The 1960s—A decade of remarkable advances in middle atmosphere research

Marvin A. Geller

School of Marine and Atmospheric Sciences,
Stony Brook, New York 11794-5000, USA
E-mail: marvin.geller@sunysb.edu

The 1960s was a decade of great advances in middle atmosphere research. In this paper, I briefly discuss some of the seminal papers and events that greatly influenced middle atmosphere research and continue to do so to this day. This paper also gives a primer on many of the basics of middle atmosphere wave-mean flow interaction.

1 Introduction

The 1960s was a decade of great social change and scientific advances. The ends of World War II and the Great Depression were more than a decade in the past. The Cold War was ongoing, and the conflicts in Southeast Asia intensified during that decade. Technology was progressing rapidly, with the applications of the World War II technologies of rockets and radar as well as newly developed satellite observations of our environment. Computers were becoming common research tools. Governments appreciated the important role of science and technology in World War II and were increasing their expenditures in these areas, particularly given Cold War competition. It was also a period where young people were enticed by science and technology, and were entering these fields in great numbers, often with government support for their education.

2 1960s Middle Atmosphere Advances

The decade of the 1960s saw many important advances in observations, theory, and numerical modeling of the middle atmosphere. Different people will come up with different lists of the seminal papers that appeared during this period. Table 1 shows my selection. Of course, there is some arbitrariness here. One is that papers published 1959, or before, and 1970, or after, are not discussed here. Also, of course, I have made the selection, and therefore, some important papers are not mentioned that might have been chosen by others, so no doubt some very important papers have not made my list. I would argue, however, that all of the papers in Table 1 are seminal!

2.1 Hines (1960)

Hines’ (1960) paper was a tour de force. He presented observational evidence for gravity waves in the atmosphere. He developed the linear theory for these waves.
<table>
<thead>
<tr>
<th>Paper(s) or Event</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hines (1960)</td>
<td>Developed the theory for acoustic-gravity waves.</td>
</tr>
<tr>
<td>Reed et al. (1961) and Veryard and Ebdon (1961)</td>
<td>Discovery of the QBO.</td>
</tr>
<tr>
<td>Charney and Drazin (1961)</td>
<td>First theory for planetary wave propagation.</td>
</tr>
<tr>
<td>Hampson (1964) and Hunt (1966)</td>
<td>First suggestion of catalytic loss for stratospheric ozone.</td>
</tr>
<tr>
<td>Leovy (1964)</td>
<td>First successful model for mesospheric structure.</td>
</tr>
<tr>
<td>Reed (1965)</td>
<td>Discovery of the SAO.</td>
</tr>
<tr>
<td>Matsuno (1966)</td>
<td>Derivation of the structure of equatorially trapped Kelvin and mixed Rossby-gravity waves.</td>
</tr>
<tr>
<td>Yanai and Maruyama (1966) and Wallace and Kosmrl (1968)</td>
<td>Observational identification of equatorially trapped Kelvin and mixed Rossby-gravity waves.</td>
</tr>
<tr>
<td>Booker and Bretherton (1967)</td>
<td>Theory for gravity wave critical levels.</td>
</tr>
<tr>
<td>Hodges (1967)</td>
<td>Beginnings of the theory for gravity wave breaking.</td>
</tr>
<tr>
<td>Manabe and Hunt (1968) and Hunt and Manabe (1968)</td>
<td>First troposphere-stratosphere general circulation model.</td>
</tr>
<tr>
<td>Wallace and Holton (1968), Lindzen and Holton (1968)</td>
<td>First successful theory for the QBO.</td>
</tr>
</tbody>
</table>

He discussed some of their effects, and he predicted which waves could be observed in ionospheric regions. Hines (1960) showed that there were two distinct classes of waves in a compressible, gravitationally stratified atmosphere: internal gravity waves with frequencies less than the Brunt-Vaisälä frequency and acoustic-gravity waves with frequencies higher than the acoustic cut-off frequency. Further, he showed that the internal gravity waves had the asymptotic behavior of internal gravity waves in incompressible fluids for low frequencies and the acoustic-gravity waves had the asymptotic behavior of sound waves for frequencies much higher than the acoustic cut-off frequency.

