How well do we understand the dynamics of stratospheric warmings?

Michael E. McIntyre

J. Meteorol. Soc. Japan., **60**, 37–65 (1982)

How Well do we Understand the Dynamics of Stratospheric Warmings?

By Michael E. McIntyre

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, U.K.
(Manuscript received 19 October 1981)

Abstract

Ever since Matsuno's pioneering numerical simulations of the stratospheric sudden warming there has been little reason to doubt that this spectacular natural phenomenon is essentially dynamical in origin. But theoretical modelling, and the use of satellite observations, are only just reaching the stage where there seem to be prospects of understanding stratospheric warmings in some detail and forecasting them reasonably well. An informal discussion of recent progress is given, and suggestions are made for future work, including a way of avoiding spurious resonances in mechanistic numerical models in which tropospheric motions are prescribed *a priori*.

1. Introduction

The stratospheric sudden warming is a large-scale experiment which nature kindly performs for us from time to time. It is one of the crucial tests of our understanding of the dynamics of the middle atmosphere, and indeed of atmospheric dynamics in general. The dynamics, in turn, is one of the necessary ingredients in attempts to understand the stratospheric circulation in general, and to predict the effects of pollution on the ozone layer in particular. Recent progress in the dynamical theory, along with the more uniform global coverage facilitated by infrared satellite observations, have led to better ways of analysing the observations and to more illuminating comparisons with computer simulations, and our understanding of stratospheric warmings is now advancing significantly. There are signs moreover that this understanding may be leading to fresh insights into other, at first sight unrelated, phenomena, for instance the nonlinear behaviour of mid-latitude, tropospheric depressions (Hoskins, 1982).

In this article I do not propose to give a comprehensive review of the literature on stratospheric warmings (for which the reader may consult Quiroz *et al.*, 1975, McInturff, 1978 and Schoeberl, 1978), but rather to concentrate on some recent developments with which I have been in touch. Indeed the subject is moving so fast that an informal discussion is probably the most appropriate thing at present. Some of the work whose implications I shall discuss (most of it not my own) was presented at the IAMAP symposium on the general circulation held at the University of Reading in August 1981. I am indebted to a number of colleagues for permitting me to draw upon the results of their work in press or in progress.

Hardly anyone who has followed observational and theoretical work on major stratospheric warmings can be in much doubt today as to the essentially dynamical nature of the phenomenon. Its enormous depth scale, covering several scale heights, and its "suddenness" compared to estimates of diabatic time scales at least in the lower stratosphere, makes it pretty obvious that dynamically-induced air-parcel descent is required to account for the observed temperature rises. The self-consistency of this view has been well checked by the results of many numerical simulations, including the first such study by Matsuno (1971). These mechanistic, or hypothesis-testing simulations have consistently reproduced at least the final, "sudden" stage of the process in a qualitatively convincing way (see the reviews by Quiroz *et al.*, 1975 and Holton, 1975). And in the simulations, at least, there is no doubt whatever that the large temperature rises are induced adiabatically, by descent of air parcels in the polar cap (*e.g.* Hsu, 1980; Dunkerton *et al.*, 1981); the fact that this is compatible with ascending Eulerian-mean motion is well known by now, Mahlman (1969) having apparently been the first to point it out in the present context. Some of the effects of experimenting with diabatic time scales in the models can be seen from studies such as those of Holton (1976) and Schoeberl and Strobel (1980a). Finally, the weight of observational and theoretical evidence leaves little room for doubt that large-amplitude "planetary waves", in the sense of large, planetary-scale disturbances to the zonal wind in the stratosphere, especially those involving zonal harmonic wavenumbers 1 and 2, are an essential ingredient in the process and not merely an accompaniment to it.

I shall take all the foregoing for granted, then, until proven otherwise, and suggest that the fundamental questions of interest today begin with those in the following list. Of course several of them are questions which we should ask about the behaviour of theoretical and numerical models as well as of the real atmosphere.

1. How and why do planetary-wave amplitudes become anomalously large?

2. When they do become large, for what stratospheric conditions is a major warming likely to occur (and why are major warmings relatively uncommon)?

3. To what extent can we use linear planetary-wave theory for the wave structure? And in particular,

4. to what extent can we think of the principal zonal wavenumbers, 1 and 2, as acting independently of each other? In other words, how much can we explain without invoking nonlinear interactions between different zonal wavenumbers?

5. Are wave-reflection and resonance phenomena important or not (*e.g.* to question 1)?

6. Are "critical lines" important or not (*e.g.* to question 2, or to question 5)?

7. Are shear instabilities involved at any stage, and are they relevant to question 1?

8. What quantities should be monitored in order to be able to forecast warmings?

9. To what extent, and in what sense, does the troposphere behave independently of the stratosphere (for the purposes of question 1 for instance), and how should we represent tropospheric-stratospheric coupling in mechanistic models?

In recent years we have been coming a good deal closer to answering some of these questions. A major reason is the impact of data from satellite-borne infrared radiometers. In particular, the global coverage from satellites has permitted more reliable estimates of the stratospheric circulation to be made on a daily basis, enabling a close dialogue between theory and observation to take place for the first time. The observations are not only giving a much better idea of how, for instance, Eulerian-mean zonal wind profiles change from day to day, but even some idea (Butchart *et al.*, 1982; Chapman and Miles, 1981; Kanzawa, 1980; Kanzawa and Hirota, 1981; O'Neill and Youngblut, 1982; Palmer, 1981a, b; Simmons, 1982b) of the harder-to-estimate quantities which theory tells us must be central to the dynamics. These include isentropic potential-vorticity gradients and associated planetary-wave refractive indices, and the convergence of the Eliassen-Palm wave flux. Of course we have been lucky, scientifically speaking, in that the essential phenomena really do seem to have the deep vertical scales already remarked on, so that they can be seen rather well by the satellite radiometers. Were this not so, the picture now emerging could hardly have the degree of dynamical self-consistency which it seems to have.

There has been at least one other piece of scientific good luck. Nature decided to present the satellites and the FGGE observers with a specially significant sequence of events in January and February 1979. As I shall now argue, that sequence of events, culminating in the major, wave-2-dominated warming of February 1979, contains some particularly important clues about the dynamics.

2. Why wave 2, and why January-February 1979?

It might be asked why we should be specially interested in wave 2, when many of the warmings observed during the past decade seem to have been more or less dominated by wave 1. I think that it is precisely the comparative rarity of wave-2-dominated warmings that makes them unusually interesting. In some ways they are the severest test of our understanding—particularly as regards question 2 on my list. It seems very likely that, in order to get a major warming, stationary planetary waves must not only attain large amplitudes, and have phase tilts of the type usually associated with propagation from below, but must also be unusually well focused into the high-latitude polar cap, say latitudes $\geq 60°$. The small mass and moment of inertia of that region gives the waves by far their best chance of causing dramatic effects. Now observation and theory both suggest that although stationary waves 1 and 2 can often propagate quite happily up from the troposphere into the wintertime stratosphere, they also have a general tendency to propagate equatorwards, *away* from the polar cap, much as one might expect from a consideration of the spherical geometry of the earth. For instance if one were to start a wave propagating horizontally along a latitude circle, at high latitudes, it would tend to go off at a tangent, along something like a great circle path (Hoskins and Karoly, 1981). The resulting tendency for the waves to avoid the high-altitude polar cap, and propagate into the much larger areas available elsewhere, could be called "defocusing" for want of a better term. It is probably one reason why major warmings do not happen more often. Defocusing tends to be more pronounced for wave 2 than for wave 1, as originally found by Matsuno and confirmed by subsequent studies, most recently the ray-tracing calculations of Karoly and Hoskins (1982) reported elsewhere in this issue.

