Sectional Lecture, Proc. 18th Internat. Congr. Theor. Appl. Mech., Haifa, Israel,
ed. S. R. Bodner, J. Singer, A. Solan, Z. Hashin. Elsevier, 1993.

On the role of wave propagation and wave breaking in atmosphere–ocean dynamics

M. E. McIntyre

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, UK (*mem2@phx.cam.ac.uk*)

Abstract

Wave propagation, wave breaking, and the concomitant wave-induced momentum transport are ubiquitous processes in the Earth's atmosphere and oceans, a classic example being the surface gravity waves from storms in the Southern Ocean that drive mean longshore currents on California beaches. Such processes are not only interesting in themselves, but are also fundamental to making sense of the various 'wave-turbulence jigsaw puzzles' with which the atmosphere and ocean, separately and in combination, confront us. For instance, what used to be regarded as an enigmatic 'negative viscosity' of the subtropical atmosphere is now straightforwardly comprehensible in terms of Rossby wave propagation and breaking. Other examples include understanding (*a*) why the mean east–west winds in the equatorial lower stratosphere reverse every 14–18 months, throughout a belt encircling the globe, (*b*) why the lowest temperatures on earth (as low as 110 K, or −163°C) are found not in the winter hemisphere but near the summer pole, at altitudes between 80 and 90 km, (*c*) why the *e*-folding atmospheric lifetimes of certain man-made chlorofluorocarbons are of the order of a century, and (*d*) why the greatest concentrations of stratospheric ozone are found where photochemical ozone production rates are least.

This lecture will discuss some of the theoretical-mechanical concepts relevant to understanding these phenomena, including the concepts of wave 'momentum' and wave 'breaking'. There emerge, somewhat unexpectedly, what might prove to be some useful new ideas about the problem of water-wave generation by wind. The main point is that the water waves can be systematically amplified by certain irreversible, ratchet-like, non-superposable effects that depend on spatio-temporal inhomogeneities, such as wind gustiness and wave 'groupiness'. These include the effects of what might be called 'Rossby lee waves' in the airflow downstream of water-wave groups. The resulting wave drag can amplify non-breaking water waves and might, for instance, help to explain the growth of the 'energy front' reported by O. M. Phillips in this Proceedings.

1. INTRODUCTION

I want to widen the context of this Minisymposium and talk about some phenomena and concepts that appear fundamental to a whole range of problems in atmosphere–ocean dynamics. I also want to say something about that old but still problematic topic, the generation of water waves by wind. I cannot claim to be an expert on that problem, let alone on air-sea interaction in general, but it is possible that a fresh look from another perspective might help to advance our understanding.

The atmosphere and oceans used to be thought of simply as 'turbulent' fluids on a vast scale, and early attempts to understand them often involved Reynolds averaging together with the hope that the resulting eddy-flux terms might be able to be characterized, at least roughly, in terms of the notion of 'eddy diffusivity'. It was therefore a surprise when even the signs of these fluxes sometimes turned out contrary to expectation. This was found, for example, from global-scale atmospheric 'general circulation statistics' (e.g. Lorenz 1967); and for a time the phenomenon was thought of as a kind of mysterious 'negative eddy viscosity' (Starr 1968). This meant of course that no-one understood what was going on. Eastward momentum was seen to be transported poleward by large-scale eddies, against the local mean horizontal shear in subtropical latitudes. The eddy viscosity became infinite a little further poleward. Many other such examples are now known, one of the most conspicuous being the celebrated 'quasi-biennial oscillation' (QBO) of the equatorial lower stratosphere, to be described in §4.

The mystery was gradually solved as it was recognized how important for this purpose, as well as ubiquitous, are the various wave propagation mechanisms that operate in the atmosphere and oceans, such as internal gravity wave and Rossby wave* propagation. Among the important pioneering contributions were those of Charney and Drazin (1961), Eady (unpublished, but see Green 1970), Booker and Bretherton (1967), Wallace and Holton (1968), Lindzen and Holton (1968), Dickinson (1969), and Rhines (1975). Today we have a relatively clear view of these problems, both through data from clever terrestrial and space-based observing techniques, and through a better understanding of the basic theoretical principles and of how to apply both numerical and idealized theoretical-mechanical modelling. Such modelling is used not only in hypothesis-testing thought-experiments, but also as pointing toward better ways to make observational data tell us, in a dynamically intelligible way, more about what is going on in the real atmosphere and oceans (e.g. Thorncroft *et al.* 1993, & refs.). Better understanding includes seeing what is robust about an idealized model, hence which aspects of it are likely to carry over to more realistic situations. This lecture will mention a few such models and their contribution to our present-day understanding. Also touched on will be the intimate relation between wave propagation mechanisms and shear instability mechanisms, and the concept of 'wave breaking' and its frequent relevance — when appropriately defined — to phenomena involving wave-induced momentum transport.

It is interesting to view the wind–wave problem from the conceptual vantage point thus arrived at. There emerge what may turn out to be some new ideas about wave generation mechanisms, to be discussed briefly at the end of the lecture. Besides being of general significance for our understanding of ocean waves and air–sea interaction, these ideas might help to explain, for instance, the growth of the 'short wave energy front' seen in the experiments described in Professor Phillips' Minisymposium Lecture in this Proceedings and in Chu *et al.* (1992). The key is to recognize all the wave propagation and wave breaking mechanisms that come into play, in the air as well as in the water.

2. THE MIDDLE ATMOSPHERE: SOME OBSERVED FACTS

The phenomena to be discussed, including the phenomenon of 'negative viscosity', are seen very clearly in what is now usually called the 'middle' atmosphere (but still, occasionally, the 'upper' atmosphere). It comprises the stratosphere, extending to about 50 km altitude, and the mesosphere above it, extending to somewhere between 80 and

*Rossby waves (historically, 'Kelvin–Kirchhoff–Rayleigh–Rossby waves') may also be called 'vorticity waves' or, more generally, 'potential-vorticity waves' (§6 below). I am following today's established usage.

100 km. Figure 1 includes the bulk of the middle atmosphere; it shows a typical mean January temperature distribution as a function of altitude and latitude. The units on the right-hand scale are e-folding pressure scale heights, roughly 7 km, so that the altitude range is from sea level to roughly 85 km. (Somewhere above this altitude, temperatures rise steeply into what is called the thermosphere, where simple thermodynamics and continuum mechanics break down and plasma physics becomes important.) This section and §§4,8 sketch briefly some of what is known and understood about the middle atmosphere; a more extensive discussion and bibliography can be found in a recent review of mine (1992) written for a non-specialist audience of physicists, and in other, more specialized review material cited therein; see for instance the book by Andrews, Holton and Leovy (1987).

Figure 1. Temperatures T between sea level and about 85 km, for typical January conditions (degrees Kelvin, longitude and time averaged). Northward is toward the right, and the right-hand altitude scale is in e-folding pressure scale heights, roughly 7 km. From satellite and other observations analyzed by Barnett and Corney (1985).

The light shading in Figure 1 shows the warmest regions, and the dark shading the coldest, in January. Some of these features can be understood largely from considerations of radiative heat transport. For instance the high temperatures $T \gtrsim 260\,\mathrm{K}$ near 7 scale heights or 50 km, defining the 'stratopause', are largely due to heating by absorption of solar ultraviolet radiation by ozone, balanced by infrared cooling to space. T increases southward at the stratopause mainly because the diurnally averaged normal solar irradiance has an absolute maximum at the south pole at the December solstice, and throughout a substantial part of December and January. The south pole is then, so to speak, the sunniest place on earth.*

A number of other features are not so simply explained. One example is the very cold region at the top left of the picture, where the lowest mean temperature plotted is $T = 150\,\mathrm{K}$. In fact individual rocket soundings have shown temperatures as low as 110 K

*For solar declination α relative to a spherical, rotating earth, the fractional length of day $\lambda(\phi)$ at latitude ϕ is $\pi^{-1}\arccos[\max\{-1, \min(1, -\tan\alpha\tan\phi)\}]$, and the diurnally averaged vertical component of solar irradiance is the full solar irradiance multiplied by $\lambda(\phi)\sin\alpha\sin\phi + \pi^{-1}\sin\{\pi\lambda(\phi)\}\cos\alpha\cos\phi$. This has an absolute maximum at the south pole when α is within 2.8° of its minimum value −23.6°.

on occasion. At these altitudes, the sunniest place on earth is also the coldest place on earth. Until about a decade ago, this was regarded as one of the great enigmas of atmospheric science, the enigma of the 'cold summer mesopause'. There is no doubt that the observed temperatures are far below what they should be on the basis of radiative heat transport and photochemistry alone. Temperatures substantially higher than their radiative values also occur. They occur, for instance, throughout most of the depth of the north-polar region on the right of Figure 1.

Let us note one other observed fact, at first sight having little to do with the observed pattern of temperature anomalies, but actually closely connected. This is the now-notorious fact (with its potentially serious implications for the stratospheric ozone layer) that the man-made chlorofluorocarbons known as CFC-11 and CFC-12 have very long atmospheric lifetimes, of the order of a century. These are *e*-folding times. Even if all leakage of these CFCs into the atmosphere could be stopped tomorrow, it would take several centuries for their concentrations to decrease to, say, 1% of present values.

How are these facts connected? The connecting link is indicated in Figure 2, which shows an estimate of the mean circulation of the stratosphere and mesosphere. It is the mean in a sense very roughly equivalent to a Lagrangian mean with suitable re-initialization of particle ensembles (see the caveats in my 1992 review and its bibliography). This mean circulation gives us a roughly correct explanation both of the temperature anomalies and of the observed lifetimes of CFCs, and the lifetimes of certain other long-lived chemical tracers. The temperature anomalies are accounted for by adiabatic expansion in the rising branches of the circulation, and adiabatic compression in the descending branches, pulling temperatures away from the radiative values toward which they would otherwise tend to relax. The rising branches carry chemical tracers upward, for example CFCs and other tropospheric tracers through the tropical stratosphere. The rate at which they are carried upward governs the rate at which they are exposed to sufficiently energetic solar ultraviolet radiation, and hence destroyed photochemically. This is the main way in which CFCs are destroyed, removal rates at sea level being relatively small. The typical strength of the mean circulation required to account for the observed temperature anomalies also gives a CFC destruction rate consistent with the observed lifetimes.