One of the most fundamental results of the Hines (1960) paper was that the vertical component of an internal gravity wave’s phase velocity is opposite to the vertical component of the internal gravity wave’s group velocity, the speed at which the gravity wave energy propagates. This is shown in Fig. 1 here, which is figure 2 in Hines (1960).

2.2 Construction of the Jicamarca radar

The Jicamarca Radio Observatory was originally constructed in Peru by the Central Radio Propagation Laboratory (CRPL), which was then part of the United States
Fig. 1. Pictorial representation of internal gravity waves. Instantaneous velocity vectors are shown, as are their instantaneous and overall envelopes. Density variations are depicted by a background lying in surfaces of constant phase. The vertical component of the phase velocity is downward while energy is being propagated upward. Note that gravity is directed vertically downward. (From Hines (1960).)

Fig. 2. Backscatter power profile from the Jicamarca radar as reported by Woodman and Guillen (1974).
Bureau of Standards. The CRPL later was absorbed into ESSA, the Environmental Science Service Administration, which ultimately became NOAA, the National Oceanic and Atmospheric Administration. Since 1979, the Jicamarca Radio Observatory has received financial support from a US National Science Foundation grant through Cornell University. Jicamarca is a unique research facility in many ways. It is essentially located on the magnetic equator (1° dip angle). It has the largest antenna of any atmospheric radar in the world (18,432 half-wave dipoles arranged over an area of a little less than 85,000 square meters). Its principal use was envisaged to be, and still is, incoherent scatter measurement of the equatorial ionosphere, but Woodman and Guillen (1974) showed that the Jicamarca radar could be used to obtain neutral winds in the stratosphere and mesosphere. Figure 2, from Woodman and Guillen (1974), shows the backscatter power profile as a function of altitude from a single pulse and from a 316 pulse average. Thus, with suitable averaging, the Jicamarca radar can measure winds throughout the middle atmosphere.

The construction of the Jicamarca radar in 1960–61 paved the way for the construction of MST radars worldwide, so at the present time, there are one or more MST radars operating on every continent except for Africa and Antarctica, and the Japanese have plans for constructing an MST radar in Antarctica. Measurements from MST radars have yielded a great deal of information on middle atmosphere dynamics, especially on gravity waves where their ability to sample frequently and with excellent altitude resolution is unique.

2.3 Discovery of the quasi-biennial oscillation (QBO)

The QBO was discovered independently by Reed et al. (1961) and by Veryard and Ebdon (1961). Figure 3 shows its structure over the Equator. Note a number of features seen in Fig. 3. Most evident is a quasi-periodic pattern of descending easterlies (unshaded) followed by descending westerlies (shaded). Looking more closely, one sees that the average period of a complete cycle is about 28 months, but that the period varies considerably, being about 21 months in 1972–1974 and about 35 months in 1983–1986. Moreover, the westerlies descend more quickly than the easterlies. The maximum amplitude of the QBO is about 20 m s$^{-1}$, occurring in the middle stratosphere.

Following the discovery of the QBO, there were many attempts to explain why this phenomenon occurred, but it was not until the late 1960s that a successful explanation was given. This will be discussed later.

2.4 Murgatroyd and Singleton (1961)

In a pioneering publication, Murgatroyd and Singleton (1961) attempted to calculate the middle atmosphere mean meridional circulation by using observations of temperature, calculating the net diabatic heating rates, and assuming that the dynamical heating resulting from the mean meridional circulation balances the radiative terms. This approach ignores the eddy heat and momentum transport terms, so gives an incorrect picture for the Eulerian mean meridional circulation. Dunkerton (1978) realized, however, that the Murgatroyd and Singleton (1961) circulation gave what we now call the “diabatic circulation,” which is a good first approximation to the Lagrangian mean meridional circulation that correctly portrays mass transports. It
Fig. 3. Time-height section of the monthly mean zonal winds (in m s$^{-1}$) over equatorial stations. (From Geller et al. (1997), which is an update from Naujokat (1986).)
Fig. 4. (Left)—The middle atmosphere diabatic circulation from Dunkerton (1978). This is essentially the circulation computed by Murgatroyd and Singleton (1961) for Northern Hemisphere winter solstice. (Right)—The January 1964 Northern Hemisphere, lower stratosphere, Eulerian circulation from Vincent (1968). Note that very different height ranges are shown in the two panels, and the winter pole is on the right in the left panel but on the left in the right panel.

is interesting to compare the discussions on pages 12 and 14 in Holton (1975) with that on pages 303–305 in Andrews et al. (1987) to see the evolution in time of our understanding on this subject.