In order to quantify such things as focusing and defocusing, it is important in practice to use a conservable measure of wave propagation. The use for instance of eddy fluxes of geopotential, as a measure of wave propagation, tends to obscure the issue. They are strongly affected by the local strength of the mean westerlies in a way that has nothing to do with focusing (Eliassen and Palm, 1961; Bretherton and Garrett, 1968). A similar thing happens with measures of wave amplitude such as eddy geopotential height, which in a tight polar-night jet tends to be roughly proportional to the local jet speed (Simmons, 1974, Eq. 10), representing the speed with which air parcels travel through a given streamline or latitudinal displacement pattern (cf. Edmon *et al.*, 1980, §2d). So inspection of the magnitudes of such quantities in a latitude-height cross-section may tell us little more than where the jet is. They are also very deceptive as indicators of such things as vertical propagation times. The eddy fluxes of geopotential do, to be sure, tell us the direction of the planetary-wave group velocity in a meridional plane, to the extent that the notion of group velocity is valid, a question which I shall touch on in section 3. But their magnitudes are potentially misleading as a guide to whether or not the waves are converging onto a given height and latitude.

It is fortunate, therefore, that there is another measure of wave flux which is just as good for indicating the direction of the group velocity (when that notion applies), but which is both easier to compute from observations and is also a true *conservable* measure of the flux of wave activity across an arbitrary zonal-wind profile. That is, it does not converge unless either the waves are building up transiently at the place in question, or there are some dissipative or other departures from conservative motion. This is the so-called Eliassen-Palm wave flux F. It has a number of other useful properties, to some of which I shall refer later. Its horizontal component is proportional to minus the Eulerian northward eddy momentum flux, and its vertical component proportional to the northward eddy heat flux. For a recent review the reader may consult section 2 of the paper by Edmon *et al.* (1980). Examples of the use of the EP flux to describe stratospheric planetary waves have been given by Butchart *et al.* (1982), Dunkerton *et al.* (1981), Kanzawa and Hirota (1981), O'Neill and Youngblut, 1982, Palmer (1981a, b), and Sato (1980). I have been using the term "Eliassen-Palm cross-section" to refer to latitude-height cross-sections showing both F and its divergence.

The pair of EP cross-sections shown in Figs. 1a and 1b gives an excellent illustration of the defocusing of wave 2 under "typical" or "climatological" conditions. Fig. 1a is taken from Dunkerton *et al.* (1981; *q.v.* for further details) and Fig. 1b from unpublished work by C.-P. F. Hsu (personal communication). They were obtained from a pair of model warming simulations using Hsu's (1980) modification of the semi-spectral model developed by Holton (1976), in which wave 2 was forced at an artificial lower boundary in much the same way as in Matsuno (1971). (The extent to which this is a valid procedure will emerge later, when we address question 9 on the list.) Figs. 1a and 1b represent an early stage in the simulations, before the mean state has changed very much. Each simulation was started with exactly the same initial state, a zonally-symmetric state close to the kind of climatological zonal mean which has often been used in such modelling studies, with a broad polar-night jet merging smoothly into the still broader mesospheric westerlies (Hsu, 1980, Fig. 1a). The convergence of F shown by the negative

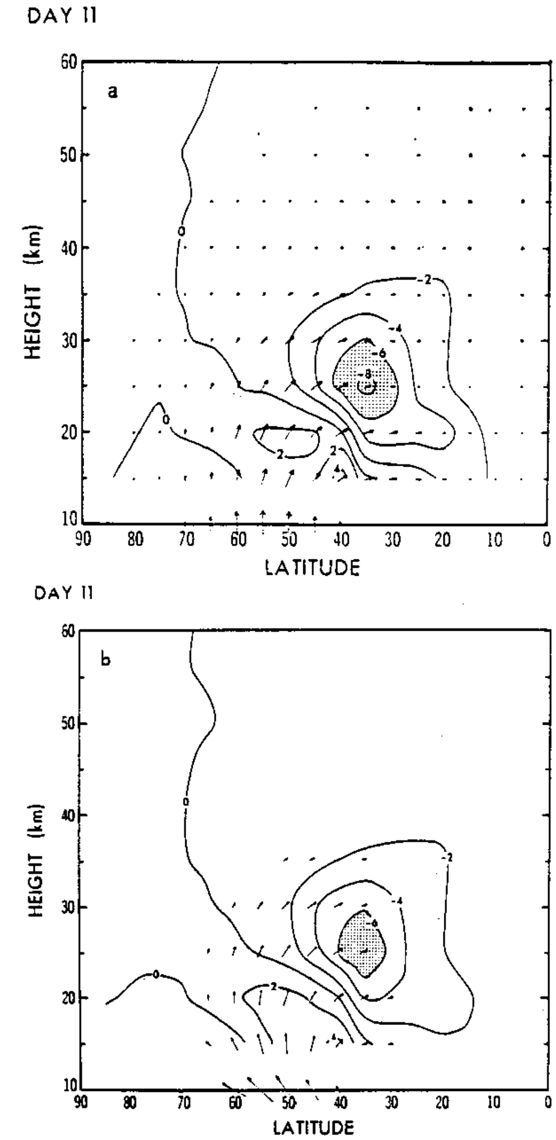

Fig. 1 Eliassen-Palm cross-sections for a pair of model simulations in which planetary waves of zonal wavenumber 2 are generated by applying two different lower boundary conditions. The waves propagate on a basic zonal wind profile typical of what has usually been taken as a representative climatological mean. The arrows represent the Eliassen-Palm wave flux and the contours its divergence, plotted in accordance with the conventions described in Dunkerton *et al.* (1981), from which case (a) is taken. Case (b) is from unpublished work by C.-P. F. Hsu, with kind permission. The arrow scales are such that the arrow patterns look non-divergent if and only if the flux is nondivergent.

contour values is attributable mainly to transience, simply showing where the waves are arriving—a process which is being slowed down by the proximity of a subtropical zero-wind line. Similar cross-sections for earlier times in the simulations show the waves emerging from the boundary and then turning equatorwards. (They also show clearly how the waves speed up as they get into slightly stronger westerlies, and *vice versa*, just as suggested by the ray propagation times shown in Fig. 7 of Karoly and Hoskins, 1982, page 119 in this issue.) The phrase "*turning* equatorwards" refers of course to the way things appear in the meridional cross-section.

The difference between the two simulations is due solely to different lower boundary conditions. The purpose of changing the boundary condition was to try to persuade the waves to focus into the polar cap simply by forcing F to point poleward at the boundary. This was done by imposing a boundary forcing with a southeast-northwest phase tilt, an experiment suggested by O'Neill and Taylor (1979). It is striking how easily the defocusing effect frustrates this attempt. The same phenomenon appears to be implied by the theoretical and observational results shown in Figs. 5b and 6b of Matsuno (1970) and in Fig. 1 of Hirota and Sato (1969). Above 20 km, the waves hardly seem to notice the difference, and turn towards the equator regardless. Indeed the subsequent evolution of the two simulations from the stage shown in Fig. 1 was astonishingly similar, even as regards the timing of the various stages of mean-flow evolution described by Dunkerton *et al.* (*op. cit.*). Essentially similar results have been independently obtained by Butchart *et al.* (1982), using a different numerical model.