Again, the rising branch of the circulation at higher altitudes, shown schematically by the heavy dashed curve at the top left of Figure 2, explains the extraordinarily low summer mesopause temperatures. The circulation also carries small amounts of water vapour upward to the mesopause. A phenomenon observed at these altitudes, the sporadic formation of 'noctilucent' and 'polar mesospheric' clouds, requires a supply of water vapour as well as exceptionally low temperatures.* Water vapour is photochemically destroyed near the mesopause, where ultraviolet photons are even more energetic than in the stratosphere. Considerations like these, and other observational evidence, give confidence in the picture suggested by Figure 2.

The mean circulation is also part of what controls the distribution of ozone, and the rate of replenishment of the ozone layer. The photochemistry of ozone is complicated and can interact more subtly with the fluid dynamics than, for instance, CFC destruction, which depends mainly on total exposure to ultraviolet photons. Ozone can be expected to be more sensitive to processes hidden by the averaging in Figure 2; we return briefly to this point in §8.

---

*See for instance Thomas *et al.* (1989, & refs.). These incidentally are the world's highest clouds, by far, occurring at altitudes around 83–85 km. Anyone who lives between about 50° and 60° latitude can observe them as an electric blue glow above the poleward horizon, sometimes intricately patterned, on a few clear nights after midnight in the two months following midsummer.

Figure 2. *Light curves*: mass transport streamlines of the longitude and time averaged global-scale circulation in an altitude range between about 3 and 9 pressure scale heights (cf. right-hand scale in Figure 1), estimated from satellite data for January 1979 by Solomon *et al.* (1986). The pressure altitude is in nominal kilometres, defined as 1/7 of a scale height. The circulation is defined in a quasi-Lagrangian sense giving a simplified, but roughly correct, indication of the vertical advective transport of chemical tracers (and is broadly consistent with direct observations of such tracers). The time for a notional fluid element to rise from the tropical tropopause to, say, 40 km is typically of the order of two years. *Heavy dashed curve*: qualitative indication of the extension of the circulation into the upper mesosphere and lower thermosphere.

How is the mean circulation driven? This again forces us to consider processes hidden by the averaging. The key point is that the Earth is a rapidly rotating planet. It is rapidly rotating in the sense that the distribution of angular momentum M per unit mass in the atmosphere is dominated by the Earth's rotation, and only weakly affected by air motion relative to the Earth. In the extratropics there is a strong latitudinal gradient of M. So a circulation like that shown in Figure 2 cannot persist unless something exerts a persistent torque on the extratropical atmosphere, in an appropriate sense (in fact, against the Earth's rotation everywhere except in the summer mesosphere, where it must be with the rotation). In a thought-experiment in which this torque is switched off, the circulation begins to die down, and temperatures to relax toward their radiative values (e.g. Haynes *et al.* 1991, & refs.).

It is now believed on good evidence that in the real middle atmosphere this torque arises from wave-induced angular momentum transport, in fact mostly from Rossby waves and internal gravity waves generated in the troposphere. The Rossby waves account for most of the lower part of circulation shown in Figure 2, and the gravity waves account for the upper mesospheric branch. The sense of the angular momentum transport is related to the phase speeds of the waves and to the different places where they are generated and dissipated, and not locally to such things as the sign of the mean shear. It is no wonder, then, that attempts to apply ideas like 'eddy viscosity' can produce incongruous results, such as negative or infinite eddy viscosities.

3. WAVE-INDUCED MOMENTUM TRANSPORT

The fact that wave propagation and diffraction are generally accompanied by a systematic flux or transport of momentum is a well-known rule in theoretical physics, applying to waves in fluids as well as to the more obvious case of waves in a vacuum, such as photons. The key phenomenon is that, if progressive waves are generated in one place A and dissipated in another B, this is accompanied by an irreversible, cumulative transport of momentum from A to B. The sense and rate of this transport is usually given, to some useful approximation, by what might be called the photon analogy, or 'quasimomentum rule' (§5).

The physical reality of such wave-induced momentum transport is easy to demonstrate in the laboratory. For instance, it manifests itself in the phenomenon known as the 'sonic wind' or 'quartz wind' (e.g. Lighthill 1978a,b), in which a piezoelectric transducer emits a beam of ultrasound that transports momentum from the transducer to locations where the waves dissipate. The resulting mean force on the fluid often generates a turbulent jet. One can demonstrate what is fundamentally the same thing without any special apparatus at all, using capillary-gravity waves at frequencies of order 5 Hz. This was done during the lecture using a small cylindrical wavemaker (Figure 3a) and a glass dish containing water with chalk dust on the surface. Making the wavemaker oscillate vertically shows that the observed mean flow (arrowed curves) is predominantly wave-induced, and not Rayleigh–Schlichting boundary-layer streaming from the surface of the wavemaker. The latter has the opposite sense (e.g. Lighthill 1978a, p.348; Van Dyke 1982, Fig. 31). Carefully stopping the wavemaker and observing the persistence of the mean flow shows that it is more than a 'Stokes drift': irreversible, cumulative momentum transport has indeed taken place. With a larger area of water, one can use a curved wavemaker to focus the waves on a more distant spot (Figure 3b), illustrating the fact that the wave generation and dissipation sites can be well separated spatially.

Figure 3. Simple demonstrations of momentum transport by 5 Hz surface capillary-gravity waves (see text). Configuration (a) works well with a wavemaker whose diameter $\sim$ 3 to 4 cm, and configuration (b) with a wavemaker whose radius of curvature $\gtrsim$ 50 cm. From McIntyre and Norton (1990).

This last point is underlined by the classic work of Munk *et al.* (1963) in which, following earlier work by Barber and Ursell, they demonstrated that surface gravity waves generated by storms in the Southern Ocean can, and often do, propagate across the Pacific all the way to beaches in California, where they break and generate longshore mean currents (see also Snodgrass *et al.* 1966). This is a clear case of irreversible wave-induced momentum transport over many thousands of kilometres! The internal gravity waves that sustain the mesospheric circulation of Figure 2 also dissipate mostly by breaking. They have sources mostly in the denser layers of the atmosphere far below, so that horizontal momentum is transported over vertical distances not of thousands, but certainly many tens, of kilometres (e.g. Fritts 1984, 1993).

Essentially the same things happen with Rossby waves, except that Rossby-wave dynamics has a peculiar 'one-signedness' that can constrain the sense of the momentum transport in a ratchet-like way. As well as being interesting for this reason, the case of Rossby waves is arguably the simplest to understand from first principles. This will be seen in §7. First, however, I want to mention an example involving internal gravity and other kinds of waves, in which the wave-induced momentum transport interacts with wave refraction to produce an interesting feedback oscillation of the *mean flow* that has been observed both in the laboratory and in the real atmosphere — the celebrated 'quasi-biennial oscillation' or 'QBO' — and then I want to touch briefly on the photon analogy and on what is sometimes called wave 'momentum'.

4. THE QBO AND ITS LABORATORY ANALOGUE

The 'laboratory QBO' was demonstrated in a famous experiment by Plumb and McEwan (1978). The system used was a salt-stratified fluid contained in a large laboratory annulus, of depth 50 cm and gap width 12 cm. Internal gravity waves are excited by making a flexible lower boundary oscillate in a standing wave, equivalent to equal amplitudes of clockwise and anticlockwise progressive waves. The response of the fluid breaks this symmetry (the annulus is at rest relative to the laboratory, and the Earth's rotation can be neglected); and a wave-induced mean flow arises, horizontally around the annulus. Soon a regime is established displaying a characteristic spacetime signature, in which the mean flow reverses periodically at a given altitude, and does so earlier at higher altitudes. This was illustrated by a movie of the original experiment shown in the lecture.

The two mechanisms involved are first the wave-induced angular momentum transport, cumulatively changing the mean velocity profile as the waves dissipate (viscously in this case), and second the effect of mean shear in Doppler shifting and refracting the waves (somewhat like the selective surface-wave refraction that can make a bathtub vortex appear, at first glance, to be rotating the wrong way). In the Plumb–McEwan experiment the wave dissipation rate is least, and the vertical group velocity greatest, when the waves propagate against the mean flow. Such waves therefore penetrate highest. They transport angular momentum in the same sense as their intrinsic angular phase speed and can therefore reverse the mean motion where they dissipate, first at higher and then at lower and lower altitudes in the annulus. Mathematical models that express these ideas tend to behave non-chaotically, and to produce the spacetime signature very robustly (e.g. Yoden and Holton 1988; Haynes *et al.* 1993), as long as waves of sufficient amplitude are excited in both senses, clockwise and anticlockwise.

The same spacetime signature (albeit not exactly periodic) is conspicuously present in the mean east–west winds of the tropical lower stratosphere, throughout a belt encircling the globe, at altitudes between about 16 and 35 km and latitudes between about $\pm 15°$. Whereas in the laboratory it typically takes about half an hour for the mean flow to reverse, in the real stratosphere it takes about 14 ± 3 months. The wind reversals are clearly seen in the tropical radiosonde data that became routinely available from the early 1950s onwards; and there is indirect evidence for their existence many decades before then (Hamilton and Garcia 1984; Teitelbaum *et al.* 1993). Although our detailed understanding is incomplete, it seems overwhelmingly likely that the resemblance to the laboratory experiment and to mathematical models of it is not accidental, and that the mean flow changes are wave-driven in much the same way. Indeed, before wave driving was thought of at all in this context — the observed QBO used to be another great enigma of atmospheric science — a strong case had already been made for the existence of some strange kind of eddy or fluctuation-induced angular momentum transport, again entailing 'negative eddy viscosity'. Without some such eddy transport, it seemed impossible to make dynamical sense of the observed mean flow changes, on the basis of careful physical arguments and numerical experiments (Wallace and Holton 1968). It was only after this that the realization came (Lindzen and Holton 1968) that wave-induced angular momentum transport could behave in just the required manner.