Figure 4 shows a comparison between the middle atmosphere diabatic circulation from Dunkerton (1978), essentially the Murgatroyd and Singleton (1961) calculation, and the lower stratosphere Eulerian circulation, as computed by Vincent (1968). Both represent January conditions. Comparing the circulations in the lower stratosphere, one sees ascending air through the tropical tropopause and descending air at high latitudes in the left panel, but the right panel shows ascending air both in the tropics and at polar latitudes with descending air at middle latitudes. Thus, the left panel is consistent with the Brewer-Dobson circulation that is required to explain observed stratospheric water vapor and ozone distributions, while the right panel is not. This is reasonable since we now know that the left panel represents the Lagrangian circulation that transports constituents while the right panel does not represent this “transport circulation.” Why this is so will become clearer in the following discussion of the “non-interaction” and “non-transport” theorems.

2.5 Charney and Drazin (1961)

During the 1960s, observations indicated that the summer hemisphere polar vortex showed little longitudinal variation while the winter hemisphere polar vortex was very asymmetric (see Hare, 1968, for example). This was explained in the classic paper of Charney and Drazin (1961).

They used the linear, quasi-geostrophic equations on a $\beta$-plane to derive the following equation for the vertical structure of planetary wave disturbances.

$$\frac{d^2 V}{dz^2} - \frac{1}{H} \frac{dV}{dz} - \frac{\beta N^2 (u_0 - c - \nu_c)}{f_0^2 u_c (u_0 - c)} V = 0$$  \hspace{1cm} (1)
where $u_c \equiv \frac{\beta}{k^2 + l^2}$, $z$ being the altitude, $H$ the pressure scale height, $\beta$ the latitudinal derivative of the Coriolis parameter, $u_0$ the mean zonal flow, $c$ the planetary wave zonal phase velocity, $f_0$ the Coriolis parameter, $V$ the vertical variation of the planetary wave northward velocity, $k$ and $l$ the zonal and meridional planetary wave wavenumbers, respectively, and $N$ the Brunt Väisälä frequency. The solution to Eq. (1) is

$$V = (Ae^{inz} + Be^{-inz}) e^{\frac{z}{H}},$$

where

$$n^2 = -\frac{1}{4H^2} - \frac{N^2}{f_0^2} \left( (k^2 + l^2) - \frac{\beta}{u_0 - c} \right).$$

(2)

Now, Eq. (2) implies that for vertical planetary wave propagation, $n$ must be real, which implies $n^2$ must be greater than zero, which in turn implies that

$$0 < u_0 - c < U_c,$$

where

$$U_c = \frac{\beta}{k^2 + l^2} + \frac{f_0^2}{4H^2N}.$$}

(3)

Condition (3) implies that for stationary planetary waves ($c = 0$), the mean zonal flow must be westerly (or eastward) and less than a critical wind speed $U_c$. This immediately accounts for the undisturbed summer polar vortex since the mean zonal winds are easterly (westward) during summer. It also accounts for the fact that even in winter, the asymmetries are large scale relative to the situation in the troposphere since $k$ and $l$ are larger for the smaller scale waves, implying $U_c$ is smaller. One difficulty though is that even when the mean zonal winds are quite large, stationary planetary waves are seen to be present, but improvements in the theory to address this were done by Dickinson (1968) and Matsuno (1970).

Finally, Charney and Drazin (1961) show that for steady waves, in the absence of dissipation, the planetary wave heat and momentum fluxes are so configured that there is no interaction between the vertically propagating planetary waves and the mean zonal flow. This, so-called non-interaction theorem is discussed next.