Results like these add to the growing body of evidence suggesting that, no matter what the troposphere is doing, conditions in the stratosphere have to be prepared in some special way before a major warming can take place (Butchart *et al.*, 1982; Dunkerton *et al.* 1981; Kanzawa, 1980; Labitzke, 1981; Palmer 1981b; Quiroz et al., 1975, §2b), especially a warming dominated by wave 2. Something is needed which can overcome the defocusing effect and guide planetary waves upwards into the polar cap. For wave 2 we may expect the requirements to be more stringent than for wave 1. Whatever these requirements are, it is clear that they were well satisfied just before the major warming around 20 February 1979. Fig. 3b of Palmer (1981a) shows strong focusing of wave 2 from below on 19 February—the direction of F was tilted well in towards the pole, a state of affairs quite the reverse of that shown in our Fig. 1a.

This is reason enough for paying special attention to the case of February 1979. But there are further reasons. Without this case, we would have significantly less evidence concerning what the requirements for focusing might actually be. Most observed warmings appear to involve not only planetary waves of very large amplitude, but also more than one wave component simultaneously. It is possible that the mere presence of one large-amplitude wave might help to focus another. If this kind of nonlinear, wave-wave interaction in the stratosphere were an essential ingredient in the warming process, a thorough investigation of it by numerical experimentation would be a daunting task indeed. The strength of such interactions is likely to be sensitive to all kinds of variables, including details of the basic state chosen, and the number of possibilities to be explored before full understanding could be claimed would be enormous. It is here that February 1979 has provided a clue of the first importance in the scientific detective story. As I shall explain, it is a specially clear example in which the focusing of wave 2 seems unlikely, in fact, to have depended crucially on the presence of other wave components. This is a very direct piece of evidence bearing on question 4 in my list.

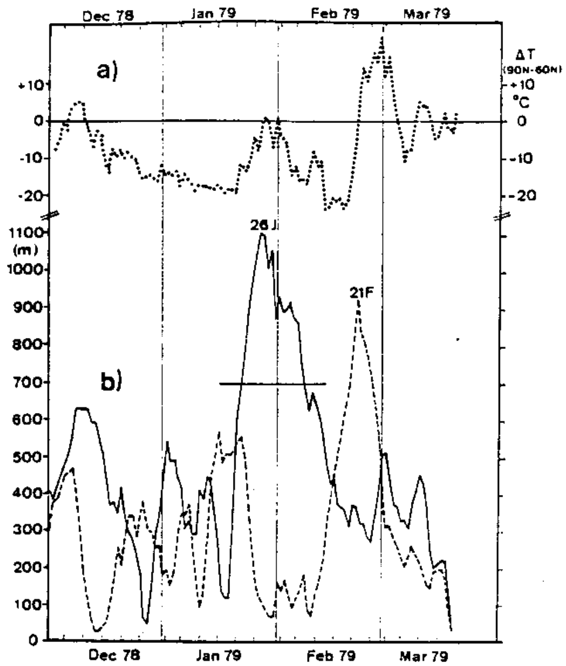

Fig. 2 (a) Difference at 30 mb, around 24 km altitude, between the temperature at the north pole and the zonally averaged temperature at 60°N; (b) amplitudes in metres of zonal harmonic geopotential height waves 1 and 2 (broken line) at 60°N and 30 mb. After Labitzke (1981).

This evidence is strengthened and its value enhanced by the fact that, in many other respects, the events of January and February 1979 seem to follow a pattern not untypical of other major warmings, especially those of 1967-68 and 1970-71. Fig. 2, adapted from Labitzke (1981), shows for instance the observed time variation of the wave amplitudes, and of the latitudinal temperature contrast in the polar cap, at 30 mb during January and February 1979. The upper curve (a) shows the zonal-mean temperature difference between the north pole and 60°N, and the lower curves (b) the geopotential height amplitudes of waves 1 and 2 at 60°N. These curves may be compared with an essentially similar series for the years 1964-81 reproduced in the article by Labitzke (1982, Fig. 1, page 127 in this issue). See also the review by Schoeberl (1978). In the present case, the time evolution of the wave amplitudes follows more or less what Schoeberl calls the "type A" pattern, in which a small wave-2 pulse (occurring in mid-January in this case) is followed by a large peak in wave 1, and finally by another wave-2 pulse of variable strength at about the time of the main temperature rise. Also quite variable are the exact timing of the main temperature rise and the final wave-2 pulse. The present case would appear to be an extreme case mainly in one respect, which is the unusually long delay of about four weeks between the final wave-2 pulse and its wave-1 predecessor. The delay is about double what has been seen in other cases. Tropospheric observations for the same period (*e.g.* Fig. 4 of Quiroz, 1979 or Fig. 3 of Labitzke, 1981) suggest, as Dunkerton *et al.* remark, that "the warming occurred late simply because no wave-2 forcing was available earlier" from the troposphere. Whatever the reason for the delay, one consequence of it was that the final wave-2 pulse, and the warming itself, took place with only a comparatively modest amount of wave 1 present. We can scarcely avoid the impression that to get a major warming of this general type it may not be essential to have waves 1 and 2 both present simultaneously.

What then was the role of the huge wave-1 peak in late January, which gave rise only to a "minor" warming? The foregoing suggests that its essential effect could only have been to *pre-*

condition the basic state so that it was able subsequently to focus wave 2. The idea that effects of this kind might be important has recently been stressed by several authors (Butchart *et al.*, 1982; Dunkerton *et al.*, 1981; Kanzawa, 1980; Labitzke, 1981; Palmer, 1981b). It is certainly an observational fact, evident from even a cursory inspection of stratospheric synoptic maps for January-February 1979, that the basic state was very different after the big wave-1 event from what it was before. Afterwards the polar-night jet was much narrower. The dynamical implications of this warrant further discussion, and I shall take them up in the next two sections. I am not saying, of course, that wave-wave interactions were of no significance at all. It would be naive to think so, and we already know from model experiments that they can matter at least for getting the details right (*e.g.* Butchart *et al.*, 1982; Hsu, 1981; Lordi *et al.*, 1980). However, it now seems very likely that there are real cases in which wave-wave interaction is not the most dynamically fundamental effect, and it seems sensible to try to develop a really good understanding of such cases first.

I have been keeping up my sleeve yet another noteworthy recent development, which once again concerns the stratospheric events of January-February 1979, and which incidentally removes any lingering doubts, if such remain, as to the essentially dynamical character of those events. The main January and February events were successfully forecast, in considerable detail, by a numerical model being developed at the European Centre for Medium Range Weather Forecasts (ECMWF) (Simmons, 1982b). The January event appeared to be quite well simulated starting from 16 January up to the main wave-1 peak in Fig. 2 ten days later, and the February event from 13 February at least up to the splitting of the vortex around 20 February. This is the first truly successful forecast of a real sudden warming which I have heard of. The model was an experimental version of the high-resolution forecasting model at present being used operationally at the ECMWF. It has a hybrid (sigma, p) vertical coordinate system with the top level at $p=10$ mb, instead of the 25 mb ($\sigma=0.025$) top level of the operational model. The operational, sigma-coordinate model did, in fact, capture some aspects of the events in the stratosphere as well, including the splitting of the vortex at 50 mb (Bengtsson *et al.*, 1982). The existence of simulations like these implies, for one thing, new opportunities for the extraction of the dynamically relevant diagnostics which put such a strain on observational data. A realistic, high-resolution simulation can provide "data" which is more dynamically consistent than raw observations.

