the simple rigid pendulum to practically any other oscillator — such as that associated with the springiness — or even just persistently jiggle the suspension point in a prescribed, deterministic way, then the single homoclinic orbit will break up into a "stochastic layer" or "chaotic layer" having finite though possibly small thickness. In KAM language, this homoclinic orbit is a torus that always breaks up under perturbation. This is intuitively reasonable. A logarithmically infinite time is available, which is plenty of time for practically any slight disturbance to nudge the pendulum into neighbouring parts of phase space, and enormously large numbers of ways for it to do so — depending on initial conditions on the fast modes, such as the states of the spring, or on the imposed jiggling. So if one replaces the exact springy-pendulum motion by a balanced model of it (in which information about the fast initial conditions or the jiggling is lost) then one is replacing the stochastic layer — which we might call a *slow quasimanifold* to emphasize that it is *not* a single invariant manifold of the system — by an artificially imposed, infinitesimally thin slow manifold. In effect, this artifice suppresses all the uncertainty due to the lost information.

By analogy, we expect that the accuracy seen in Figure 6 cannot mean that the exact shallow-water equations naturally have a true “slow manifold”, a precisely invariant manifold within the full phase space whose existence implies the possibility of exact “superbalance”, exact PV inversion, and therefore (incredibly) exact action-at-a-distance despite finite inertia–gravity wave speeds. Rather, even though strict mathematical proof in this infinite-dimensional problem seems well beyond anyone’s capability, the pendulum analogy leads us to expect that exact shallow-water motion close to balance — such as the motion corresponding to the bottom half of Figure 6 — must have a trajectory within a thin, though not infinitesimally thin, slow quasimanifold or infinite-dimensional stochastic layer within the infinite-dimensional phase space of the exact shallow-water equations.

6. Lighthill radiation — geostrophic it’s not!

It is here that James Lighthill’s work enters the picture. Lighthill’s pioneering thinking about aerodynamic sound generation, the emission of sound waves by, in his case, three-dimensional vortical motion such as turbulent jets, can be adapted to our problem even though there are some nontrivial technicalities (Ford *et al.* 2000). Lighthill’s thinking beautifully complements Poincaré’s homoclinic-tangle argument. It gives us a profound physical insight (a) showing, quite independently of Poincaré’s argument, why one expects a slow quasimanifold rather than a true slow manifold (in the infinite-dimensional phase space of the exact fluid equations), and (b) some idea of why the slow quasimanifold should be so remarkably thin in the fluid case, as Figure 6 emphasizes, certainly thinner than any simple order-of-magnitude consideration could possibly predict.

Lighthill’s essential insight is contained in the phrases “quadrupole radiation” and “destructive interference”. It tells us (a) that unsteady vortical motions, like those illustrated in Figures 1–3 and 6 of this paper, will, in continually adjusting toward balance, emit inertia–gravity waves spontaneously but (b) that such “spontaneous-adjustment emission” is weak because of destructive interference. This begins to explain why PV inversion can work better than it has a right to, and why any approach via simple order-of-magnitude analysis, or any other approach neglecting the full subtlety of the problem, will tend to overestimate the strength of spontaneous-adjustment emission.

The adjective *spontaneous*, incidentally, is essential for clarity here. Spontaneous-adjustment emission is not to be confused, as sometimes occurs in the literature, with inertia–gravity wave emission due to what is today called “Rossby adjustment”. Rossby or initial-condition adjustment is the process famously encountered by Richardson, in his first attempt at numerical weather prediction. It is simply and straighforwardly the emission of inertia–gravity waves due to imbalanced initial conditions, artificially imposed. There is nothing subtle here. If you kick the system, you will excite inertia–gravity waves. I am also avoiding the term “geostrophic adjustment” because — when the adjustment is spontaneous — it is more likely, if anything, to be adjustment away from geostrophic balance (7), $\mathbf{u} = \mathbf{u}^{\mathrm{G}}(\mathbf{x}; h(\cdot))$, as adjustment toward it. It is likely to be toward a more accurate balance $\mathbf{u} = \mathbf{u}^{\mathrm{B}}(\mathbf{x}; h(\cdot))$. Part of the trouble may be that the term "geostrophic" is often used not in its precise technical sense, $\mathbf{u} = \mathbf{u}^{\mathrm{G}}(\mathbf{x}; h(\cdot))$, but rather as a word to mean anything vaguely to do with the general concept of balance. The next thing that happens, of course, is that the word gets used in both senses at once!