2.6 Eliassen and Palm (1961)

While Charney and Drazin (1961) treated the non-interaction theorem for planetary waves, Eliassen and Palm (1961) focus on the non-interaction theorem more generally. A more general form of their equation (2.13), for the case of a steady gravity wave propagating in a shear flow in the absence of diabatic effects, is

$$\rho'w' = -\rho_0(u_0 - c)ww',$$

(4)

where $p$, $u$ and $w$ are pressure and horizontal and vertical velocities, the overlines denote averaging over wave phase, and the primes indicate the wave perturbations. Equation (4) is sometimes referred to as Eliassen and Palm’s 1st Theorem. It implies that for upward wave energy flux ($\rho'w' > 0$), the wave momentum flux ($\rho_0 uu'$) is negative when the mean flow, $u_0$, is greater than the wave phase velocity $c$, and is positive when $u_0 < c$. Thus, any physical process that leads to a decrease of the wave amplitude as it propagates (e.g., dissipation) will force the mean flow toward the wave phase velocity.
For gravity waves with phase velocity \( c \), Eliassen and Palm’s 2nd Theorem can be written as

\[
\rho_0 (u_0 - c) \bar{u}'w' = \text{constant} \tag{5}
\]

in the case of no wave transience, no diabatic effects, and \( u_0 - c \neq 0 \). Thus, in this case, there is no gravity wave interaction with the mean flow. This is consistent with the non-interaction planetary wave result of Charney and Drazin (1961), which was also obtained by Eliassen and Palm (1961).

The implications of Eliassen and Palm’s 1st and 2nd theorems are far-reaching. These indicate that unless there are dissipation, other diabatic effects, wave transience, or \( u_0 = c \), atmospheric waves do not interact with the main flow. Conversely, if any of these are present, the waves do interact with the mean flow, and this interaction gives rise to a deceleration or acceleration of the mean flow toward the wave’s phase velocity. These interactions are fundamental in understanding why the quasi-biennial and semi-annual oscillations occur in the equatorial middle atmosphere and what gives rise to stratospheric sudden warmings.

2.7 Leovy (1964)

By the 1960s, it was known that summer mesopause temperatures were much colder than winter mesopause temperatures, which was opposite to the influence of solar heating—the summer mesopause being constantly sunlit while the winter mesopause is subject to continuous darkness. Leovy (1964) constructed the first sim-
ple model of the zonally averaged mesospheric circulation. He carried out numerical experiments where the middle atmosphere circulation was forced by realistic solar differential heating. He used two different formulations of mechanical damping, one being Rayleigh friction and the other being momentum diffusion by viscosity. Figure 5 shows the zonal wind simulation for the Rayleigh friction case. Leovy (1964) also showed results for viscosity and thermal conduction. He was able to get reasonable results, but only with viscosities and thermal conduction coefficients that were excessively large.

Leovy’s (1964) calculations were followed up much later by Schoeberl and Strobel (1978) and by Holton and Wehrbein (1980), at which time there was a greater appreciation that Rayleigh drag was likely a crude representation of gravity wave effects. Research into proper representation of gravity wave effects on mean flows then became a popular subject for research following the pioneering works of Matsuno (1982) and Lindzen (1981).

2.8 Early research on stratospheric ozone loss by catalytic reactions

Chapman (1930) suggested the following set of chemical reactions might explain the stratospheric ozone layer.

\[
\begin{align*}
    O_2 + h\nu(\lambda < 240 \text{ nm}) & \rightarrow O + O \\
    O + O_2 + M & \rightarrow O_3 + M \\
    O_3 + h\nu(\lambda < 320 \text{ nm}) & \rightarrow O + O_2 \\
    O + O_3 & \rightarrow 2O_2
\end{align*}
\]

Thus, ultraviolet radiation dissociates molecular ozone in (6); ozone is formed in the presence of a third body, M, in (7); ozone is photochemically dissociated in (8), but the atomic oxygen quickly transforms back into ozone by (7); and the ozone is lost by (9).

Benson and Axworthy’s (1957) measurements of reaction (9), however, indicated that reaction proceeded much more slowly than had been assumed, implying that reactions (6)–(9) give excessive stratospheric ozone. Hampson (1964) suggested that the following catalytic chemical reactions might resolve the issue.

\[
\begin{align*}
    OH + O_3 & \rightarrow HO_2 + O_2 \\
    HO_2 + O_3 & \rightarrow OH + 2O_2
\end{align*}
\]

Reactions (10) and (11) constitute a set of reactions for a catalytic cycle for stratospheric ozone loss, since OH initiates the cycle and remains after the cycle is completed. Hunt (1966) suggested a set of reaction rates for (10) and (11) that was consistent with observed values for stratospheric ozone. In fact, these reactions were not found to be very important in the stratosphere, but are important in the mesosphere.