Daily maps of Ertel's potential vorticity in isentropic surfaces would be one important example. The availability of such maps would open the way to all kinds of refinements in our understanding of the dynamics—things which are more or less well hidden by present-day diagnostics, EP cross-sections included! For instance I suspect that isentropic potential-vorticity maps might make it obvious why there was a second phase of mean zonal deceleration in the polar cap around 26 February, well after the splitting of the polar vortex (Palmer 1981a). A rough estimate of advection times in each half of the split vortex seems consistent with a simple explanation in terms of the advection of potential-vorticity "debris" around the two cutoff lows. This idea might also help to explain the failure of a lower-resolution simulation (Butchart *et al.*, 1982, to be referred to in the next section) to reproduce this further episode of mean zonal deceleration, since following the "debris" in detail after the vortex splits would immediately place a considerable strain on numerical resolution. Such a regime of motion marks the point at which the eddies have largely ceased to be wave-like, dynamically speaking; one may say that they have "saturated", or "broken". It is quite like what happens to ocean waves on a beach. An even closer analogy is the breaking of tides and internal gravity waves in the mesosphere, since in that case there is no question of having two immiscible fluids like air and water, and so the basic gradient (static stability) to which the waves owe their existence is mixed irreversibly. Similarly, irreversible mixing of the isentropic potential-vorticity gradient may be regarded as the distinguishing feature of a breaking planetary or Rossby wave. As soon as such wave-breaking occurs, the detailed potential-vorticity distribution will become very complicated, and "wave-wave" interactions (between very many zonal harmonics) will be prominent in any detailed description. Indeed thinking too literally in terms of "waves", in the dynamical sense, may not then be very profitable.

3. A test of the focusing hypothesis

The main hypothesis suggested by the observations, as discussed so far, is that the unusual focusing of wave 2, just before the major warming of February 1979, was simply due to an unusual configuration of the basic zonal state. This hypothesis has been directly tested in a beautifully-conceived series of numerical experiments carried out at the U.K. Meteorological Office by Butchart *et al.* (1982). The results provide very strong support for the essential correctness of the hypothesis, and in the process resolve the long-standing question as to why a different overall behaviour was found in Matsuno's wave-2 simulation and its successors. The model used was a more elaborate one than Matsuno's, being a finite-difference, primitive-equation model with somewhat more zonal resolution than would correspond to Matsuno's semi-spectral truncation to one zonal wavenumber. The model did not attempt to represent the troposphere. It was forced in much the same way as Matsuno's, by prescribing the geopotential artificially at 100 mb.

The essential point which these numerical experiments establish is simply that the results are, indeed, sensitive to the basic state adopted.* They clearly vindicate "the need for special care in the choice of initial conditions for model simulations" suggested by Quiroz *et al.* (1975, §2b). If a climatological mean state is used as in Matsuno's simulation and its successors, then the model exhibits just the familiar behaviour first found by Matsuno for wave 2, which begins in the defocused way illustrated by Fig. 1 above and hence cannot produce a warming without first undergoing a long period of mean-flow evolution, typically twenty days or more. The way in which enough focusing is eventually achieved to give rise to a warming in this kind of wave-2 model experiment has been elucidated by Dunkerton *et al.* (1981). The focusing depends upon the partially-reflecting properties of the nonlinear "critical layer" associated with a zero-wind line which moves northwards from the subtropics and reflects the waves back into higher latitudes, a bit like an artificial side wall.

If, on the other hand, the actual mean state on 16 February 1979 is used, with its much narrower polar-night jet, then there is immediate focusing without any preliminary period of mean-flow evolution. Moreover, if care is taken to use a forcing having a zonal phase speed of the order of that observed at 100 mb, which was eastward and significantly different from zero, then even with pure wave-2 forcing the focusing persists long enough to produce a strong warming. This takes only about ten days, despite the fact that the initial state is zonally symmetric. It should be remembered of course that the real stratospheric circulation on 16 February was far from being zonally symmetric, as is obvious from Fig. 2. When the actual initial conditions and the actual 100 mb forcing were used, so far as could be determined from the observations, then the model came closer still to imitating the warming that actually occurred. A warming looking quite like the real one was achieved in less than *five* days.

Butchart *et al.* argue persuasively that some important aspects of the model's behaviour could have been anticipated by inspection of meridional cross-sections of the refractive index squared, R^2, appropriate to a linear, wave-2 disturbance propagating steadily at the prescribed phase speed on the Eulerian zonal-mean state (Charney and Drazin, 1961; Matsuno, 1970). R^2 is the basic quantity entering into the "ray theory" of planetary-wave refraction in a meridional cross-section. It contains a term proportional to the latitudinal isentropic gradient of potential vorticity, divided by the velocity of the mean zonal wind relative to the wave. Rays tend to bend towards regions of large, positive R^2 (Palmer, 1981b; Karoly and Hoskins, 1982, §2d, page 112 in this issue); and Butchart *et al.* found that the EP wave flux in the model tended to behave in a similar way. In particular, the spatial distribution of R^2 seemed to account satisfactorily for the focused and defocused initial EP flux patterns found for the actual and climatological mean states. Results carrying similar implications have been obtained by O'Neill and Youngblut (1982), from an observational study of the January 1977 warming which included some ray-tracing calculations based on suitably smoothed observational estimates of R^2 at different times. It is quite remarkable how well the refractive-index and ray-tracing concepts, and by implication the concept of group velocity, seem to succeed in predicting important aspects of the behaviour of the wave fluxes. Meridional and vertical wavelengths, and temporal rates of change of the mean state, are all too large for the relevance of those concepts to be self-evident—to say nothing of the presence of additional complications such as interference between stationary and travelling wave components, which can cause transient fluctuations in the direction and magnitude of the wave fluxes (*e.g.* Boyd, 1976; Madden, 1975; Palmer, 1981a,

* *Note in proof:* See also Bridger and Stevens (1982).

appendix; Schoeberl and Strobel, 1980a).

It seems certain, then, that computations of R^2 and associated ray paths are going to be an important aid to understanding the results of future numerical experiments on stratospheric planetary waves. A cross-section of R^2, at least, would be virtually indispensable, before one could tell *a priori* whether there was any possibility of a given, narrow polar-night-jet profile focusing wave 2. In this situation the refractive index is very sensitive to the precise shape of the jet velocity profile. For any given profile there is always a "tunnelling" region of negative R^2, or imaginary R, near the pole, which rays by definition cannot enter. Rather, they bend away from it, accounting for some of the defocusing effect already discussed. The negative values are due to a term in R^2 proportional to minus the zonal wavenumber squared, which for the moment I shall call somewhat loosely the "defocusing term" even though it is not in fact the only term which can cause defocusing. Its magnitude quadruples when we go from wave 1 to wave 2, and as the pole is approached it always dominates the term involving the potential-vorticity gradient, as a result of geometrical factors multiplying the two terms.* In order to have positive R^2 somewhere in the polar cap, which is a necessary condition for rays to be able to enter that region at all, the competition between the two terms has to go the other way—the potential-vorticity-gradient term must dominate the defocusing term—somewhere in the polar cap. Because of the geometrical factors this has its best chance of happening if the largest potential-vorticity gradients are concentrated towards the *south* side of the jet maximum, the side furthest from the pole, as suggested by the heavy curves in Fig. 5a below. (Potential vorticity Q is shown on the left and zonal velocity u on the right: the thin curves suggest climatologically typical profiles.) It appears from Butchart *et al*'s results that the mean state on 16 February had essentially this configuration in the lower stratosphere, and that R^2 for wave 2 was not only positive on the south flank of the jet, but had a positive maximum there, both for waves with zero phase speed and for waves with the observed phase speed. Because of the local maximum this is a configuration capable of causing focusing, but as we shall see shortly this is not the whole story! Generally speaking, the competition between the two terms under discussion means that in the polar cap, especially for wave 2, even the sign of R^2 can be sensitive to the precise shape of the zonal velocity profile, which has to be differentiated twice to get the gradient of Q. A recent series of numerical simulations of linear planetary-wave behaviour (Lin, 1982) confirms the expected sensitivity by showing how it takes only small changes in the shape of the jet to change the refractive-index configuration completely and give a drastically different pattern of wave propagation.