We do not yet have a convincing *quantitative* model of the real QBO including, for instance, the effect of the mean upwelling illustrated in Figure 2. This advects angular momentum upward on a comparable timescale, and we lack sufficient quantitative knowledge of its strength and seasonal variation. There is also uncertainty over exactly which wave types are significant in the tropical stratosphere. It used to be assumed that the principal such wave types are the equatorially trapped Kelvin and Rossby–gravity waves (e.g. Andrews *et al.* 1987, Gill 1982). There is observational support for this in the case of the Kelvin wave, whose observed amplitude is not far from having an appropriate order of magnitude. But it has seemed more and more likely, for instance, that planetary-scale Rossby waves originating in the extratropical troposphere are more significant than equatorial Rossby–gravity waves — meaning locally significant in the tropics (e.g. Dickinson 1968; Lindzen and Tsay 1975; Andrews and McIntyre 1976; Dunkerton 1983; Takahashi and Boville 1992; O'Sullivan and Hitchman 1992, see §8 below) as well as indirectly significant through the strength of the mean upwelling.

We have even less knowledge, either observational or theoretical, of possible wave generation mechanisms. The real QBO seems to involve highly complicated, chaotic, nonlinear wave generation processes (mainly in the troposphere, both tropical and extratropical) to which there is a robustly non-chaotic response in the tropical stratosphere — with, it might be added, potentially important consequences for long-range weather forecasting, such as a feedback on the depth of cumulonimbus convection and hence on tropical cyclone intensities and El Niño timings (Gray *et al.* 1992). These aspects remain a challenge and opportunity for the future.

5. LONG-DISTANCE TRANSPORT AND THE PHOTON ANALOGY

The 'photon analogy' or 'quasimomentum rule' says that the rate at which momentum is transported from location A to location B, when a wave packet is generated at A and dissipated at B, is the same as if

(*a*) the fluid were absent, and
(*b*) the wave packet had a certain amount of momentum **q** that it carries around with it, like a photon in a vacuum.

The quantity **q** is not actually a momentum. It is a property of the wave packet that may more aptly be called its *quasimomentum* or *pseudomomentum*, in order to distinguish it from momentum. Momentum, in the presence of a material medium, is a fundamentally different quantity.* Hence the alternative term 'quasimomentum rule'. Indeed a better, more precise statement is that "the rate... is the same as if

(*a*) the fluid were absent, and
(*b*) the wave packet had momentum equal to its quasimomentum."

The quasimomentum is a wave property in the sense that it can be evaluated from linearized wave theory alone. In the simplest theoretical approximations (slow modulation

*This is because its conservation corresponds to a different translational symmetry (of the propagating medium rather than of the total physics). There is no widely agreed name for **q**. Other names used with different kinds of waves include 'wave-vector', 'Poynting's momentum', 'Minkowski momentum', 'radiation momentum', 'acoustic momentum', 'crystal momentum', 'phonon momentum', 'tensor momentum', 'field momentum', 'canonical momentum', 'wave momentum', and 'momentum'. This has led to some confusion. For the basics plus some history going back to the time of Rayleigh, Poynting, Abraham, and Minkowski, and for keys to the literature, the reader may consult my 1981 essay and its bibliography, also Loudon and Paige (1991, p.236), Peierls (1991, §§2.4–6), and Shepherd (1990, §5).

as well as small amplitude), we may take $\mathbf{q} \simeq \hat{E}\mathbf{k}/\hat{\omega}$ where $\hat{E}$ is the intrinsic wave-energy in the sense discussed e.g. by Bretherton and Garrett (1968), $\mathbf{k}$ is the wavenumber vector, and $\hat{\omega}$ is the intrinsic frequency, or frequency Doppler-shifted to a reference frame moving with the local mean flow. Like $\hat{E}$, $\mathbf{q}$ is $O(a^2)$ in wave amplitude a.

The analogy summarizes a body of special and general results from theories in which a is considered small and in which wave-induced momentum transport, and all the associated mean effects, are self-consistently described correct to $O(a^2)$. As is well known the theoretical calculations can be elaborate and tricky, equatorial Rossby–gravity waves being a case in point (Andrews and McIntyre 1978a §9, & refs.). Such theories are often qualitatively applicable, and may also be quantitatively applicable to important parts of the problem, such as the part concerned with wave-induced momentum transport between the sites of wave generation and dissipation. Amplitudes in the intervening wave field can in some cases be truly small (in the relevant sense, measured by wave slopes), as for instance in the case of surface gravity waves crossing the Pacific. Some aspects of the theory and the photon analogy extend to finite amplitude as well (Andrews and McIntyre 1978a,b). The analogy has relevance to all the situations and all the wave types mentioned in this lecture, including the equatorial Kelvin and Rossby–gravity waves that are thought to contribute to the real QBO.

But wait, I hear someone say, why all this hair-splitting about an 'analogy' — why the 'as if' and the 'quasi', to say nothing of the 'pseudo'? Isn't this complicating things unnecessarily? Surely wave packets in a fluid really do have momentum $\mathbf{q}$, which they really do carry around with them, just like photons in a vacuum. Is it not well known that dissipating waves exchange 'their momentum' (meaning $\mathbf{q}$) with the mean flow? How could they do such a thing if they didn't really have momentum $\mathbf{q}$ to exchange? Besides, how else could those waves that propagate across the Pacific, after generation by storms in the far south, drive longshore currents on northern beaches? In that situation we do not have a steady wavemaker, and we do not have a steady 'radiation stress' spanning the whole ocean for weeks on end. So surely wave packets, and finite wavetrains, must just carry momentum with them. And what about the mesospheric circulation and the noctilucent clouds? Real internal gravity waves are highly intermittent; and isolated wave packets are again, arguably, a more relevant idealization than steady waves spanning the whole depth of the middle atmosphere.

Well, what really happens is interesting, and worth a brief digression. Take for instance an idealized version of the situation in mid-Pacific. Figure 4, from my 1981 essay, shows an isolated, non-dissipating, two-dimensional packet of surface gravity waves on deep water. More precisely, it shows semi-schematically the leading-order theoretical solution describing the disturbance and its accompanying $O(a^2)$ velocity field, derivable from the work of Longuet-Higgins and Stewart (1962). That solution can be used to compute the total momentum of the propagating disturbance. In order for the photon idea to be literally true, in the manner just envisaged, it would be necessary for that momentum to be well defined and equal to $\mathbf{q}$. It turns out that in this case the momentum is indeed well defined. But a careful computation shows what Longuet-Higgins and Stewart also found, namely that the momentum is *not* equal to $\mathbf{q}$ (see figure caption). To leading order, it is zero!

This is not a paradox. Rather, it is one of many counterexamples showing that, for waves in material media, the photon idea cannot be taken literally as a general principle. In order for the photon idea to make sense in general — in fact to make sense outside a very limited set of circumstances — it is indeed crucial to regard it as an analogy, i.e., to retain the words 'as if' preceding items (a) and (b) above, and to continue to recognize the distinction between momentum and quasimomentum. Thus stated, the analogy is both useful, and capable of general theoretical justification. One approach is to use ‘generalized Lagrangian means’ in conjunction with Kelvin’s circulation theorem (Andrews and McIntyre 1978a); the importance of the circulation theorem in this kind of problem was recognized by Rayleigh (1896), and its connection with the photon analogy was, I think, first recognized by Bretherton (1971). Additional considerations that have improved our general understanding, but have yet to be fully worked out, can be found in my paper with Norton (1990). But, to return to the idealized ‘Pacific’ problem, how then *does* the whole thing work fluid-mechanically?

The main point is this. As a wave packet propagates past any given fluid element between its generation and dissipation sites, it gives rise to an $O(a^2)$ mean forcing *whose time integral is zero* for that fluid element. Details are complicated but the most significant aspect of this forcing can be thought of, for present purposes, as coming from the divergence of a radiation stress spanning the region occupied by the wave packet (and satisfying Newton’s third law of ‘action and reaction’). This causes the given fluid element first to feel a mean push, and then a mean pull, against other fluid elements in the region. Consequently, the time-integrated force on the given fluid element is zero. Corresponding statements are true of the other aspects of the $O(a^2)$ mean forcing, such as apparent mass sources and sinks (e.g. Andrews and McIntyre 1978a).

Computing the $O(a^2)$ response to a given $O(a^2)$ mean forcing is a linear problem. Therefore, during a time interval in which one or more wave packets propagate from A to B, one can regard the $O(a^2)$ mean forcing as the sum of two contributions: first a steady forcing, corresponding to a steady wavetrain and its radiation stress spanning the entire region between A and B — and conforming to the photon analogy — and second an oscillatory forcing in the same region whose time integral vanishes everywhere.

It is only the first of these two contributions that is interesting from the present viewpoint, i.e. that corresponds to the notion of a cumulative, irreversible wave-induced transport of momentum from A to B. The second, oscillatory contribution evokes a response that is non-cumulative, because of the vanishing of its time integral. It tends moreover to be strongly dependent on circumstances such as conditions at remote boundaries, and how the waves were generated. Its details can be complicated. For instance if one were to include stable stratification and Coriolis effects in a less idealized model of the Pacific ocean, then the $O(a^2)$ response to the passage of a non-dissipating wave packet would be quite different from that shown in Figure 4. It would involve the excitation of very weak $O(a^2)$ internal Coriolis–gravity (inertio–gravity) waves over a large area of ocean. In fact something similar happens even in the special case of Figure 4 (see my 1981 essay for further discussion) since in general there are very distant, very weak, fast-propagating, ultra-long $O(a^2)$ surface gravity waves, which I have not attempted to depict in the figure but which embody significant amounts of momentum, and which depend on how the wave packet was generated.

In summary, then, what is complicated, and circumstance-dependent, is the detailed, unsteady $O(a^2)$ mean response of the fluid medium to the generation, propagation and dissipation of a wave packet, or of many wave packets. What is simple, and general, is the fact that the time integral of this $O(a^2)$ response is zero apart from the cumulative contribution given by the photon analogy. This is true, in a wide range of problems of this kind, whether or not any well-determined, well-localized $O(a^2)$ mean momentum appears temporarily within the fluid as wave packets propagate past, or whether for instance $O(a^2)$ mean momentum is temporarily taken up by distant boundaries, as is sometimes signalled by divergent momentum integrals in idealized versions of the problem — or whether $O(a^2)$ mean momentum is taken up by distant long-wave disturbances, as in the problem of Figure 4 and its variants.