Crutzen (1970) and Johnston (1971) showed that the nitrogen catalytic cycles are very important in the natural stratosphere, and Molina and Rowland (1974) pointed
out that the chlorine resulting from the photodissociation of the industrially pro-
duced chlorofluoromethanes would likely destroy significant amounts of stratospheric ozone. Of course, Paul Crutzen, Sherwood Rowland, and Mario Molina went on to receive the 1995 Nobel Prize in Chemistry for their contributions, but I believe it is fair to say that this line of investigation into catalytic destruction of stratospheric ozone had its genesis during the 1960s.

2.9 Discovery of the semi-annual oscillation by Reed (1965)

Reed (1965) used a new two-year time series of wind measurements over Ascension Island (70°55′S, 14°25′W) to examine how the quasi-biennial oscillation evolved with increasing altitude between 30 and 50 km. Figure 6, from his paper, shows his results.

Note that the altitude scale for Fig. 6 essentially starts at the top of Fig. 3. Looking at Fig. 6, we see a gradual transition from a long period oscillation at 28 km, to shorter-term variability at 40 km, to a clear 6-month oscillation at 52 km. Thus, while the equatorial lower stratosphere is characterized by the quasi-biennial oscillation, the upper stratosphere—lower mesosphere is characterized by the semi-annual oscillation.

Fig. 6. Two-year time series for measured zonal wind over Ascension Island at 28, 32, 36, 40, 44, 48, and 52 km. Circles indicate measurements, and curves are harmonic curve fits. (From Reed (1965).)
2.10 Discovery of missing tidal modes by Kato (1966) and Lindzen (1966)

The theory for atmospheric tides goes back to the work of Laplace (1799), who derived the idealized equations for the free and forced oscillations for a thin atmosphere on a spherical planet. While the Moon’s gravitation forces the oceanic tide, which is semidiurnal, it is the Sun’s heating that forces the Earth’s atmospheric tides. Under simplifying assumptions, Laplace used the traditional method of separation of variables to obtain an ordinary differential equation for the latitudinal structure of the tide as well as one for its vertical structure. The equation for the latitude structure is called Laplace’s tidal equation. Its eigenvalues are often referred to in terms of equivalent depths (an analogy to the ocean), and its eigenfunctions are called Hough functions. The equivalent depth then occurs as a parameter in the vertical structure equation. One can then expand the latitudinal structure of the solar heating in terms of these Hough functions.

Early observations contradicted simple intuition, in that the surface pressure showed tidal variations that were predominantly semidiurnal, while the solar forcing is predominantly diurnal. Figure 7, from Chapman and Lindzen (1970), illustrates two things. One is that surface pressure variations in connection with extratropical systems are much greater than surface pressure variations in the tropics, except when tropical cyclones occur. The other is that tidal variations in the tropics are larger in the tropics and are predominantly semidiurnal rather than diurnal, as expected.

Initial suggestions to explain the dominance of the semidiurnal tide in surface pressure over the expected diurnal tide were that the semidiurnal solar forcing excited a resonance in the atmosphere while the diurnal forcing was off resonance. As more and more was learned about atmospheric tides and their forcing, this did not prove to be the case. It was not until Kato (1966) and Lindzen (1966) independently discovered the existence of negative equivalent depth eigenmodes to Laplace’s tidal equation that this problem was solved. The solution is best explained with the aid Fig. 7. Barometric variations (on two different scales) at Batavia (the present Jakarta at 6°S) and Potsdam (52°N) in November 1919. This figure is from Chapman and Lindzen (1970), who in turn took it from Bartels (1928).
of Fig. 8, from Chapman and Lindzen (1970). In this figure, the two profiles represent the vertical variation of the solar heating of atmospheric water vapor (V1) and ozone (V2), along with the latitudinal variations of these heating functions. Note the different scales for the diurnal and semidiurnal heating functions. When these heating functions are expanded in the diurnal and semidiurnal Hough function solutions of Laplace’s tidal equation, it turns out that the semidiurnal heating only gives rise to positive equivalent depth modes, the principal one of which has a very long vertical wavelength of about 100 km, while the diurnal heating gives rise to negative equivalent depth modes, which cannot propagate vertically and principally a positive equivalent depth mode of rather short vertical wavelength (about 25 km). The ozone heating principally forces a negative equivalent depth mode that cannot propagate down to the surface plus a short vertical wavelength mode that does reach the surface. In contrast, the semidiurnal ozone heating principally forces a long wavelength mode that does reach the surface. For the semidiurnal solar tide, the vertical wavelength is longer than the depth of the ozone heating and it propagates down to the surface, so the tidal pressure variations forced by ozone and water vapor add to produce a sizable response in surface pressure. For the diurnal ozone heating, not only does the negative equivalent depth mode not propagate down to the surface, but there is also destructive interference over the depth of the ozone heating for the short vertical wavelength solution at low latitudes. Therefore, the sizable semidiurnal variation in surface pressure is a consequence of the additive solutions from water vapor.
The 1960s