Cross-sections of R^2 can hardly be expected to tell us everything, on the other hand. And there is one specially important limitation on their validity which has not yet been discussed. Broadly speaking, the notions of refractive index, ray theory and group velocity are likely to work best in strong westerlies, but worst near "critical lines" where mean zonal wind velocity equals zonal phase velocity. Their validity, even for linear Rossby waves, fails utterly *at* a critical line (Andrews and McIntyre, 1976 Appendix B; Grimshaw, 1980). Moreover, if we use synoptic maps to estimate the nonlinear advective terms for typical planetary-wave amplitudes we find that linear wave theory itself fails utterly as well. In the real atmosphere, a critical line will always be surrounded by a nonlinear "critical layer" in which the waves saturate, or break, in the sense referred to at the end of the last section, and within which R^2 is completely irrelevant.

As originally suggested by the idealized models of Benney and Bergeron (1969), Davis (1969), and others, such a nonlinear critical layer can act as a reflector once the waves have broken

* See e.g. Matsuno (1970, eq. 11). It should be noted that some authors prefer to work with a quantity corresponding to R^2 with the defocusing term omitted, so as to be able to plot cross-sections which apply to more than one zonal wavenumber. The defocusing term must be included, however, if tunnelling regions are to correspond to regions of negative R^2. In Fig. 5c of Karoly and Hoskins (1982, page 117 in this issue), the tunnelling regions for zonal wavenumber n are those with contour values less than n. Note also that Karoly and Hoskins' definitions incorporate the geometrical effects into a transformation to Mercator coordinates. Palmer (1981b) adopts a different coordinate transformation which leads to another definition of refractive index, corresponding to Matsuno's definition divided by the sine of latitude squared. This transformation assumes that static stability is independent of latitude, but is then very convenient since it makes the propagation appear uniformly isotropic in height and latitude.

(quite irrespective of how R^2 behaves in its immediate neighbourhood). The general circumstances under which this nonlinear Rossby-wave reflection can occur have been clarified recently, and they will be discussed in section 5. They appear to include the circumstances of present interest, to a large extent, both for the real stratosphere and for mechanistic models of it. Zonally truncated mechanistic models cannot properly represent the wave-breaking process itself, but surprisingly (and very fortunately) they do manage to imitate the nonlinear reflection in a crude but qualitatively not unreasonable way, as was first shown by Geisler and Dickinson (1974) and further explained by Dunkerton *et al.* (1981, appendix B).

Butchart *et al.* note the possibility that reflection from a critical layer situated to the south of the polar-night jet might have played a role in maintaining the focusing which occurred in their most realistic numerical experiments, just as it did in the late stages of the simulation discussed by Dunkerton *et al.* The most realistic experiments, in which the focusing persisted long enough to induce a strong warming, were just those in which the phase speed of the wave forcing was realistic, as well as the initial mean state. With a phase speed of the right order (Butchart *et al.*, Fig. 1b) one finds, for instance from the mean zonal velocity cross-section in Butchart *et al*'s Fig. 2a, that a critical line did exist in middle latitudes on 16 February, extending from below 30 mb to above 5 mb. Its shape and position were, in fact, remarkably similar to the shape and position of the mid-latitude zero-wind line found a few days before the major warming in Dunkerton *et al*'s model simulation (*op. cit.*, Fig. 2c), in which the waves were nearly stationary. The subsequent evolutions were also quite similar. If we set these facts alongside the theoretical evidence presented by Dunkerton *et al.*, it seems not only possible, but indeed practically certain, that reflection from the associated critical layer must have been taking place in all the experiments with realistic phase speeds.

Not only must reflection from the critical layer have been taking place, but I see no escape from the further conclusion that the reflection must have been the primary reason, if not the *only* reason, for the persistence of focusing which characterized all the experiments with realistic phase speeds. The persistence cannot be explained simply by persistence of the initial local maximum in R^2. That maximum disappeared as the polar-night jet decelerated, at least in those cases for which cross-sections of R^2 were displayed. These included two experiments with realistic initial mean states one of which had zero and the other a realistic phase speed. In the former case there was no critical layer to stop the defocusing effect from reasserting itself as soon as the local maximum in R^2 disappeared, and defocusing was exactly what then happened. The disappearance of the maximum in both cases is hardly surprising when one recalls its sensitivity to the jet configuration.

It would be tempting at this point, and not unreasonable on the evidence so far, to conclude that the foregoing statements apply also to the real warming of February 1979. The critical line was certainly present*, and it is likely that the associated wave-breaking region would have tended to act as a reflector. What is less easy to be sure about is the behaviour of the refractive index, not only because of data problems but also because in the real atmosphere, as opposed to a truncated model, there is less reason to suppose that the Eulerian zonal mean is a good basis for estimating R^2. I shall return to the latter problem in my concluding remarks.

It is worth adding one more remark about critical lines. If a critical line happens to lie within a region where large-scale, isentropic gradients of potential vorticity are anomalously weak in the first place, then the region will act as a reflector even before wave breaking takes place. All critical-layer theories agree on this (*e.g.* Tung, 1979, Eq. 46 with $\hat{\beta}_c=0$).† The reason is that the absorption predicted by the usual linear, transient, critical-layer theory (Dickinson, 1970; Warn and Warn, 1976) depends on the development, through advection in the early stages preceding wave breaking, of a certain pattern of eddy potential vorticity in the critical-layer region (Stewartson, 1978; Warn and Warn, 1978, Fig. 2b). That pattern could not develop if there were no large-scale isentropic potential-vorticity gradient across the critical-layer region to start with. Of course when the waves do break, the consequent mixing of potential vorticity tends to ensure that the large-scale potential-vorticity gradients *become* weak even if they were not weak in the first place; this in essence is the nonlinear critical-layer reflection

* Contrary to a tacit assumption by Dunkerton *et al.*, that wave 2 was stationary in the real atmosphere on the relevant days in February 1979.

† See also p. 53a below

mechanism. In the present instance, however, there is good reason to suppose that in mid-February the large-scale, isentropic potential-vorticity gradients were already weak, on average, in middle latitudes, as suggested schematically by the heavy curve in Fig. 5a below. It is here that the wave-1 precursor comes into the story.