Another thought-experiment of fundamental interest in this connection is to scatter the wave packet of Figure 4 from an immersed obstacle, and ask what the mean recoil force is. One finds that it is given by the photon analogy, appropriately re-stated. This again is 'despite' the fact that, in this case, the wave packet has a well defined momentum equal to zero. What happens is that more very weak, fast-propagating, ultra-long-wave $O(a^2)$ disturbances, of the kind already referred to, are radiated during the reflection process. One can derive a general result, comprising a non-trivial extension of Noether's theorem, that shows why the net effect of all these $O(a^2)$ phenomena must be given by the photon analogy in many such cases including this one. The result also shows why there are some 'exceptional' problems for which the analogy fails.*

Figure 4. Packet of surface gravity waves propagating toward the right in deep water, and its accompanying $O(a^2)$ velocity field plotted quantitatively except that the Stokes drift (near the surface) is not depicted. The total momentum of the wave packet is well defined, and comprises the momentum of the Stokes drift, which, for these particular waves, equals $\mathbf{q}$, plus the momentum of the return flow underneath, which equals $-\mathbf{q}$ (because return-flow acceleration reactions feel the free surface as effectively rigid — for further discussion see McIntyre 1981).

6. ROSSBY WAVES, VORTICES AND SHEAR INSTABILITIES

Of all the examples of irreversible wave-induced momentum transport, some of those associated with Rossby waves are arguably the simplest as well as among the most important. Rossby waves and related phenomena are ubiquitous in the atmosphere and oceans, and are fundamental to almost every aspect of large-scale atmosphere–ocean dynamics. For example, it is Rossby waves and related phenomena that drive most of the mean circulation illustrated in the lower part of Figure 2. The 'Rossby-wave elasticity' to which the wave propagation owes its existence is important also, for instance, in strongly inhibiting the turbulent transport of chemicals into the Antarctic ozone hole, or of drier air into the moist eye wall of a tropical cyclone. A still wider interpretation of 'related phenomena', meaning phenomena depending on 'Rossby elasticity', would

*The key step is to consider a certain translational symmetry operation giving a conservation law for the sum of the quasimomentum of the waves and the momentum of the immersed obstacle. There are 'exceptional' cases because this translational symmetry operation is singular. The singularity is strong enough, in some cases, to invalidate the extension of Noether's theorem. Details will be given in a forthcoming paper with S. D. Mobbs (1993). I discovered one of these exceptions by chance some time ago, via some very careful $O(a^2)$ calculations (McIntyre 1972, 1973). The existence of such exceptions is a telling confirmation that the photon analogy is only an analogy, highly useful when valid but not, in general, to be taken literally.

include many important types of *shear instabilities* all the way from ordinary small-scale shear instabilities to the large-scale, buoyancy-powered 'baroclinic instabilities' that can lead to the formation of common types of atmospheric and oceanic eddies and vortices, including extratropical weather cyclones (Hoskins *et al.* 1985, & refs.).

Rossby waves and related phenomena occur in dynamical systems of the generic form

$$DQ/Dt = 0 \;, \qquad \mathbf{u} = \mathbf{I}(Q) \;, \tag{6.1a,b}$$

where $D/Dt = \partial_t + \mathbf{u}\cdot\nabla = \partial_t + u\partial_x + v\partial_y$, the two-dimensional material derivative, and where $\mathbf{I}(\cdot)$ is a time-independent functional of the materially conserved scalar field Q. The simplest case is the familiar case of two-dimensional inviscid, incompressible vortex dynamics, for which

$$\mathbf{I}(Q) = (-\partial_y,\ \partial_x)\nabla^{-2}Q_{\mathrm{r}} \tag{6.2}$$

where $Q_{\mathrm{r}} = Q = v_x - u_y$, the ordinary vorticity and, to make the inverse Laplacian unambiguous, suitable boundary conditions are understood such as evanescence of $|\mathbf{u}|$ at infinity. Note that (6.2) implies $\nabla\cdot\mathbf{u} = u_x + v_y = 0$. The next simplest case is the same thing on a rotating earth whose vertical component of absolute vorticity is Q_{e}, say, a prescribed function of horizontal position but not of time. Then, retaining the notation $\mathbf{u} = (u, v)$ for the relative velocity and $Q_{\mathrm{r}} = v_x - u_y$ for the relative vorticity, we may take $Q = Q_{\mathrm{e}} + Q_{\mathrm{r}}$ in (6.1) and retain (6.2) unchanged, remembering that Q_{e} is prescribed. The notation in (6.1) has been chosen to emphasize the fact that the single scalar field Q contains all the dynamical information. At every instant, the Q field can be 'inverted' to recover the velocity field $\mathbf{u}$. One may call $\mathbf{I}(\cdot)$ the 'inversion functional'.

In more realistic models of the rotating, stratification-constrained, vortical flows that occur in the real atmosphere and oceans, the same generic mathematical structure (6.1) applies, in many cases to remarkable accuracy. This is why simple two-dimensional vortex dynamics has always been such an important idealization in the context of atmosphere–ocean dynamics. The coordinates x and y now measure horizontally-projected distances along the (approximately horizontal) stratification surfaces. D/Dt and $\mathbf{u} = (u, v, 0)$ are still two-dimensional on each such surface, and $u_x + v_y$ is still 'zero' to some useful approximation (more precisely, has typical magnitudes $\ll$ typical magnitudes of $v_x - u_y$). What is new is that the inversion functional $\mathbf{I}(\cdot)$ is now three-dimensional. Away from the equator, it still has the qualitative character of (6.2) but with ∇^{-2} more like a three-dimensional inverse Laplacian, in coordinates vertically stretched by Prandtl's ratio, the ratio of the buoyancy to Coriolis frequencies. Distortions of the stratification surfaces are also determined as part of the inversion operation. One may generally characterize such stratified, rotating flows as approximately 'layerwise two-dimensional'. In the most accurate models, which include models whose validity extends into the tropics, Q is the Rossby–Ertel potential vorticity, and inversion is no longer a linear operation (e.g. Hoskins *et al.* 1985, & refs; Thorpe 1985; Davis 1992; Raymond 1992).

One has here, incidentally, a framework for the general characterization of coherent structures such as vortices and vortex pairs, in atmosphere–ocean dynamics — amounting to variations on a theme from Professor Roshko's Opening Lecture in this Proceedings. For instance a 'vortex', in any dynamical system of the form (6.1), is the coherent structure represented by a strong, isolated anomaly in the Q field together with the induced velocity and any other relevant fields in and around it, where 'induced' means given by whatever inversion operator $\mathbf{I}(\cdot)$ characterizes that dynamical system. Atmospheric cyclones and anticyclones, and oceanic Gulf Stream rings and 'Meddies' (e.g. Armi *et al.* 1988), are all cases in point. Figure 5 illustrates one such structure in a model atmosphere, somewhat idealized but instantly recognizable to any meteorologist familiar with large-scale atmospheric behaviour. It is a cyclonic (earth-co-rotating) extratropical vortex induced by a strong, compact anomaly in the Rossby–Ertel potential vorticity field near the tropopause. Coherent structures such as these are often important for weather developments, in which fast advection of potential-vorticity anomalies near the tropopause is an important aspect of the dynamics, approximately satisfying eq. (6.1a) over timescales of several days.

Figure 5. Section across the axisymmetric structure induced by an isolated, axisymmetric, cyclonic potential-vorticity anomaly (stippled region) near a model tropopause (heavy curve) across which the Rossby–Ertel potential vorticity has a strong discontinuity, by a factor of 6. The family of thin curves some of which are closed are isopleths of tangential velocity, at 3 $\mathrm{ms^{-1}}$ intervals — the greatest velocities $> 21\ \mathrm{ms^{-1}}$ being at the tropopause — and the other family of thin curves, more nearly horizontal, are the stratification surfaces. These are isopleths of potential temperature θ defined to coincide with actual temperature at pressure 1000 mb or 1000 hPa, plotted at 5 K intervals. 'Anomaly' means a potential-vorticity contrast on a θ surface (Hoskins *et al.* 1985, eq. 29). The induced surface-pressure minimum is 41 hPa below ambient. The structure is typical for middle latitudes; the Coriolis parameter is $10^{-4}\ \mathrm{s^{-1}}$ (as at latitude 43.3°N). The domain shown has a radius of 2500km. From the work of Thorpe (1985).

Now to 'Rossby-wave elasticity'. The simplest example occurs in the strictly two-dimensional dynamical system specified by (6.2) with $Q_\mathrm{r} = Q - Q_\mathrm{e} = v_x - u_y$, and Q_e a linear function of y so that when $\mathbf{u} \equiv 0$ we have

$$\partial Q/\partial y = \partial Q_\mathrm{e}/\partial y = \beta = \text{constant} . \tag{6.3}$$

This is Rossby's famous 'β-plane' or 'nearly flat earth' model, with (x, y) taken to be Cartesian. Then (6.1) is satisfied — in this case without linearization, as it happens — by expressions of the form $\mathbf{u} = (-\partial_y,\ \partial_x)\psi(x, y, t)$ $(Q_\mathrm{r} = \nabla^2\psi)$ with $\psi(x, y, t) \propto \cos ly \cos\{k(x - ct)\}$, provided that the phase speed c and the wavenumber components k, l satisfy the dispersion relation

$$c = -\beta/(k^2 + l^2) . \tag{6.4}$$

The one-signedness previously referred to shows up here: (6.4) is a formula for c, and not c^2 as in classical small-vibration problems in non-rotating reference frames. We can see the reasons for the one-signedness and appreciate its robustness as follows.