and ozone heating while for the diurnal tide, there is very little surface response to the ozone heating.

Not only did Lindzen’s (1966) and Kato’s (1966) papers lead to a resolution of the long-standing questions about the solar atmospheric tides, it also led to an appreciation of the completeness of the Hough functions. This, in turn, led to Longuet-Higgins (1968) paper that examined the family of oscillations on a fluid envelope on a rotating sphere, which gave a theoretical framework for understanding global atmospheric wave motions.

2.11 Matsuno’s (1966) discovery of equatorial wave modes

Matsuno (1966) examined the quasi-geostrophic wave modes that exist on an equatorial $\beta$-plane. His lowest-order wave modes correspond to what we now refer to as the equatorially trapped Kelvin and mixed Rossby-gravity waves. The structures of these waves are shown in Fig. 9. Note that both the Kelvin and mixed Rossby-gravity waves become geostrophic in that the winds are parallel to the geopotential contours off the Equator. Also, note that the Kelvin wave is symmetric around the Equator while the mixed Rossby-gravity wave is asymmetric about the Equator. Not so obvious from this figure is the fact that the Kelvin wave propagates to the east ($c > 0$) and the mixed Rossby-gravity wave propagates to the west ($c < 0$).

Fig. 9. Schematic illustration of the geopotential and wind fields for the equatorial trapped Kelvin (top) and mixed Rossby-gravity (bottom) waves. (Adapted from Middle Atmosphere Dynamics, Andrews et al., Copyright 1987 Academic Press, with permission from Elsevier. Andrews et al. (1987) in turn adapted theirs from Matsuno (1966).)
Not so long after Matsuno’s discovery of these wave modes, they were identified observationally in the stratosphere. Yanai and Maruyama (1966) identified the mixed Rossby-gravity wave and Wallace and Kousky (1968) identified the Kelvin wave. Table 2, from Andrews et al. (1987) shows the observed characteristics of these waves in the equatorial lower stratosphere. Note that these waves are of planetary scale—wave number 1–2 for the Kelvin wave and 4 for the mixed Rossby-gravity wave. Note also that they are trapped within a couple of thousand km of the Equator.

2.12 Gravity wave critical level theory by Bretherton (1966) and Booker and Bretherton (1967)

We saw in the earlier discussion of the Eliassen and Palm (1961) paper that one of the conditions for non-interaction between gravity waves and the mean flow is that the wave phase velocity is not equal to the mean zonal flow, or \( u_0 \neq c \). Bretherton (1966) examined the case of a gravity wave in a shear flow where \( u_0 = c \) (the critical level) and the Richardson number is very large. He found that in this case the gravity wave vertical group velocity \( \to 0 \) as \( u_0 \to c \). Thus, the gravity wave energy flux \( \rho \bar{u}'w' \) vanished on the far side of the critical level, in which case the momentum flux \( \rho_0 u'w' \) was also zero. Since there was no wave interaction below the critical level, this implies a convergence (or divergence) of the wave momentum flux at the critical level.
Booker and Bretherton (1967) generalized this result to the case of finite Richardson number, in which case they derived the results that in passing through the critical level, the wave momentum flux is attenuated by a factor of $e^{-2\pi \sqrt{Ri}^{-1}}$. Thus, at a gravity wave critical level, the absorption of the wave will tend to bring the mean flow toward the wave phase velocity.