I should not risk leaving the reader with the impression that Butchart *et al.* claimed to have explained every last detail of the February 1979 warming solely in terms of interaction between the mean state and wave 2. On the contrary, they found that even the modest amount of wave 1 present in the actual initial conditions and 100 mb forcing seemed to be significant for approximating the observed behaviour in a very small region within a radius of ten degrees' latitude or so from the north pole. By way of comparison, the main region of zonal-mean easterlies resulting from the February 1979 warming covered a much larger area, out to a radius of about thirty degrees, *i.e.* to about 60°N. Butchart *et al.* point out that truncation errors due to finite differencing near the pole may have been significant in their simulations, and in view of the fact that the otherwise very similar final stages of Dunkerton *et al*'s pure-wave-2 simulation did not corroborate this detail, it is perhaps a cause for some concern. But it can also be remarked that sensitivity to wave 1 very near the pole is what one might expect in any case from a synoptic viewpoint. The somewhat artificial process of taking zonal means within ten degrees of the pole can obviously give quite variable results with even the slightest departure from wave-2 symmetry.

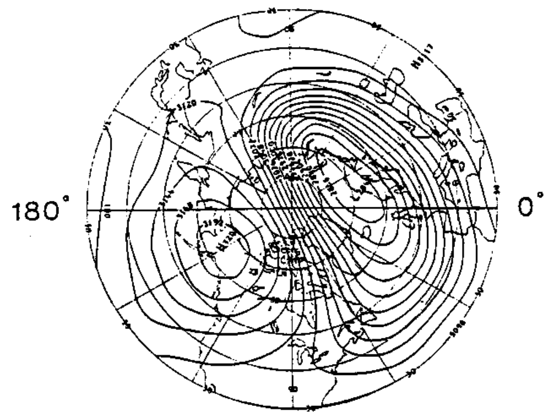

Fig. 3 Breaking planetary wave at 00Z on 27 January 1979, as shown by the height of the 10 mb constant-pressure surface. Contour interval is 24 dekametres; H=32.09 km, L=28.42 km. From Pick (1979). Latitudes north of 30°N are shown.

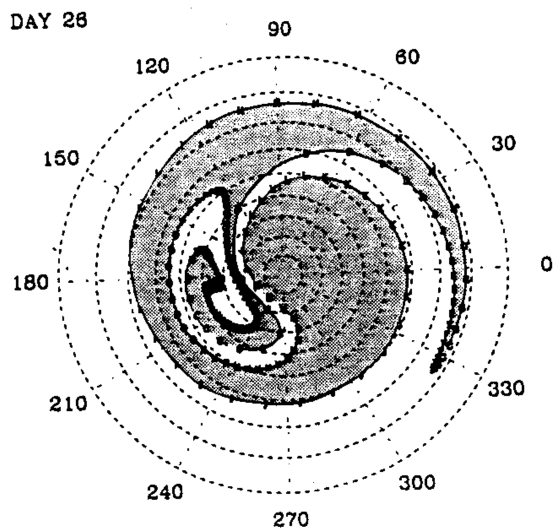

Fig. 4 Shape of a material line, originally coincident with the 30°N latitude circle at an altitude of about 31 km, and then advected by a wind field of a qualitatively similar pattern to that implied geostrophically by Fig. 3. The wind field was generated in a mechanistic model simulation. From Hsu (1981). Whole hemisphere is shown.

4. The wave-1 precursor

What then of events prior to 16 February? In particular, how did the mean state take up a "non-climatological" configuration with a narrow polar-night jet favourable to wave-2 focusing? The synoptic maps, for example those for geopotential height at 10 mb between say 23 January and 10 February, show beyond reasonable doubt that the wave-1 event which dominated that period must have caused a great deal of quasi-horizontal mixing of potential vorticity in middle latitudes. Throughout that period, the Aleutian high was well developed as a cutoff high spanning a range of latitudes reaching from the subtropics to about 80°N, as illustrated for 27 January by Fig. 3, taken from Pick (1979). The closed-streamline circulation around this huge system must inevitably have been advecting potential vorticity straight across the planetary gradient, twisting up the isopleths of potential vorticity in each isentropic surface like spaghetti on a fork. Much the same thing would have been happening to the isopleths of ozone and other quasi-conservative tracers. A model calculation giving some idea of the kinematics of the process is shown in Fig. 4, taken from Hsu (1981), in which the light areas represent low-potential-vorticity air from the tropics and *vice versa*. The boundary between the two air masses is a material boundary which was initially coincident with the 30°N latitude circle, before a pure wave-1 disturbance was switched on. This of course is another example of what I have been calling a breaking, or saturating, planetary wave. To that extent it is essentially the same as what goes on in a nonlinear critical layer, but now on such a grand latitudinal scale that the term "layer" begins to be something of a misnomer.

Note incidentally that in terms of streamfunction, approximately equal to geopotential height divided by Coriolis parameter, the Aleutian high would have been centered further south in Fig. 3, but that Fig. 3 stops at 30°N whereas Fig. 4 shows the whole hemisphere. Of course the correspondence is intended to be qualitative at most.

Exactly what an actual potential-vorticity map would look like for the real stratosphere, as the wave-1 peak died down during the first half of February, would be extremely hard to guess without the help of a very accurate numerical simulation. By that time, the horizontal mixing would have produced a complicated, sheared-out pattern of small-scale potential-vorticity debris—the phenomenon underlying the so-called "potential enstrophy cascade" (*e.g.* Rhines, 1979)—with some bits and pieces still being pulled round the weakening Aleutian high, and others in the outer part of the displaced polar vortex, which incidentally undergoes considerable fluctuations in shape and position during the period of large wave-1 amplitude. The very important dynamics involved in the whole process seems not always to have been fully appreciated by synopticians—one sometimes hears about the "mere" strengthening of the Aleutian high. Perhaps this has been because of the near-impossibility of drawing isentropic maps of potential vorticity from even the best data analyses and thus seeing directly what is going on. The experimental, high-resolution ECMWF forecasts for late January 1979 would seem to be our best hope at present of being able to say anything quantitative about this, and I am hoping that Dr. Simmons will be able to do something about it before too long!

However, we can guess the qualitative effect on the zonal-mean state easily enough, as already hinted, using the quasi-conservative property of potential vorticity. The ideas date back to the old arguments of G. I. Taylor, C.-G. Rossby and others concerning the mixing of absolute vorticity in barotropic flows, and of course are intimately related, in various ways, to the ideas developed for instance in Dickinson (1969), Davies (1981), Geisler (1974), Green (1970), Holton and Dunkerton (1978), Rhines and Holland (1979), and Rhines and Young (1982). From the synoptic evidence we can expect that most of the mixing was centered on middle latitudes in this case, and therefore that as far as the net effect on the larger scales are concerned, *i.e.* ignoring the small-scale "debris", the isentropic potential-vorticity gradients would have tended on average to be smeared out in middle latitudes, as suggested schematically by the heavy curve in Fig. 5a to which we have already referred. A smearing-out of large-scale mean gradients in middle latitudes implies a sharpening of gradients at the edge of what is left of the polar-night jet, giving rise to a tighter and narrower jet as suggested (again schematically) by the heavy curve in Fig. 5b. As we have already seen, the basic-state

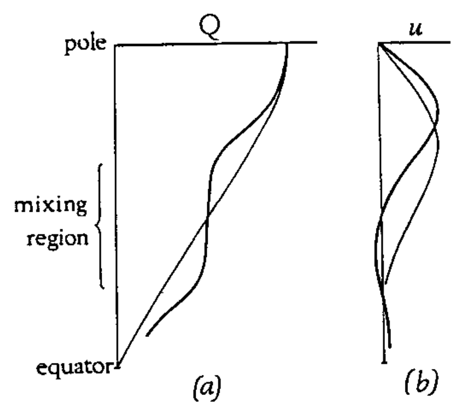

Fig. 5 (a) Schematic latitudinal distributions of Ertel's potential vorticity Q on an isentropic surface in the stratosphere, before and after a large-amplitude planetary wave breaking event centred on middle latitudes. Thin curves are "before", and heavy curves "after". (b) Corresponding polar-night jet profiles u. A broad jet (thin curve) is converted into a narrow jet (heavy curve) with a region of small $\partial Q/\partial y$ to the south of it where the potential vorticity has been most strongly mixed. No attempt is made to suggest the small-scale structure of Q due to "debris" from the wave-breaking event. The profiles may be thought of as representing Eulerian zonal means *after* the wave has largely decayed; see remarks near the end of section 4 about reversible and irreversible changes in the Eulerian-mean state.

configuration suggested by Fig. 5 is precisely the sort of configuration needed to focus an upward-propagating wave-2 pulse, if and when wave 2 decides to amplify subsequently in the troposphere.