For our rotating systems, the appropriate generalization of Lighthill's original argument is as follows. Again we use the shallow-water equations with constant f. The momentum and mass-conservation equations are taken in flux form:

$$\frac{\partial}{\partial t}(hu_i) + \frac{\partial}{\partial x_j}(hu_iu_j) - \varepsilon_{ij} f h u_j + \frac{g}{2}\frac{\partial}{\partial x_i}(h^2) = 0 \,, \tag{21}$$

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x_i}(hu_i) \;=\; 0 \,, \tag{22}$$

where u_i is the ith Cartesian component of $\mathbf{u}$ ($i = 1, 2$), and ε_{ij} is the two-dimensional alternating tensor, defined by $\varepsilon_{12} = -\varepsilon_{21} = 1$ and $\varepsilon_{11} = \varepsilon_{22} = 0$. A little manipulation (Ford *et al.* 2000) produces

$$\left(\mathcal{L} - \frac{\partial^2}{\partial t^2}\right)\frac{\partial h}{\partial t} \;=\; -\,\frac{\partial^2}{\partial x_i \partial x_j} T_{ij} \,, \tag{23}$$

where $\mathcal{L}$ is the modified Helmholtz operator (19) and where

$$T_{ij} \;=\; \frac{\partial}{\partial t}(hu_iu_j) + \frac{f}{2}(\varepsilon_{ik} h u_j u_k + \varepsilon_{jk} h u_i u_k) + \frac{g}{2}\frac{\partial}{\partial t}(h'^2)\,\delta_{ij} \,. \tag{24}$$

Equation (23) has the form of the linear inertia–gravity wave operator $\mathcal{L} - \partial^2/\partial t^2$ acting on $\partial h/\partial t$, on the left-hand side, and an *apparent wave source* consisting of the nonlinear terms on the right-hand side.

The essential point noted in Lighthill (1952) is that, in the case $F \ll 1$, the right-hand side is known to good approximation from the vortical flow alone, and can therefore be regarded as a *given* source of inertia–gravity waves.

The only essential assumption is that corrections $O(F)$ or weaker cannot change the qualitative character of the right-hand side. It is crucial, however — and this was Lighthill's most important point — that any approximate representation of the vortical flow be first substituted into T_{ij} before the differentiations $\partial^2/\partial x_i \partial x_j$ are carried out. This is because the weakness of the radiation depends on the cancellation, or destructive interference, already mentioned. It is the second-derivative form of the right-hand side of (23), rather than the precise form of T_{ij} itself, that is crucial to the cancellation and corresponds to the celebrated "quadrupole radiation". It is the weakness of quadrupole radiation, in other words, that begins to account for the possible accuracy of PV inversion. It can be added that the introduction of Coriolis effects should weaken the radiation still further, because of the cutoff at the inertia frequency $\omega = f$. Further discussion may be found in Ford *et al.* (2000) and in Saujani and Shepherd (2001).

Of course Lighthill's argument leaves unanswered the question "How does the wave emission remain weak even when the Froude number F is not small?" It is not obvious from Lighthill's argument why such a thing should come about. Here there is still considerable mystery. Part of the answer seems to be that the PV inversion operator takes on more and more of a *short-range character* as the Froude number F increases. Quasi-geostrophic theory, though quantitatively inaccurate, is enough to give a qualitative feel for this point. Within that theory — based, as mentioned earlier, not only on approximating $\mathbf{u}^{\mathrm{B}}(\mathbf{x}; h(\cdot))$ by (7) but also on neglecting squares $(h'/h_0)^2$ and products $(h'/h_0)(\zeta/f)$ in Q, replacing it by a linearized counterpart Q_{qg} say — it is well-known that the PV inversion operator becomes essentially $\mathcal{L}^{-1}$, where $\mathcal{L}$ is again the modified Helmholtz operator defined in (19), relevant to cases in which the Coriolis parameter f is exactly or approximately constant. The resulting PV inversion sets the divergence to zero and can be written explicitly, in an unbounded xy-domain, in terms of the stream function

$$\psi = -\frac{1}{2\pi}\iint K_0\Big(\frac{|\mathbf{x}-\mathbf{x}'|}{L_0}\Big)\,Q_{\mathrm{qg}}(\mathbf{x}')\,d^2\mathbf{x}'\;; \tag{25}$$

$K_0(\cdot)$ is the modified Bessel function with exponential decay at large argument. This represents a short-range interaction, because of the exponential tail of the Bessel function, and is such that the range decreases in proportion to L_0 and in inverse proportion to F, for given typical velocities $|\mathbf{u}|$, as F increases.