By (6.1), the contours of constant Q are also material contours. If a disturbance makes these contours undulate as suggested in Figure 6, then Q_r in (6.2) will be alternately positive and negative as indicated by the plus and minus signs in Figure 6. Then since $\psi = \nabla^{-2}Q_\mathrm{r}$ the contours of ψ can be pictured as the equipotentials of the electrostatic field due to a pattern of alternating positive and negative charges (with the sign changed), or as the topographical contours giving the displacement of a stretched elastic membrane that is pulled up $(-)$ and pushed down $(+)$ alternately in the same pattern. Hence ψ will have hills and valleys centred respectively on the minus and the plus signs, implying that the strongest north-south velocities (at right angles to the electric field in the electrostatic analogy) will occur at intermediate positions, a quarter wavelength out of phase with the displacement, and in the sense shown by the heavy, dashed arrows in Figure 6. If one now makes a moving picture in one's mind's eye of what this induced velocity field will do to the material contours, one can see at once that the behaviour must be oscillatory, and also such that the undulations propagate from right to left only.

The same physical picture exhibiting one-way propagation applies when $Q_{\rm e} = 0$ but Q has a background gradient $\partial Q_0/\partial y$ due to a mean shear flow $\mathbf{u} = \{u_0(y), 0\}$,

$$\partial Q_0/\partial y = -\partial^2 u_0/\partial y^2 = -u_{0yy} \tag{6.5}$$

so that $Q = Q_0(y) + Q'(x,y,t)$ and $\psi = -\int^y u_0(\tilde{y})\, d\tilde{y} + \psi'(x,y,t)$ and, for small disturbances,

$$(\partial_t + u_0\partial_x)Q' - u_{0yy}\psi'_x = 0\ , \qquad \psi' = \nabla^{-2}Q'\ , \tag{6.6}$$

essentially the Rayleigh equation. It is no surprise, therefore, to find that the simplest classical shear instabilities, such as the instability of the $u_0 = \tanh y$ shear layer, or its Rayleigh (piecewise linear) counterpart, can be understood (consistently with the Rayleigh, Fjørtoft and Arnol'd stability theorems) in terms of a coupled pair of Rossby waves, in a certain frame of reference, each of which propagates against the local mean flow and phase-locks with the other in such a way as to bring it to rest. Furthermore, each Rossby wave makes the other grow exponentially, via a reduced phase shift between disturbance velocities v' and the sideways displacements η' of the Q contours (Lighthill 1963, Bretherton 1966, Hoskins *et al.* 1985 §6b). The Miles (1957) wind–wave instability is in some ways fundamentally similar, except that one of the Rossby waves is replaced by the surface gravity wave, and there is a mismatch between the strengths of the gravity and Rossby elasticities leading among other things to relatively slow growth.

7. ROSSBY-WAVE BREAKING: A DEFINITIVE EXAMPLE

In Figure 6 the Q (material) contours are depicted as undulating reversibly, the situation described by linearized, dissipationless wave theory. Indeed, for linearized theory to be self-consistently and generally applicable, along with associated concepts such as the principle of superposition, the undulations must also be gentle. Strictly speaking, the sideways slopes $(\partial y/\partial x)_Q$ must be infinitesimal. The opposite extreme, that of infinite sideways slopes followed by sideways overturning, and rapid, irreversible deformation, lengthening and folding of the Q contours, is a commonplace occurrence and can be recognized as a Rossby-wave version of 'wave breaking'. An idealized example is shown in Figure 7; see also Figure 9 below. This phenomenon, in a variety of forms, is ubiquitous in the real atmosphere and oceans and plays an important role in the irreversible transport of momentum and angular momentum by Rossby waves.

The case of Figure 7, known as the Stewartson–Warn–Warn (SWW) Rossby-wave critical-layer solution, is now described in more detail. It is important out of all proportion to the restrictive idealizations used because it is an unequivocal example, described by an analytical solution, of irreversible wave-induced angular momentum transport due *solely* to wave breaking, with no other wave dissipation mechanisms involved.

We return to the dynamical system described by (6.1)–(6.3), but now introduce an undisturbed flow $u_0(y)$ having constant shear u_{0y}, say. Again the Q contours are made to

Figure 6. Sketch of the Q contours and the Q-anomaly (Q') pattern, and the induced velocity field giving rise to the sideways 'Rossby elasticity' in a simple, non-breaking Rossby wave.

undulate, this time by introducing a gently undulating boundary near some value y_0 of y, exciting Rossby waves of constant phase speed c and y-lengthscale u_{0y}/β. The boundary displacement amplitude is of order $\epsilon\, u_{0y}/\beta$ where ϵ is a small dimensionless parameter. The x-wavelength $2\pi/k$ of the undulation is assumed long enough, $\gg 2\pi\, u_{0y}/\beta$, for the inverse Laplacian ∇^{-2} in (6.2) to be approximated by $(\partial/\partial y)^{-2}$, simplifying the mathematics. (The lengthscale u_{0y}/β then corresponds to the lengthscale l^{-1} in (6.4) when $c \simeq l^{-1}u_{0y}$.) Under these restrictions and with suitable choices of y_0, this problem can be solved analytically (Stewartson 1978, Warn and Warn 1978). The Q contours behave in an approximately undular manner except in a narrow region surrounding the critical line, or y-location where $u_0(y) = c$. This narrow region, the 'critical layer', is separated from the undulating boundary by a distance $\gtrsim u_{0y}/\beta$ and has width of order

$$b = \epsilon^{1/2}\, u_{0y}/\beta \ . \tag{7.1}$$

The 'inner problem' for this region is mathematically the same as the nonlinear pendulum problem, incompressible fluid flow replacing incompressible phase-space flow. It is analytically soluble in terms of elliptic functions, and the solution confirms that the region is, indeed, a region of Rossby wave breaking in the sense envisaged. Figure 7 shows the solution at four successive times. Q contours are overturning sideways, and deforming in a manifestly irreversible way.

Let us suppose that the resulting rearrangement of the Q field has, in a coarse-grain view, something of the character of a mixing layer as shown in idealized form in Figure 8a. Then it is a trivial matter to see that there must be an associated irreversible transport of momentum. Let $\delta\bar{Q}(y)$ be the change in Q represented by the difference between the solid and dashed lines in Figure 8a, and $\delta\bar{u}(y)$ the corresponding change in u. Application of the inversion operator (6.2) gives (since $\delta\bar{Q} = \delta\bar{Q}_{\rm r} = -\delta\bar{u}_y$ here) the parabolic profile

$$\delta\bar{u}(y) = -\int_{-\infty}^{y} \delta\bar{Q}(\tilde{y})\, d\tilde{y} = \begin{cases} \tfrac{1}{2}\beta(y^2 - \tfrac{1}{4}{b_{\rm m}}^2) & (-\tfrac{1}{2}b_{\rm m} < y < \tfrac{1}{2}b_{\rm m}) \\ 0 & (y < -\tfrac{1}{2}b_{\rm m} \text{ or } y > \tfrac{1}{2}b_{\rm m}) \end{cases} \tag{7.2}$$

where $b_{\rm m}$ is the breadth of the mixing layer. Negative momentum

$$\delta M \propto \int_{-\infty}^{\infty} \delta\bar{u}(y)\, dy = -\tfrac{1}{12}\beta {b_{\rm m}}^3 \tag{7.3}$$

has been transported irreversibly into the region where Q has been rearranged. This phenomenon is robust: any change $\delta\bar{Q}(y)$ qualitatively like that in Figure 8a will be associated with a momentum change of the order of magnitude, and sign, indicated by (7.3). For instance the SWW solution has a limiting value of δM as $t \to \infty$, which can be expressed in the form (7.3) with the value of $b_{\rm m}$ shown by the bar at the centre of Figure 7d. This value of $b_{\rm m}$ may be thought of as an 'effective mixing width' for the SWW solution. In this case the momentum (7.3) has come from the wave source comprising the undulating boundary at large Y. When β, or more generally the initial gradient $\partial(Q_{\rm e}+Q_0)/\partial y$, is positive, this momentum is necessarily negative — another manifestation of the one-signedness of Rossby dynamics.

The SWW solution provides us, incidentally, with a definitive counterexample to arguments saying that Q-mixing scenarios like that of Figure 8a are impossible because they violate momentum conservation. These arguments overlook the possibility of wave-induced momentum transport from outside the region. Further historical remarks and references, going back to various issues surrounding, for instance, G.I. Taylor's 'vorticity transfer theory', and its recent developments including links with the photon analogy, can be found in my 1992 review and in the paper with Norton (1990).

Another interesting point about Figure 7 is that the predicted Q configuration becomes shear-unstable after the contours first overturn. So if any noise is present initially, the actual evolution is quite different in detail. In typical cases the result is an apparently chaotic form of Rossby wave breaking, and an increase in the effective mixing width $b_{\rm m}$ by a modest factor such that δM ($\propto {b_{\rm m}}^3$) increases in magnitude by a factor 2 or 3. Figures 8b,c, from the definitive study by Haynes (1989), show the $\bar{Q}$ and $\delta\bar{u}$ profiles, defined as Eulerian x-averages, in one such case. The $\delta\bar{u}$ profile has the approximately parabolic form suggested by the idealization (7.2).

The foregoing examples are conceptually important in another way, already hinted at. The inviscid, two-dimensional fluid-dynamical system under consideration is known to have mathematically regular behaviour, over arbitrarily long time intervals. The examples are therefore cases of wave dissipation and irreversible momentum transport that do not depend on overtly dissipative processes like viscosity. The irreversibility involved is a purely fluid-dynamical irreversibility, precisely that associated with the persistent lengthening of the Q contours as time goes on, and familiar from other fluid-dynamical paradigms such as 'random straining' and 'turbulence' (e.g. Batchelor 1952).

8. THE DEFINING PROPERTY OF WAVE BREAKING, THE STRATOSPHERIC 'SURF ZONE', AND OZONE CHEMISTRY

What should one mean by wave breaking for general, non-acoustic* waves in fluids? Even in the most familiar case, ordinary surface gravity waves, the phenomenon usually recognized as breaking has an extensive 'zoology' of shapes and time-evolutions. The same is true of internal gravity and Rossby waves. The question does not seem to have any natural answer from a 'zoological' or morphological viewpoint. However, a natural answer does suggest itself if one wants the concept of 'wave breaking' to be relevant to the general question of wave-induced momentum transport, or, more precisely, to the question of when wave-induced momentum transport becomes irreversible.