Of course an ensuing major warming—the end result of the subsequent pulse, if any, being focused and then breaking *in* the polar cap—can itself be looked at from the viewpoint of potential-vorticity mixing (Davies, 1981). From this viewpoint the difference between a minor warming and a major warming is simply that middle latitudes are mixed in the first case, but the polar cap is mixed in the second.

Presumably the configuration sketched in Fig. 5 would be able to focus wave 1 just as well if not better than wave 2; and a subsequent wave-1 pulse in February 1979 could presumably have led to a major warming just as well as the wave-2 pulse that actually occurred. That is more or less what seems to have happened in the final warming of 1973-74, if the diagram corresponding to Fig. 2 is any guide (Labitzke, 1982, Fig. 1, page 127 in this issue). For that matter, if large wave-1 amplitudes had simply persisted a little longer in the present, January-February 1979 case, there could have been a more or less continuous evolution into a wave-1 dominated major warming. Whether the evolution is continuous or "pulsed" does not seem to be a specially fundamental distinction. There are suggestions in Labitzke's Fig. 1 that wave-1-dominated examples with roughly continuous evolution occurred in February 1980, and in 1969-70. Schoeberl (1978) calls the 1969-70 case a "type B" warming, defining this category by the criterion that wave 1 "maintains a large amplitude for a long period." O'Neill (1980) describes in some detail another example of a strong, wave-1-dominated warming which evolved in a more or less continuous way, taken not from observations of the real atmosphere but from a 13-level general circulation model. See also the nonlinear mechanistic model simulation reported by Hsu (1981).

These remarks have obvious implications for question 8 on my list. The forecasting of a major warming seems certain to depend equally crucially on two separate things. One is an accurate estimate of the initial potential-vorticity gradients in the polar-night jet, together with relevant phase speeds and critical-line positions. The other is an accurate estimate of just how long the anomalous forcing from the troposphere will persist. It is hardly surprising that no simple rule of thumb for predicting major warmings has been found.

A complementary view of the process suggested by Fig. 5 is given by wave, mean-flow interaction theory. The zonal deceleration in middle latitudes suggested by Fig. 5b is precisely what that theory would predict for a not-too-well-focused planetary-wave pulse which saturates somewhere in middle latitudes instead of in the polar cap. Wave 1 can probably do this under a wider range of conditions than wave 2, which may explain why "an intense development of height wave 1" is usually necessary, according to Labitzke (1978), before a major warming can occur.

In an Eulerian-mean description of the effect of a wave pulse like the big wave-1 peak in Fig. 2, an appropriate measure of "where" the waves break or saturate is the convergence of the EP flux integrated over the time of the whole wave event. The time integration gets rid of the purely temporary, reversible changes which may complicate the Eulerian-mean picture from moment to moment as wave amplitudes fluctuate. This ties in with our previous view of Fig. 5 because the convergence of the EP flux is approximately proportional to the isentropic flux of potential vorticity, as is well known (*e.g.* Green, 1970, Eq. 11; Edmon *et al.*, 1980, Eq. 3.5); and the time integration over the wave event picks out the net contribution representing irreversible, downgradient mixing of potential vorticity (cf. Rhines and Holland, 1979; Edmon *et al.*, *op. cit.*, p. 2610; Hoskins, 1982).

The way in which the time integrated EP flux convergence enters the wave, mean-flow interaction theory can be seen directly from eqs. (4.1a, d) of Dunkerton *et al.* (1981) for the rates of change $\partial\bar{u}/\partial t$ and $\partial\bar{\theta}/\partial t$ of the Eulerian-mean zonal velocity $\bar{u}$ and potential temperature $\bar{\theta}$. Those equations are the prognostic members of the set of transformed Eulerian-mean equations presented by Andrews and McIntyre (1976, 1978a; see also Boyd, 1976, eq. below 3.9). If changes in static stability are neglected, the transformed equations for $\partial\bar{u}/\partial t$ and $\partial\bar{\theta}/\partial t$ may be integrated over any given time interval to give equations of the same mathematical form but involving only the net changes $\Delta\bar{u}$ and $\Delta\bar{\theta}$ in $\bar{u}$ and $\bar{\theta}$, instead of the instantaneous rates of change $\partial\bar{u}/\partial t$ and $\partial\bar{\theta}/\partial t$. The time-integrated EP flux convergence now appears in place of the instantaneous convergence.

Wave-mean theory highlights two interesting points about the net mean-flow change $\Delta \bar{u}$ suggested by Fig. 5b. First, the theory predicts not only mid-latitude deceleration where the strongest EP wave flux convergence occurs, but also a general tendency for acceleration to occur north of that location, if there is comparatively little EP convergence there. This point was noted by Palmer (1981b). Second, there is a strong tendency for the main deceleration region to be narrower latitudinally, and deeper vertically, than the region of actual EP wave flux convergence. This phenomenon is nicely illustrated, from observational data for the real atmosphere, by Figs. 5a and 5b of O'Neill and Youngblut (1982). It is one reason why the amount of focusing or defocusing is so important for questions 2 and 8 on my list. As far as the net effect on the $\bar{u}$ profile is concerned, it really does matter at what latitudes the waves saturate. For a given state of the troposphere, the precise degree of wave focusing in the stratosphere could easily make all the difference between getting a major warming such as that of February 1979, and a minor warming such as that of January 1979.

Examples which nicely illustrate these two points have been given by Dunkerton *et al.* (1981) and Hsu (1981), and Fig. 6 recalls one of them. It is taken from a later stage in the same mechanistic simulation that produced Fig. 1a, but still well before the final major warming. The heavy curve marked D_F represents the divergence of the EP flux, in the quasi-geostrophic approximation, rescaled so as to represent the effective Eulerian-mean zonal force per unit mass due to the waves (with dimensions of acceleration). This makes D_F *equal* to the northward flux of quasi-geostrophic potential vorticity (within the approximations usually associated with quasi-geostrophic theory). D_F incorporates the principal Eulerian eddy heat as well as momentum fluxes, so that quasi-geostrophically there is no other wave-induced forcing of the mean state, as described by the transformed Eulerian-mean equations. That is why the response, as measured by the actual mean zonal acceleration $\partial \bar{u}/\partial t$, follows D_F much more closely than either the eddy momentum flux convergence or the Eulerian-mean Coriolis acceleration, shown by the thin curves in the figure. It can be seen, however, that the region of deceleration is narrower in latitudinal extent than D_F itself (and it is, in fact, very much deeper vertically, although the figure does not show this). It can also be seen that positive, westerly acceleration is indeed occurring in high latitudes. This phenomenon was remarked on by Holton (1976) in connection with a similar model simulation, and it can also be seen in Fig. 7b of Hsu (1981), which relates to a different, wave-1-dominated simulation.