So even though the inertia–gravity waves cannot travel infinitely fast in order to mediate the action-at-a-distance implied by balance and PV inversion, the influence has a shorter and shorter distance to travel, as F increases and L_0 decreases.

Another consequence of this shortening interaction distance is that the vortex dynamics becomes distinctly more sluggish or slow moving. Thus the spontaneous-adjustment emission, which depends on unsteadiness of the vortex dynamics, is in this problem weaker, in most circumstances, than Lighthill's scaling laws would suggest. This must be another part of the explanation of the near-miracle of Figure 6, in which it should be noted that the Rossby length $L_0 \sim 1000\,\mathrm{km}$ in middle latitudes. For further insight the reader is referred to a paper by Ford (1994), presenting some very careful numerical experiments on Lighthill radiation.

7. Concluding remarks

It has sometimes been argued that spontaneous-adjustment emission is not properly described as "Lighthill emission" or "Lighthill radiation" because, for instance, a domain such as that of Figure 6 isn't really infinite, as in Lighthill's original problem or its rotating counterpart studied by Ford (1994) and Ford *et al.* — or, again, because in a rotating system the emission is a lot weaker than Lighthill's original (non-rotating) scaling laws would say. I would argue that that misses Lighthill's main point. It comes down to saying that in *any* situation where the spontaneous-adjustment emission of inertia–gravity waves is in fact weak, for any reason, one can regard the right-hand side of (23) as known to leading order and thus make a conceptual separation between vortical motion and wave emission. This point seems to me to be quite independent of whether the domain is bounded or unbounded or whether the Coriolis parameter f is zero or nonzero. It is important because of the help it gives in beginning to understand the ubiquity yet weakness of spontaneous-adjustment emission, hence the physical cause of slow-quasimanifold fuzziness and the remarkable thinness of the slow quasimanifold in at least some cases of interest such as that of Figure 6.

I want to end by mentioning briefly some recent developments in the Hamiltonian theory of balanced motion that I have been involved in. The unapproximated shallow water system, and analogous stratified, rotating fluid systems, are Hamiltonian in the classical sense. Ways of making this explicit mathematically are now very well known. In three pioneering papers, Salmon (1983, 1985, 1988) took the first steps toward developing systematic procedures to construct balanced models from their exact "parents" that inherit the parent Hamiltonian structure, including all the associated conservation relations such as that for PV. Recently, Ian Roulstone of the U.K. Meteorological Office and I have succeeded in simplifying and clarifying these procedures in such a way as to make plain a number of generic or universal properties of such Hamiltonian balanced models (McIntyre and Roulstone 1996, 2001). One of these is a property we call "velocity splitting", which can be viewed as an immediate consequence of imposing the balance condition $\mathbf{u} = \mathbf{u}^{\mathrm{B}}(\mathbf{x}; h(\cdot))$.

We find it mnemonically useful to say that the imposition of the balance condition — for any choice of the functional $\mathbf{u}^{\mathrm{B}}(\mathbf{x}; h(\cdot))$, accurate or inaccurate — splits the parent velocity field into two distinct velocity fields, whose difference provides a natural intrinsic measure of the inaccuracy of the model. Because of Lighthill radiation, this inaccuracy can never be zero, no matter how delicately one tries to refine the balance condition, save in a tiny set of exceptional special cases where the vortex dynamics is steady and the Lighthill radiation vanishes.

Recently, Mohebalhojeh and I have shown that velocity splitting is not peculiar to Hamiltonian balanced models (paper in preparation). It is a generic property of all accurate *non-Hamiltonian* balanced models as well, with just one class of exceptions. The most important member of that class — an exception that has long diverted attention from what is generic — is the balanced model called the Bolin–Charney "balance equations" in the form described by Gent and McWilliams (1984) and Whitaker (1993). Our result is essentially that all PV-conserving balanced models significantly more accurate than the Bolin–Charney model, including for instance the balanced model defined by (10)–(19) above, must suffer velocity splitting *as a direct price for their accuracy.*

And now I am out of time and out of space! How can I capture something of all this in a limerick, as seems to be *de rigueur* at this Symposium? Well, here goes:

On balance, consider the angle
Of the pendulum's upside-down dangle:
  Geostrophic it's not,
  But I don't care a lot,
'Cause it's all in a bit of a tangle.