One can then use the rapid, irreversible material-contour deformation illustrated above as the defining property of wave breaking. 'Rapid' means that deformation rates are comparable, at least, to the local intrinsic wave frequency. Such a definition is entirely compatible with the accepted phenomenology, and relevance, of wave breaking in the case of surface gravity waves and longshore ocean-beach currents. The case for such a generalization is carefully argued in three papers with T. N. Palmer (1983–5). It avoids 'zoological' definitions, and requirements to decide whether 'turbulence' is involved, but does take account of the relevant general theorems, particularly Kelvin's

*Shock formation in acoustic waves seems best kept conceptually separate, if only because its essential dependence on the existence of overtly dissipative processes (microscopic irreversibility).

Figure 7. Contours of constant Q for the SWW solution, a special but clear-cut example of Rossby-wave breaking (contrast the purely undular, non-breaking Rossby wave in Figure 6), and the consequent irreversible wave-induced momentum transport. See also Figure 9. The y-scale has been expanded using the re-scaled coordinate $Y = y/b$ with $b = \epsilon^{1/2}\, u_{0y}/\beta$; the range $-5 \le Y \le 5$ is plotted. Four successive stages in the evolution are shown, at times 1, 1.5, 2, 3 in units of $2^{1/2}(k\, b\, u_{0y})^{-1}$, where k is the x-wavenumber. The vertical bar in panel (d) gives the effective mixing scale $b_{\mathrm{m}} \simeq b$; cf. (7.2), (7.3) and Figure 8a. From Killworth and McIntyre (1985), after Stewartson (1978) and Warn and Warn (1978).

(a) y (b) Y (c) Y

Figure 8. Q-mixing scenarios and associated momentum change; see (7.2), (7.3): (a) the simplistic, but qualitatively relevant, idealization that assumes perfect Q-mixing over width b_{m}; (b) an actual Eulerian-mean $\bar{Q}(y)$ profile from an accurate, quasi-chaotic Rossby-wave critical-layer solution (P. H. Haynes, personal communication); and (c) the approximately parabolic $\delta\bar{u}(y)$ profile corresponding to (b), showing the momentum change due to the Rossby wave breaking (first of (7.2) and of (7.3)). As in Figure 7, the re-scaled coordinate $Y = y/b$ is used, over the same range -5 to 5 ($b = \epsilon^{1/2}\, u_{0y}/\beta$).

circulation theorem and the way it manifests itself in exact, formally complete theories of wave–mean interaction. The relevant material contours are defined to be those that would otherwise undulate reversibly under the generalized elasticity, or restoring mechanism, that gives rise to the wave propagation.

Figure 9, of which an animated version was seen in the lecture, shows a less idealized example of Rossby wave breaking and its effects, taken from the work of Norton (1993). The parameter conditions are far closer to those in the real stratosphere than those assumed by critical-layer theory. Here the dynamical system comprises the shallow-water equations on a sphere, solved numerically by a high-resolution pseudospectral method. The projection is polar stereographic, and the winter northern hemisphere is shown. An axisymmetric initial state is disturbed by smoothly distorting the lower boundary in a large-scale pattern so as to imitate the effect, on the real stratosphere, of planetary-scale Rossby waves propagating up from the much denser troposphere below (Charney and Drazin 1961). The result looks remarkably similar to what is seen in the real winter stratosphere at altitudes of the order of 25 to 50 km.

The left panel of Figure 9 shows Q, now the shallow-water potential vorticity, defined as absolute vorticity over local layer depth. The model problem still has the generic form (6.1) to excellent approximation, but with a shorter-range potential-vorticity inversion operator* $\mathbf{I}(\cdot)$. The central region is the model's 'stratospheric polar vortex', where the relevant material contours, which in this experiment lie initially along latitude circles, are almost coincident with the Q contours and undulate nearly reversibly. In this region there is little Rossby-wave breaking. This is illustrated by comparison with the right panel, which shows the behaviour of some material contours computed very accurately using a high-precision 'contour advection' technique adapted from the work of Dritschel (1988). Outside the polar vortex, in middle latitudes, is a region in which the waves are breaking vigorously. There, the initially-latitudinal material contours are deformed rapidly and irreversibly, and mixed into a broad, 2D-turbulent 'Rossby-wave surf zone'; e-folding times for contour lengthening were estimated to be about 4 days (Norton, *op. cit.*). This 'surf zone' is the real-stratospheric counterpart of the idealized Rossby-wave critical layer. It is far broader, and very different in detail, but it illustrates equally well the robustness and one-signedness of the angular momentum transport associated with irreversible Q-rearrangement. As already suggested, such Rossby-wave breaking is an important part of how the mean circulation illustrated in Figure 2 is driven. Furthermore, its recognition in models of the global-scale stratospheric transport of trace chemicals like CFCs is beginning to lead to improved realism in the predictions of such models, both via a more realistic mean circulation, and also via a more realistic representation of the quasi-horizontal turbulent transport (Garcia *et al.* 1992).

This latter aspect may be especially important for ozone photochemistry (and relevant to some current controversies about ozone depletion — see §10 of my 1992 review). This is because ozone photochemistry by its nature could be more sensitive than, for instance, CFC photochemistry, to the timing of a typical molecule's excursions across the 'surf zone'. Timescales for such excursions are comparable, at certain altitudes, to photochemical timescales. Together with the mean circulation itself, these complicated fluid motions control the rate at which ozone is produced photochemically, mainly in the high tropical stratosphere, and carried thence to the extratropical lower stratosphere where it accumulates (unless destroyed by 'ozone-hole chemistry') in far greater concentrations than can be produced by tropical photochemistry alone.

Model simulations like that of Figure 9 are also relevant to understanding the Antarctic ozone hole. The contours in the right-hand panel of Figure 9 can be regarded as isopleths of an advected passive tracer, the advection being very accurately simulated, with no artificial diffusion. The simulation shows that, at least in the model, chemical substances in the surf zone do not penetrate past the region of strong Rossby elasticity concentrated in the steep Q gradient near the vortex edge. This is believed to be important

*To rough approximation, this $\mathbf{I}(\cdot)$ is given by (6.2) with ∇^{-2} replaced by $(\nabla^2 - \kappa^2)^{-1}$, corresponding to an elastic membrane tethered by local springs, somewhat like a spring mattress, with a latitude-dependent e-folding scale κ^{-1} (the 'Rossby radius') of about 1400 km at the pole, 2000 km at 45°N, and 3000 km in the tropics. McIntyre and Norton (1990, 1993) give examples of much more accurate (but much more elaborate) potential-vorticity inversion operators $\mathbf{I}(\cdot)$ for shallow-water models.

Figure 9. Winter hemisphere in a high-resolution, shallow-water numerical model of the stratosphere, from Norton (1993). The mean depth is 4 km, and the numerical resolution (triangular truncation at total wavenumber 127) corresponds to a mesh size roughly 1° latitude. *Left panel*: potential vorticity Q, contour interval $4 \times 10^{-9}\,\mathrm{m}^{-1}\mathrm{s}^{-1}$, zero contour dotted. (Negative Q values would, in a three-dimensional model, signal the three-dimensional mode of Rossby-wave breaking pointed out by O'Sullivan and Hitchman 1992.) *Right panel*: isopleths of an advected tracer field, initially axisymmetric and coincident with the Q contours, showing the fluid-dynamical irreversibility characteristic of Rossby-wave breaking and two-dimensional 'turbulence'. This was computed with near-perfect accuracy using a high-precision 'contour advection' techique, introduced independently by Norton (1993) and by Waugh and Plumb (1993) using an algorithm developed in another context by Dritschel (1988).

for ozone-hole chemistry. The same phenomenon has been demonstrated in the laboratory by Sommeria *et al.* (1989, 1991). The contour-advection technique, conceived of as a benchmark numerical tracer advection algorithm, was introduced independently by Norton (*op. cit.*) and by Waugh and Plumb (1993).

Recently, O'Sullivan and Hitchman (1992) have shown that an entirely different mode of Rossby-wave breaking is possible near the equator, where the potential vorticity itself changes sign. It conforms to the general wave-breaking definition, with three-dimensional rather than layerwise-two-dimensional material contour deformations, arising from an asymmetric inertial (quasi-centrifugal) instability. Among other things, this may have new implications for the quantitative modelling of the QBO.

9. WIND-GENERATED WATER WAVES: TWO NEW MECHANISMS?

What does this old but elusive problem look like from the foregoing perspective? The first point is that Rossby-wave dynamics is involved, albeit on much faster timescales than before. The velocity profile in the air, whether or not approximately logarithmic, will usually have a strong curvature u_{0yy} near the water surface (y vertical). Therefore there is a vertical gradient of spanwise vorticity and hence, in the present language, a Rossby elasticity, in the airflow just above the water waves; recall (6.5) and Figure 6. That is why the Miles inviscid wind–wave instability, for instance, can be regarded as a coupled 'Rossby-wave, gravity-wave' instability. One can think in terms of a pair of phase-locked, counterpropagating waves, a backward-propagating Rossby wave in the air coupled to a forward-propagating gravity wave in the water. (Other examples of Rossby-wave, gravity-wave instabilities go back to G. I. Taylor's work in the 1930s on the 'Taylor–Goldstein equation'; see also, e.g., Griffiths *et al.* (1982), Hayashi and Young (1987), and Sakai (1989).) In the Miles instability, the Rossby elasticity is relatively weak. This implies not only that the water wave largely determines the phase speed c, but also that the growth rate is slow and that the most significant Rossby effects occur near the critical line $u_0(y) = c$, where intrinsic phase speeds are slow. The mathematical underpinning for these statements is well known and is given in detail by Miles (1957) and Lighthill (1962).