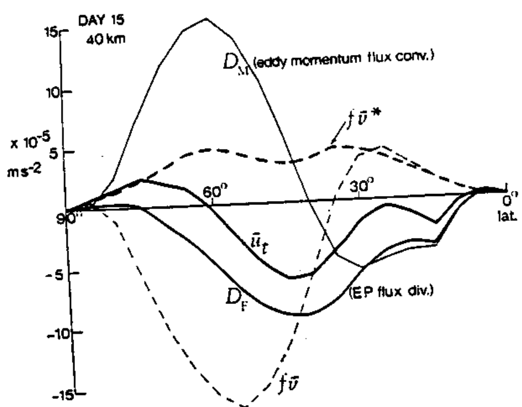

Fig. 6 Terms in equations governing the mean zonal acceleration $\bar{u}_t$, in a situation where the EP wave flux is converging mainly in middle and low latitudes. D_F is the EP flux divergence normalized so as to represent a zonally-directed force per unit mass, and $f\bar{v}^*$ is the Coriolis force due to the residual meridional circulation, which appears in the transformed Eulerian-mean equations. The light curves show the principal terms in the conventional mean momentum equation for comparison. From Dunkerton *et al.* (1981).

Both the high-latitude acceleration, and the narrowing and deepening of the region of deceleration, are immediate consequences of the general way in which a balanced zonal flow responds to a given zonal force D_F per unit mass concentrated at a given height and latitude, a classical problem studied by Eliassen (1951) and discussed in this context by Dunkerton *et al.* (*op. cit.*, §4), by Palmer (1981b, §5), and by O'Neill and Youngblut (1982). In Fig. 7, also taken from Dunkerton *et al.*, the shaded region shows the height and latitude where D_F is greatest in our example. Eliassen's theory tells us that the response will include a meridional circulation whose Coriolis force redistributes the effect of the force D_F, and whose vertical advection re-orients the isentropic surfaces, in such a way as to preserve thermal-wind balance. In the transformed Eulerian-mean formalism, the relevant meridional circulation is what Andrews and McIntyre called the "residual circulation" ($\bar{v}^*$,

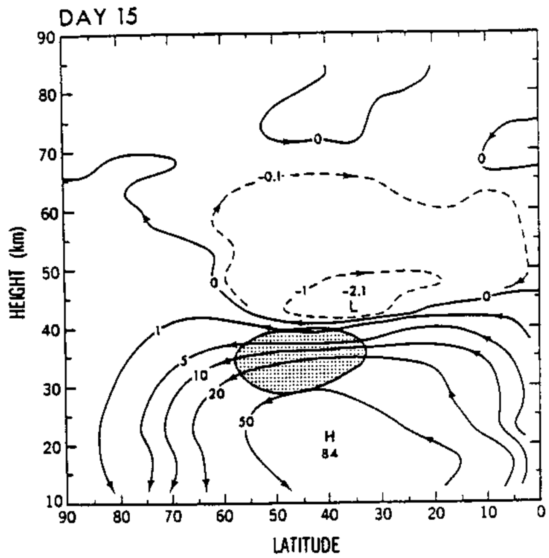

Fig. 7 Region of maximum D_F (shown shaded), for the same situation as in Fig. 6, and mass streamfunction $\bar{\chi}^*$ for the residual Eulerian-mean circulation ($\bar{v}^*$, $\bar{w}^*$). The shaded region represents values of $-D_F$ exceeding $10^{-4}\mathrm{m\,s^{-2}}$. $\bar{\chi}^*$ is defined such that $\partial\bar{\chi}^*/\partial\phi = a\rho_0\bar{w}^*\cos\phi$ and $\partial\bar{\chi}^*/\partial z = -\rho_0\bar{v}^*\cos\phi$, where ϕ is latitude, a the radius of the earth, $\rho_0(z)$ a standard density, and z is 7 km times $-\ln$ (pressure/1000 mb). Contour values are to be multiplied by 1.46 $\mathrm{kg\,m^{-1}s^{-1}}$. From Dunkerton *et al.* (1981).

$\bar{w}^*$). The residual circulation induced by D_F in the present example is also shown in Fig. 7. It extends over many scale heights and has the simple, two-cell structure predicted by Eliassen's theory. The lower cell is acting to tilt the isentropic surfaces anticlockwise in the picture, and the upper cell clockwise. The northward flow through the region where D_F is concentrated extends into the polar cap, where the associated Coriolis acceleration ($f\bar{v}^*$ in Fig. 6) causes the positive, high-latitude zonal acceleration already noted. The Coriolis forces in the two branches of the return flow effectively extend the region of *deceleration* high up into the mesosphere and, apparently, well down into the troposphere, so far as we can tell from a mechanistic model with an artificial lower boundary condition. I should emphasise that none of these general features of the response depend in any way whatever upon linear wave theory. Eliassen's theory applied to the transformed Eulerian-mean equations shows that the response will have the character just described whenever the EP wave flux converges, for whatever reason, onto a sufficiently well localized region in middle latitudes in the manner suggested by the shaded region in Fig. 7.

This picture seems to apply quite well to the 1979 wave-1 episode, provided that we are careful to interpret it in the time-integrated sense. For quantitative purposes the effects of diabatic cooling would have to be added. The relevant period of time appears to be mid-January to mid-February. Fig. 4c of Labitzke (1981) gives a latitude-time section of the mean zonal wind $\bar{u}$ at 10 mb, and Fig. 5 of the same paper gives three meridional cross-sections of $\bar{u}$ during that period. Those figures confirm that $\bar{u}$ did undergo a net change $\Delta\bar{u}$, over the whole time interval, very like that suggested by our Fig. 5b, involving deceleration in middle latitudes and acceleration in high latitudes. But they show also that it happened in at least two distinct stages. (Fig. 3 in the same paper suggests that even that may be a considerable oversimplification, when transient events at higher altitudes are taken into account.) At 10 mb, Labitzke's Fig. 4c shows a net Eulerian-mean deceleration over the first few days, from about 18 to 25 January, in a fairly broad region spanning middle and high latitudes. Much of this Eulerian-mean deceleration is attributable to wave transience associated with the rapid growth of wave-1 amplitude during that period. Viewed synoptically, part of the effect of wave-1 "transience" on the Eulerian mean is just the kinematical effect of displacing the main, cyclonic polar vortex out of line with the latitude circles around which the mean is taken, as is illustrated by our Fig. 3.

Most of the high-latitude acceleration appears to have taken place, somewhat erratically, during a second stage which occupied the period from about 26 January to mid-February. Weak deceleration continued, on average, in middle latitudes during most of that period. From Labitzke's figures and the daily synoptic maps it can be anticipated that the EP wave flux would have been diverging from high latitudes and converging (weakly) into the main region of wave-breaking to the south, when time-averaged over this second stage. Some preliminary EP cross-sections constructed by Dr. Palmer appear to confirm this, in addition showing considerable day-to-day fluctuations as expected.

The EP divergence in high latitudes, time-averaged over the second stage from 26 January to mid-February, was probably due mainly to transience in the opposite sense from before, associated with a net reversible decay of the local wave amplitude as planetary-wave activity propa-