Acknowledgements: Special thanks are due to my colleagues Rupert Ford, Peter Haynes, Björn Haßler, Ali Mohebalhojeh, Philip Mote, Warwick Norton, Dirk Offermann, Alan Plumb, Martin Riese, Ian Roulstone, Jürgen Theiss, Darryn Waugh, and Christof Wunderer, all of whom have contributed in important ways to the most recent developments described or mentioned in this all-too-brief survey, which barely does justice to recent advances. Darryn Waugh kindly allowed me to reproduce Figure 1 from his ozone-layer research, and Martin Riese Figure 3. This work received generous 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 Institute's Programme in the Mathematics of Atmosphere and Ocean Dynamics. Finally, I want to thank Frank Hodnett and his Organizing Committee for inviting me here and for an unusual, delightful, and interesting Symposium.

References

Arnol'd, V. I., 1978: *Mathematical Methods of Classical Mechanics*, New York: Springer-Verlag, 462 pp.

Bokhove, O., and Shepherd, T. G., 1996: On Hamiltonian balanced dynamics and the slowest invariant manifold. *J. Atmos. Sci.*, **53**, 276–297.

Charney, J. G., 1948: On the scale of atmospheric motions. *Geofysiske Publ.*, **17(2)**, 3–17.

Ertel, H., 1942: Ein Neuer hydrodynamischer Wirbelsatz. *Met. Z.*, **59**, 271–281.

Ford, R., 1994: Gravity wave radiation from vortex trains in rotating shallow water, *J. Fluid Mech.*, **281**, 81–118.

Ford, R., *et al.*, 2000: Balance and the slow quasimanifold: some explicit results, *J. Atmos. Sci.*, **57**, 1236–1254.

Gent, P. R., McWilliams, J. C., 1984: Balanced models in isentropic coordinates and the shallow water equations. *Tellus*, **36A**, 166–171.

Gough, D. O., McIntyre, M. E., 1998: Inevitability of a magnetic field in the Sun's radiative interior. *Nature*, **394**, 755–757.

Haynes, P. H., 1989: The effect of barotropic instability on the nonlinear evolution of a Rossby-wave critical layer, *J. Fluid Mech.*, **207**, 231–266.

Hoskins, B. J., *et al.*, 1985: On the use and significance of isentropic potential-vorticity maps, *Q. J. Roy. Meteorol. Soc.*, **111**, 877–946; *Corrigendum*, **113**, 402–404.

Killworth, P. D., McIntyre, M. E., 1985: Do Rossby-wave critical layers absorb, reflect or over-reflect? *J. Fluid Mech.*, **161**, 449–492.

Kleinschmidt, E., 1950–1: Über Aufbau und Entstehung von Zyklonen (1–3 Teil). *Met. Runds.*, **3**, 1–6; 3, 54–61; 4, 89–96.

Lahoz, W. A., *et al.*, 1996: Vortex dynamics and the evolution of water vapour in the stratosphere of the southern hemisphere. *Q. J. Roy. Meteorol. Soc.*, **122**, 423–450.

Lighthill, M. J., 1952: On sound generated aerodynamically, *Proc. Roy. Soc. Lond.*, **A 211**, 564–587.

Manney, G. L., *et al.*, 1994: Simulations of the February 1979 stratospheric sudden warming: model comparisons and three-dimensional evolution. *Mon. Wea. Rev.*, **122**, 1115–1140.

McIntyre, M. E., 1982: How well do we understand the dynamics of stratospheric warmings? *J. Meteorol. Soc. Japan*, **60**, 37–65 [Special middle-atmosphere issue, ed. K. Ninomiya. *NB:* "latitude" should be "altitude" on page 39a line 2.]

McIntyre, M. E., 2000: On global-scale atmospheric circulations. In: *Perspectives in Fluid Dynamics: A Collective Introduction to Current Research*, ed. G. K. Batchelor, H. K. Moffatt, M. G. Worster, 557–624. Cambridge, University Press, 631 pp. *Corrigenda:* (1) The symbols $\times$ for 3-dimensional vector products were all misprinted as $\wedge$; (2) the evaporation feedback at the tropical sea surface (p. 621) may be strongly compensated radiatively; see Held and Soden 2000, Ann. Rev. Energy Environment, **25**, 441–475.

McIntyre, M. E., Norton, W. A., 2000: Potential-vorticity inversion on a hemisphere. *J. Atmos.*

Sci., **57**, 1214–1235. *Corrigendum* **58**, 949 and on www.atm.damtp.cam.ac.uk/people/mem/

McIntyre, M. E., Palmer, T. N., 1983: Breaking planetary waves in the stratosphere. *Nature*, **305**, 593–600.