Now the self-consistency of the linearized instability theory requires that conditions near the critical line resemble those in Figure 7a. In the present language, the Rossby wave is not only weak and slow, but is also in the earliest stages of wave breaking, in the generalized sense already referred to. The later stages can be expected to look more like those of Figure 7b–7d and beyond, with a small value of the effective mixing scale $b_{\rm m}$, viewed in a suitably undulating coordinate system. This suggests that the idealized scenario represented by the inviscid Miles mechanism lacks robustness, when taken literally. The delicate phase-locking and synchronization of the normal mode's displacement fields will be disrupted by Rossby wave breaking before much vertical rearrangement of spanwise vorticity can take place (and long before the water waves break). The outcome will be net vorticity rearrangement, in the airflow, over only a small effective mixing depth $b_{\rm m}$, of the same order as the scale b given by (7.1) except that the lengthscale u_{0y}/β is replaced by a typical value of $-u_{0y}/u_{0yy}$ near the critical line. There is a correspondingly small net wave-coherent momentum transport $\delta M \propto {b_{\rm m}}^3$ from air to water, having the order of magnitude implied by (7.3) with β again replaced by a typical value of $-u_{0yy}$ near the critical line.

However, there is one important feature of this idealized scenario that does, on the other hand, look robust. This is the *sign* of the net wave-coherent momentum transport, which tends to be such that the water waves are amplified when travelling in the same direction as the wind. More precisely, the sign is determined by the sign of $-u_{0yy}$ and the one-signedness of the associated Rossby-wave dynamics — the same one-signedness that so strongly controls the sense of the mean circulation throughout most of Figure 2. In the wind–wave problem the sign is determined quite independently of whether or not we have strictly x-periodic waves and delicately synchronized exponential, normal-mode growth.

This sign-robustness must imply a kind of 'ratchet effect'. Almost any spatio-temporal intermittency — whether it be any tendency of the water waves to arrange themselves in groups through, for instance, subharmonic instabilities, or any gustiness of the wind that might be modelled as an intermittency in quantities like $-u_{0yy}$ — will tend to favour intermittent wave-coherent momentum transport whose effects are cumulative. They may possibly also be such as to reinforce the intermittency. There is no longer any reason, moreover, why the effective depths $b_{\rm m}$ associated with any Rossby wave breaking should be especially small.

More generally, the sign-robustness points to the likely effectiveness of any process in which wave-coherent undulation of the airflow over the water waves causes Rossby wave activity in the air to increase in total amount, or to dissipate, or both. The relevant measure of 'activity' for this purpose is a suitably defined quasimomentum (Killworth and McIntyre 1985; Shepherd 1990, & refs.), since the photon analogy can be shown to apply here, in the required sense. Any such process will result in wave-coherent momentum transport in such a sense as to amplify the water waves. This could well be important for the real wind–wave problem, despite the added complexities of turbulence and other three-dimensional effects. There appear, in particular, to be two distinguishable mechanisms whose possible role deserves closer attention, both theoretically and experimentally.

One is simply local Rossby wave breaking characterized by relatively large effective depths b_{m}. If, for instance, a gust produces an effective $-u_{0yy}$ that then encounters a wave group, leading to rapid vertical rearrangement of spanwise vorticity (reducing $-u_{0yy}$), then the water waves will be correspondingly amplified for a short time. The transient nature of such phenomena would add to the well-known difficulties of observing the associated pressure phase shift experimentally. The other mechanism, which in reality might tend to occur at the same time, is the formation of Rossby lee waves in the airflow. This will give rise to 'wave resistance' in the usual way, not depending on wave breaking of any kind. Indeed the lee-wave mechanism should be at its most efficient when both Rossby wave breaking and water wave breaking are unimportant. Then one should get unseparated flow over the water waves (e.g. Banner 1990, & refs.) and, for at least some $-u_{0yy}$ profiles, an efficient shedding of Rossby-wave quasimomentum into the downstream airflow. It seems possible that the amplifying wave group in the experiments of Chu *et al.* reported by Professor Phillips, which they call the 'short wave energy front', might depend on the Rossby lee-wave mechanism — and the experiments might offer a chance to study it under controlled conditions.

These ideas might also help resolve some longstanding questions about the effects of three-dimensional turbulence on different timescales, from the relatively long, 'dissipative' timescales required to restore $-u_{0yy}$ profiles, to the relatively short, 'elastic' timescales on which rapid-distortion theory might be appropriate, merely modifying the Rossby elasticity in the air — as with internal gravity waves in the important observational study by Finnigan and Einaudi (1981), in which eddy-viscosity or other steady-state dissipative turbulent modelling was shown to be wrong by 'an order of magnitude'. In some wind–wave models, turbulent processes are modelled semi-empirically for instance by adding something like an eddy viscosity to the Miles theory. Such devices (drastically modifying the type of behaviour suggested by Figure 7) would seem unavoidable in models that insist on x-periodicity, or on horizontal statistical stationarity and linear superposability of wave-coherent processes. But a physically correct turbulence model might need to bring in a wider range of timescales. This suggests that it could be fruitful to try to model the spatio-temporal intermittency of the water waves and the airflow explicitly.

Acknowledgements: Early versions of the ideas developed here were part of an essay that shared the 1981 Adams Prize in the University of Cambridge. Their development has been influenced by many colleagues going back to early ideas from F. P. Bretherton, D. O. Gough, T. Matsuno, R. E. Peierls, E. A. Spiegel, and most recently P. H. Haynes, W. A. Norton, and other co-authors with whom it has been a pleasure to work. The experimental demonstrations in Figure 3 were stimulated by one of M. S. Longuet-Higgins' beautiful wave-tank demonstrations, and G. I. Barenblatt stimulated me to think harder about real turbulent airflow over water. Support from the Natural Environment Research Council (through the UK Universities' Global Atmospheric Modelling Project and through the British Antarctic Survey), the Science and Engineering Research Council, the UK Meteorological Office, and the US Office of Naval Research is gratefully acknowledged.

REFERENCES

Andrews, D. G., Holton, J. R., Leovy, C. B., 1987: *Middle Atmosphere Dynamics.* Academic Press, 489 pp.

Andrews, D. G., McIntyre, M. E., 1976: Planetary waves in horizontal and vertical shear: the generalized Eliassen–Palm relation and the mean zonal acceleration. J. Atmos. Sci., **33**, 2031–2048. [See pp. 2043–5]

Andrews, D. G., McIntyre, M. E., 1978a: An exact theory of nonlinear waves on a Lagrangian-mean flow. J. Fluid Mech., **89**, 609–646. [Note the caveats about Lagrangian means in §10 (regarding cases like that of Figure 7 above), and in McIntyre 1980.]

Andrews, D. G., McIntyre, M. E., 1978b: On wave-action and its relatives. J. Fluid Mech., **89**, 647–664 (Corrigendum **95**, 796).

Armi, L., Hebert, D., Oakey, N., Price, J., Richardson, P. L., Rossby, T., Ruddick, B., 1988: The history and decay of a Mediterranean salt lens. Nature, **333**, 649–651.

Banner, M. L., 1990: The influence of wave breaking on the surface pressure distribution in wind–wave interactions. J. Fluid Mech., **211**, 463–495.

Barnett, J. J., Corney, M., 1985: Middle atmosphere reference model derived from satellite data. In: Handbook for MAP, vol. **16**, ed. K. Labitzke, J. J. Barnett, B. Edwards, pp. 47–143. Available from SCOSTEP secretariat, University of Illinois, 1406 W. Green St, Urbana, Ill. 61801, U.S.A.

Batchelor, G. K., 1952: The effect of homogeneous turbulence on material lines and surfaces. Proc. Roy. Soc. **A 213**, 349-366.

Booker, J. R., Bretherton, F. P., 1967: The critical layer for internal gravity waves in a shear flow. J. Fluid Mech., **27**, 513–539.

Bretherton, F. P., 1966: Baroclinic instability and the short wavelength cut-off in terms of potential vorticity. Q. J. Roy. Meteorol. Soc., **92**, 335–345.

Bretherton, F. P., 1971: The general linearised theory of wave propagation. Amer. Math. Soc. Lectures in Applied Mathematics, **13**, 61–102. [See §6.]

Bretherton, F. P., Garrett, C. J. R., 1968: Wavetrains in inhomogeneous moving media. Proc. Roy. Soc. Lond., **A 302**, 529–554.

Brillouin, L., 1925: On radiation stresses. Annales de Physique, **4**, 528–586. (In French.)

Charney, J. G., Drazin, P. G., 1961: Propagation of planetary-scale disturbances from the lower into the upper atmosphere. J. Geophys. Res., **66**, 83–1O9.

Chu, J. S., Long, S. R., Phillips, O. M., 1992: Measurements on the interaction of wave groups with shorter wind-generated waves. J. Fluid Mech., **245**, 191–210

Davis, C. A., 1992: Piecewise potential vorticity inversion. J. Atmos. Sci., **49**, 1397–1411.

Dickinson, R. E., 1968: Planetary Rossby waves propagating vertically through weak westerly wind waveguides. J. Atmos. Sci., **25**, 984–1002. [See p. 1001 for the suggestion about the QBO.]

Dickinson, R. E., 1969: Theory of planetary wave–zonal flow interaction. J. Atmos. Sci., **26**, 73–81.

Dritschel, D. G., 1988: Contour surgery: a topological reconnection scheme for extended integrations using contour dynamics. J. Comput. Phys., **77**, 240–266.

Dunkerton, T. J., 1983: Laterally-propagating Rossby waves in the easterly acceleration phase of the quasi-biennial oscillation. Atmos.–Ocean, **21**, 55–68.

Finnigan, J. J., Einaudi, F., 1981: The interaction between an internal gravity waves and the planetary boundary layer. Part II: effect of the wave on the turbulence structure. Q. J. Roy. Meteorol. Soc., **107**, 807–832.

Fritts, D. C., 1984: Gravity wave saturation in the middle atmosphere: a review of theory and observations. Revs. Geophys. Space Phys. **22**, 275–308.

Fritts, D. C., 1993: Gravity wave sources, source variability and lower and middle atmospheric effects. In: Coupling Processes in the Lower and Middle Atmosphere, ed. E. V. Thrane, T. A. Blix, D. C. Fritts; Dordrecht, Kluwer, 191–208.

Garcia, R. R., Stordal, F., Solomon, S., Kiehl, J. T., 1992: A new numerical model of the middle atmosphere. 1. Dynamics and transport of tropospheric trace source gases. J. Geophys. Res., **97**, 12967–12991.