There followed a period of very active research into the nature of gravity wave critical levels. Hazel (1967) showed that the Booker and Bretherton (1967) result was essentially correct in the case of a fluid with viscosity and heat conduction. Breeding (1971) suggested that nonlinear effects might lead to some wave reflection in addition to absorption, and Geller et al. (1975) suggested that as the wave approached a critical level, it produces turbulence that would likely lead to wave absorption before nonlinear effects would lead to wave reflection.

2.13 Hodges (1967) and gravity wave breaking

In an isothermal atmosphere, the density decreases exponentially with increasing altitude $z$ as $e^{-z/H}$, $H$ being the pressure scale height. Without dissipation or critical levels, the gravity wave kinetic energy per unit volume $\rho_0 \bar{u}^2$ should remain constant, in which case the amplitude of the wave’s horizontal velocity (and as it turns out temperature) fluctuations should grow as $e^{+z/B}$. This being the case, the wave eventually becomes unstable. Hodges (1967) was the first to point out that this will be a source of turbulence in the middle and upper atmosphere.

This provided the starting point for Lindzen’s (1981) seminal paper that suggested a self-consistent way of parameterizing the effects of unresolved gravity waves in climate models. The principle for this parameterization is illustrated in Fig. 10. On the right is illustrated a gravity wave whose wind and temperature amplitudes are exponentially increasing with height, and as pictured, the wave momentum flux $\rho_0 \bar{u}'w'$ is...
constant with height. Since the vertical wavelength of this wave is fixed \( \frac{\partial \bar{v}}{\partial z} \) and \( \frac{\partial T}{\partial z} \) also increase with height exponentially, as illustrated by the outer envelope, eventually, the wave becomes either convectively unstable or shear unstable and breaks down. Lindzen (1981) made the assumption that above the level where the wave breaks down, it loses just enough energy to turbulence to keep the wave amplitude constant above that level, as illustrated. This means that \( \rho_0 u' w' \) decreases with height above the breaking level so that there is a divergence of wave momentum flux above the level where the gravity wave begins to break, also as pictured in Fig. 10. Of course, one could make different assumptions of what occurs above the breaking level. For instance, Alexander and Dunkerton (1999) assume that gravity waves deposit all of their momentum at the breaking level.

Developing and implementing ways of parameterizing the effects of unresolved gravity waves in climate models is a research topic of great current interest, but one might say that this had its intellectual roots in the papers of Eliassen and Palm (1961),
Hodges (1967), and Booker and Bretherton (1967).

2.14 The beginnings of troposphere-stratosphere general circulation models

Manabe and Hunt (1968) and Hunt and Manabe (1968) published the results of the first troposphere-stratosphere general circulation models. These papers used a model that contained a full treatment (for that time) of tropospheric physics and also had transport tracers. Figure 11 shows the results of their simulations of the zonally averaged temperature and zonal wind for January for their 18-level model that had
its top at 37.5 km. The agreement between the model temperatures and winds and observations are pretty good. Of course, there are noticeable differences between the model results and observations, some of which are the modeled tropopause jet is too strong and too far poleward; the modeled tropopause temperatures are too cold; and the polar stratosphere is too cold, which leads to the lower portion of the polar night jet being too strong in comparison with observations.

Hunt and Manabe (1968) conducted some tracer experiments with this model. One of their results is shown in Fig. 12. These results are for a passive ozone-like tracer that has its source in the tropical upper stratosphere. Note the near cancellation of the flux divergence terms from the mean meridional circulation and the large-scale eddies. This suggests that, in an analogous fashion to the non-interaction theorem, there is also a non-transport theorem; that is to say, that for steady state conditions, no dissipation, no critical levels, and for a conserved tracer, there is no net transport by the planetary-scale eddies since these eddies give rise to canceling transport effects by the mean meridional circulation.

2.15 Theory for the QBO

After the discovery of the quasi-biennial oscillation, there followed a number of suggestions for its cause. The first successful explanation for the QBO was given by Lindzen and Holton (1968). This followed a very important precursor paper by Wallace and Holton (1968), who showed that one could not explain the QBO by periodically varying heating terms. Rather, they showed that the forcings for the QBO must be periodically varying terms in the momentum equation. Lindzen and Holton (1968) first advanced the theory that, with few modifications, is accepted as the explanation of the QBO today. Basically, they proposed that waves carrying eastward and westward momentum upward were forced in the tropics and their interaction with the mean flow gave rise to alternating momentum source terms for zonal mean momentum, leading to the QBO. This theory was altered a bit in Holton and Lindzen (1972) when dissipation replaced critical level interaction as the mechanism for the wave mean flow momentum interactions.