McIntyre, M. E., Palmer, T. N., 1984: The "surf zone" in the stratosphere. *J. Atm. Terr. Phys.*, **46**, 825–849.

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., Roulstone, I., 1996: Hamiltonian balanced models: constraints, slow manifolds, and velocity splitting. *Forecasting Research Scientific Paper* **41**, *Meteorological Office, U.K.*. Shortened and corrected version in revision for *J. Fluid Mech.*; full text and corrections available at http://www.atm.damtp.cam.ac.uk/people/mem/

McIntyre, M. E., Roulstone, I., 2001: Are there higher-accuracy analogues of semigeostrophic theory? In: *Large-scale Atmosphere–Ocean Dynamics: II: Geometric Methods and Models* (Proc. Newton Inst. Programme on Mathematics of Atmosphere and Ocean Dynamics), ed. I. Roulstone and J. Norbury. Cambridge, University Press, in press.

Mohebalhojeh, A. R., Dritschel, D. G., 2001: Hierarchies of balance conditions for the f-plane shallow water equations. **J. Atmos. Sci.**, in press.

Mote, P. W., *et al.*, 1996: An atmospheric tape recorder: the imprint of tropical tropopause temperatures on stratospheric water vapor. *J. Geophys. Res.*, **101**, 3989–4006.

Mote, P. W., *et al.*, 1998: Vertical velocity, vertical diffusion, and dilution by midlatitude air in the tropical lower stratosphere. *J. Geophys. Res.*, **103**, 8651–8666.

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., *et al.*, 2000: Brewer–Dobson Workshop (report plus transcripts of invited lectures by J. R. Holton and A. W. Brewer). *SPARC Newsletter* **15**, 25–32. Available from sparc.office@aerov.jussieu.fr, also http://www.aero.jussieu.fr/~sparc/ [sic].

Plumb, R. A., McEwan, A. D., 1978: The instability of a forced standing wave in a viscous stratified fluid: a laboratory analogue of the quasi-biennial oscillation. *J. Atmos. Sci.*, **35**, 1827–1839.

Plumb, R. A., *et al.*, 1994: Intrusions into the lower stratospheric Arctic vortex during the winter of 1991–92. *J. Geophys. Res.*, **99**, 1089–1105.

Rhines, P. B., 1975: Waves and turbulence on a beta-plane. *J. Fluid Mech.*, **69**, 417–443.

Riese, M., *et al.*, 2001: Stratospheric transport by planetary wave mixing as observed during CRISTA-2. *J. Geophys. Res.*, submitted. *See also* Riese, M., 2000: *Remote Sensing and Modeling of the Earth's Middle Atmosphere: Results of the CRISTA Experiment*, Habilitation thesis, Bergische Universität, Fachbereich 8, 42097 Wuppertal 1, Germany.

Robinson, W. A., 1988: Analysis of LIMS data by potential vorticity inversion. *J. Atmos. Sci.*, **45**, 2319–2342.

Rossby, C. G., 1936: Dynamics of steady ocean currents in the light of experimental fluid mechanics. Mass. Inst. of Technology and Woods Hole Oc. Instn. *Papers in Physical Oceanography and Meteorology*, **5(1)**, 1–43. [Material PV conservation: Equation (75).]

Salmon, R., 1983: Practical use of Hamilton's principle. *J. Fluid Mech.*, **132**, 431–444.

Salmon, R., 1985: New equations for nearly geostrophic flow. *J. Fluid Mech.*, **153**, 461–477.

Salmon, R., 1988: Semigeostrophic theory as a Dirac-bracket projection. *J. Fluid Mech.*, **196**, 345–358.

Saujani, S., and Shepherd, T.G., 2001: Comments on "Balance and the Slow Quasimanifold: Some Explicit Results". *J. Atmos. Sci.* **58**, submitted; FORD, R., *et al.*, 2001: Reply, *J. Atmos. Sci.* **58**, submitted.

Stewart, R. W., Thomson, R. E., 1977: Re-examination of vorticity transfer theory. *Proc. Roy. Soc. Lond.*, **A354**, 1–8.

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

Waugh, D. W., *et al.*, 1994: Transport of material out of the stratospheric Arctic vortex by Rossby wave breaking. *J. Geophys. Res.*, **99**, 1071–1088.

Whitaker, J. S., 1993: A comparison of primitive and balance equation simulations of baroclinic waves. *J. Atmos. Sci.*, **50**, 1519–1530.