Garrett, C., Smith, J., 1976: On the interaction between long and short surface waves. J. Phys. Oceanography, **6**, 925–930.

Gill, A. E., 1982: *Atmosphere-Ocean Dynamics.* Academic Press, 662 pp.

Gray, W. M., Sheaffer, J. D., Knaff, J. A., 1992: Influence of the stratospheric QBO on ENSO variability. J. Meteorol. Soc. Japan, **70**, 975–995.

Green, J. S. A., 1970: Transfer properties of the large-scale eddies and the general circulation of the atmosphere. Q. J. Roy. Meteorol. Soc., **96**, 157–185.

Griffiths, R. W., Killworth, P. D., Stern, M. E., 1982: Ageostrophic instability of ocean currents. J. Fluid Mech., **117**, 343–377.

Hamilton, K., Garcia, R. R., 1984: Long-period variations in the solar semidiurnal atmospheric tide. J. Geophys. Res., **89**, 11705–11710.

Hayashi, Y.-Y., Young, W. R., 1987: Stable and unstable shear modes of rotating parallel flows in shallow water. J. Fluid Mech., **184**, 477–504.

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.

Haynes, P. H., Marks, C. J., McIntyre, M. E., Shepherd, T. G., Shine, K. P., 1991: On the "downward control" of extratropical diabatic circulations by eddy-induced mean zonal forces. J. Atmos. Sci., **48**, 651–678.

Haynes, P. H., *et al.*, 1993: Notes on the UK National Summer School in Geophysical and Environmental Fluid Dynamics. In preparation.

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. Also **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.

Lighthill, M. J., 1962: Physical interpretation of the mathematical theory of wave generation by wind. J. Fluid Mech., **14**, 385–398.

Lighthill, M. J., 1963: Boundary layer theory. In: Laminar Boundary Layers, ed. L. Rosenhead; Oxford University Press, 46–113. [See p. 93.]

Lighthill, M. J., 1978a: *Waves in Fluids.* Cambridge University Press, 504 pp.

Lighthill, M. J., 1978b: Acoustic streaming. J. Sound Vib., **61**, 391-418.

Lindzen, R. S., Holton, J. R., 1968: A theory of the quasi-biennial oscillation. J. Atmos. Sci., **25**, 1095–1107.

Lindzen, R. S., Tsay, C. Y., 1975: Wave structure of the tropical stratosphere over the Marshall Islands area during 1 April – 1 July 1958. J. Atmos. Sci., **32**, 2008–2021.

Longuet-Higgins, M. S., 1969: Action of a variable stress at the surface of water waves. Phys. Fluids, **12**, 737–740.

Longuet-Higgins, M. S., Stewart, R. W., 1962: Radiation stress and mass transport in gravity waves, with application to 'surf beats'. J. Fluid Mech., **13**, 481–504.

Lorenz, E. N., 1967: *The Nature and Theory of the General Circulation of the Atmosphere.* Geneva, World Meteor. Org., 161 pp. [See especially pp. 85, 150.]

Loudon, R., Paige, E. G. S., 1991: Alan Frank Gibson. Biogr. Mem. F. Roy. Soc. London, **37**, 221–244.

McIntyre, M. E., 1972: On Long's hypothesis of no upstream influence in uniformly stratified or rotating flow. J. Fluid Mech., **52**, 209–243.

McIntyre, M. E., 1973: Mean motions and impulse of a guided internal gravity wave packet. J. Fluid Mech., **60**, 801–811.

McIntyre, M. E., 1980: Towards a Lagrangian-mean description of stratospheric circulations and chemical transports. Phil. Trans. Roy. Soc. Lond., **A 296**, 129–148.

McIntyre, M. E., 1981: On the 'wave momentum' myth. J. Fluid Mech., **106**, 331–347.*

McIntyre, M. E., 1992: Atmospheric dynamics: some fundamentals, with observational implications. In: Proc. Internat. School Phys. "Enrico Fermi", CXV Course, ed. J. C. Gille, G. Visconti; Amsterdam, Oxford, New York, Toronto, North-Holland, pp. 313–386. [A list of bibliographical updates and minor corrections is available from the author, on paper or by email (*mem2@phx.cam.ac.uk*). For caveats about Lagrangian means, see also the 1980 paper cited above.]

McIntyre, M. E., Mobbs, S. D., 1993: On the 'quasimomentum rule' for wave-induced mean forces on obstacles immersed in a fluid medium. To be submitted to Proc. Roy. Soc. Lond. [See McIntyre 1973 for an *exception* to the rule.]

McIntyre, M. E., Norton, W. A., 1990: Dissipative wave-mean interactions and the transport of vorticity or potential vorticity. J. Fluid Mech., **212**, 403–435 (G. K. Batchelor Festschrift Issue.); Corrigendum **220**, 693.

McIntyre, M. E., Norton, W. A., 1993: Potential-vorticity inversion on a hemisphere. J. Atmos. Sci., to appear.

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.

Miles, J. W., 1957: On the generation of surface waves by shear flows. J. Fluid Mech., **3**, 185–204.

Munk, W. H., Miller, G. R., Snodgrass, F. E., Barber, N. F., 1963: Directional recording of swell from

*There is a slip on p. 339, sixth line from bottom: 'length of the wavetrain' should read 'wavelength'. Also, I now think that the statement on p. 338 about a horizontal distance of order $H \gg ct$ (with $c = (gH)^{1/2}$, tenth line from bottom) is wrong. The main points are unaffected.

distant storms. Phil. Trans. Roy. Soc. Lond., **A 255**, 505–584.
Norton, W. A., 1993: Breaking Rossby waves in a model stratosphere diagnosed by a vortex-following coordinate system and a contour advection technique. J. Atmos. Sci., submitted.
Okuda, K., 1982: Internal flow structure of short wind waves. Part 2. The streamline pattern. J. Oceanog. Soc. Japan, **38**, 313–322.
O'Sullivan, D. J., Hitchman, M. H., 1992: Inertial instability and Rossby wave breaking in a numerical model. J. Atmos. Sci., **49**, 991–1002.
Peierls, R., 1991: *More Surprises in Theoretical Physics.* Princeton, 106 pp.
Phillips, O. M., 1977: *The Dynamics of the Upper Ocean*, 2nd edition. Cambridge University Press, 336 pp.
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.
Rayleigh, Lord, 1896: The theory of sound, volume 2. New York, Dover (reprinted 1945), 504 pp. [See §352.]
Raymond, D. J., 1992: Nonlinear balance and potential-vorticity thinking at large Rossby number. Q. J. Roy. Meteorol. Soc., **118**, 987–1015.
Rhines, P. B., 1975: Waves and turbulence on a beta-plane. J. Fluid Mech., **69**, 417–443.
Sakai, S., 1989: Rossby–Kelvin instability: a new type of ageostrophic instability caused by a resonance between Rossby waves and gravity waves. J. Fluid Mech., **202**, 149–176. [See also Griffiths *et al.* 1982, Hayashi and Young 1987.]
Saravanan, R., 1993: Idealized modelling of tropospheric processes in the mechanism of the quasi-biennial oscillation. Manuscript in preparation. (The author may be contacted at the National Center for Atmospheric Research, Boulder, CO., U. S. A.)
Shepherd, T. G., 1990: Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics. Adv. Geophys., **32**, 287–338.
Snodgrass, F. E., Groves, G. W., Hasselmann, K. F., Miller, G. R., Munk, W. H., Powers, W. H., 1966: Propagation of ocean swell across the Pacific. Phil. Trans. Roy. Soc. Lond., **A 259**, 431–497. [All the way to Alaska, via several recording stations including New Zealand.]
Solomon, S., Kiehl, J. T., Garcia, R. R., Grose, W., 1986: Tracer transport by the diabatic circulation deduced from satellite observations. J. Atmos. Sci., **43**, 1603–1617.
Sommeria, J., Meyers, S. D., Swinney, H. L., 1989: Laboratory model of a planetary eastward jet. Nature, **337**, 58–61.
Sommeria, J., Meyers, S. D., Swinney, H. L., 1991: Experiments on vortices and Rossby waves in eastward and westward jets. In: Nonlinear Topics in Ocean Physics, ed. A. R. Osborne; Amsterdam, North-Holland, 227–269.
Starr, V. P., 1968: *Physics of negative viscosity phenomena.* McGraw-Hill, 256 pp.
Stewartson, K., 1978. The evolution of the critical layer of a Rossby wave. Geophys. Astrophys. Fluid Dyn., **9**, 185-200.
Takahashi, M., Boville, B. A., 1992: A three-dimensional simulation of the equatorial quasi-biennial oscillation. J. Atmos. Sci., **49**, 1020–1035.
Teitelbaum, H., Vial, F., Bauer, P., 1993: The stratospheric quasi-biennal oscillation observed in the semidiurnal ground pressure data. J. Climate, submitted.
Thomas, G. E., Olivero, J. J., Jensen, E. J., Schröder, W., Toon, O. B., 1989: Relation between increasing methane and the presence of ice clouds at the mesopause. Nature, **338**, 490–492.
Thorncroft, C. D., Hoskins, B. J., McIntyre, M. E., 1993: Two paradigms of baroclinic-wave life-cycle behaviour. Q. J. Roy. Meteorol. Soc., **119**, 17–55.
Thorpe, A. J., 1985: Diagnosis of balanced vortex structure using potential vorticity. J. Atmos. Sci., **42**, 397–406.
Wallace, J. M., Holton, J. R., 1968: A diagnostic numerical model of the quasi-biennial oscillation. J. Atmos. Sci., **25**, 280–292.
Van Dyke, M., 1982: *An Album of Fluid Motion.* Stanford, Parabolic Press.
Warn, T. and Warn, H., 1978. The evolution of a nonlinear critical level. Stud. Appl. Math., **33**, 2021-2024.
Waugh, D. W., Plumb, R. A., 1993: Contour advection with surgery: a technique for investigating fine scale structure in tracer transport. J. Atmos. Sci., submitted.
Yoden, S., Holton, J. R., 1988: A new look at equatorial quasi-biennial oscillation models. J. Atmos. Sci., **45**, 2703–2717.