This theory is best explained by the schematic shown in Fig. 13, which is taken from Plumb (1984). If there exists two vertically propagating gravity waves, one with phase velocity $+c$ and the other with phase velocity $-c$, and if the flow is slightly biased to the eastward direction, the group velocity for the $+c$-wave is diminished relative to that of the $-c$-wave. Therefore, in the presence of any dissipation, there will be preferential momentum absorption of the $+c$-wave. Eliassen and Palm’s 1st theorem implies acceleration of the mean flow toward $+c$. Profile (2) shows that the absorption of the $+c$-wave momentum flux leads to descending eastward flow. Since the $-c$-wave is not being absorbed at lower altitudes, it propagates freely to higher levels where it is also subject to dissipation, thus leading to higher level westward flow in profile (2). Profile (3) shows a later stage where the eastward flow has descended further, and is close to $+c$, while the upper level westward flow is tending toward $-c$ and is also descending. In this idealized situation, the eastward flow descends toward the forcing level, and the shears become so sharp that diffusion brings the flow toward zero, but at this point, there is a descending westward flow, which is
Fig. 13. Schematic representation of the Lindzen and Holton (1968)/Holton and Linden (1972) theory for the QBO. Panel (a) shows the initial state, which corresponds to profile (1) in panel (b). Profiles (2) and (3) show successive stages of evolution, as explained in the text. (After Plumb (1984).)

similar to the mirror image of panel (a), i.e., a westward bias to the flow, and the mirror image of the cycle proceeds, producing one complete cycle of alternating descent of eastward flow followed by a descent of westward flow, and so on.

Interestingly enough, the waves that carry eastward and westward momentum need not be equatorial waves, rather it is the very small Coriolis force near the Equator that makes the wave momentum fluxes influence \( \frac{\partial \bar{u}}{\partial t} \), rather than forcing a merdional flow. This has been elegantly shown by Haynes (1998).

2.16 Dickinson’s (1969) paper on planetary wave-mean flow interactions

There have been several generalizations of the Eliassen-Palm (1961) and Charney-Drazin (1961) noninteraction theorems. Most of these have occurred outside of our window of the 1960s (e.g., Holton, 1974; Andrews and McIntyre, 1976; Boyd, 1978), but one such generalization that did occur during the 1960s was by Dickinson (1969). He used a quasi-geostrophic formulation for small amplitude planetary waves on a \( \beta \)-plane geometry to demonstrate that there must exist eddy transport of quasi-geostrophic potential vorticity for the waves to cause an acceleration of the mean zonal flow, and that such potential vorticity transports will occur only in the presence of critical lines (where the wave phase speed equals the mean flow) or in the presence of dissipation.

It turns out that this condition for non-zero quasi-geostrophic potential vorticity transports is quite equivalent to the modern-day Eliassen-Palm flux divergence. For example, Andrews et al. (1987) show on pages 129–130 that

\[
\overline{v'q'} = \frac{1}{\rho_0} \vec{\nabla} \cdot \vec{F},
\]

(12)
where $v'$ is the eddy northward velocity, $q'$ is the eddy quasi-geostrophic potential vorticity, the overbar is a zonal average, $\rho_0$ is the basic state density, and $\vec{F}$ is the quasi-geostrophic Eliassen-Palm flux vector. Thus, a framework for today’s Eliassen-Palm flux diagnostics had been established during the 1960s.

3 Some Closing Comments

The 1960s were a very special time for me since that period spanned the time I finished high school, did my undergraduate education, earned my PhD, and became an Assistant Professor. I was very fortunate to have known, and still know, many of the individuals whose work I have cited here. It is no coincidence that I have worked on many of the topics discussed here since such an excellent foundation for my work had already been established by others by the time I entered the field.

While preparing this paper, I realized that what began as a historical review, also might serve as a good introduction to middle atmosphere dynamics. It was my pleasure to prepare this paper.

Acknowledgments. Support for M. Geller, while preparing this paper, came from the National Science Foundation’s Large-scale Dynamics program and from the National Aeronautics and Space Administration’s Atmospheric Chemistry and Modeling Program.

References


