Generation of Inertia–Gravity Waves in a Simulated Life Cycle of Baroclinic Instability

DONAL O'SULLIVAN and TIMOTHY J. DUNKERTON
Northwest Research Associates, Bellevue, Washington
(Manuscript received 3 June 1994, in final form 5 April 1995)

ABSTRACT

The excitation and propagation of inertia–gravity waves (IGWs) generated by an unstable baroclinic wave was examined with a high-resolution 3D nonlinear numerical model. IGWs arose spontaneously as the tropospheric jetstream was distorted by baroclinic instability and strong parcel accelerations took place, primarily in the jetstream exit region of the upper troposphere. Subsequent propagation of IGWs occurred in regions of strong windspeed—in the tropospheric and stratospheric jets, and in a cutoff low formed during the baroclinic lifecycle. IGWs on the flanks of these jets were rotated inward by differential advection and subsequently absorbed by the model’s hyperdiffusion. Although absorption of IGWs at the sidewalls of the jet is an artifact of the model, IGW propagation was for the most part confined to regions with an intrinsic period shorter than the local inertial period. Only a few IGWs were able to penetrate the middle stratosphere, due to weak winds or an unfavorable alignment of wavevector with respect to the mean flow.

IGWs are important both as a synoptic signal in the jetstream, which may influence subsequent tropospheric developments, and as a source of isentropic or cross-isentropic mixing in the lower stratosphere. The authors’ results demonstrated for the first time numerically a significant isentropic displacement of potential vorticity isopleths due to IGWs above the tropopause. Since conditions for IGW propagation are favorable within a jet, a region of strong isentropic potential vorticity gradient, it is likely that inertia–gravity waves affect the permeability of the lower stratospheric vortex and may in some instances lead to stratosphere–troposphere exchange.

1. Introduction

Inertia–gravity waves (IGWs) are ubiquitous within the atmosphere, especially the upper troposphere and lower stratosphere. These motions are characterized by short vertical wavelengths (≈1–5 km) and small intrinsic frequencies \( \omega \) near the Coriolis frequency \( f \) (Barat 1983; Yamanaka and Tanaka 1984; Thomas et al. 1992; Sato 1994). Their horizontal scale, which is usually estimated rather than measured, is thought to be large (≈200–1000 km) compared to mesoscale gravity waves. While momentum transport is ascribed to mesoscale waves (Fritts 1984), it is the role of IGWs in constituent mixing, and not momentum transport, that is probably more relevant. IGW mixing may be isentropic (Pierce and Fairlie 1993; Ledwell et al. 1993), like that of synoptic-scale baroclinic waves, or cross isentropic due to instabilities arising within the wave field (Dunkerton 1984; Fritts and Rastogi 1985). Notwithstanding the importance of IGW propagation, breaking, and absorption, one would also like to know something about the excitation of these waves and whether their geographical and climatological variability can be understood in terms of the postulated forcing mechanisms.

The origin of inertia–gravity waves is ambiguous because regions of strong mean flow are a favored locus of IGW excitation and propagation (Dunkerton 1984). IGWs observed within or above the tropospheric jetstream were not necessarily excited there; due to their slow propagation, the source may have been as far away as the Tropics (Usimaru and Tanaka 1990). Case studies often suggest a connection between IGWs and nearby synoptic weather systems (Sato 1989, 1993, 1994; Chan et al. 1991; Thomas et al. 1992; Eckermann and Vincent 1993). Therefore it is likely that extratropical sources dominate. That gravity wave variance increases above oceanic frontal systems and jetstreams (Fritts and Nastrom 1992) indicates that sources other than topography are important—an essential observation for understanding gravity wave transport in the Southern Hemisphere and Tropics. Part of the increase is undoubtedly due to high-frequency waves launched by moist convection or shear instabilities within frontal zones. (It was primarily mesoscale variance that was observed by Fritts and Nastrom.) However, one also expects a low-frequency IGW component due to the unstable baroclinic wave that would exist in the absence of convection, arising from the transiency of tropospheric flow as it relaxes to a balanced state. In a rotating fluid this process is referred to as geostrophic adjustment (Blumen 1972; Luo and Fritts 1993, and

© 1995 American Meteorological Society
references therein). During the adjustment, gravitational oscillations emanate and disperse away from the source. Geostrophic adjustment was interpreted by Vallis (1992) as a process that minimizes the energy for a fixed distribution of potential vorticity (PV). This interpretation assumes that the adjustment is rapid compared to the evolution of PV. General methods for finding energy extrema that allow rearrangements of PV as a materially conserved quantity were discussed by Vallis et al. (1989) and Shepherd (1990).

Many studies examined geostrophic adjustment in an idealized way, as linear relaxation from an initial imbalance of mass or momentum, with simplified (Cartesian or axisymmetric) geometry. In real life the adjustment involves complicated, three-dimensional motion with short vertical and (ultimately) horizontal scales. The need for consistent vertical resolution in primitive equation models was stressed by Lindzen and Fox-Rabinovitz (1989), and their comments pertain equally well to geostrophic adjustment. Only within the last few years have computational resources become available to make a realistic modeling study of cyclostrophic or geostrophic adjustment practical (e.g., Lelelong and McWilliams 1993). One rather fundamental point that has not been extensively addressed is the mechanism by which an initially balanced flow evolves into an unbalanced state, followed by geostrophic adjustment (Van Tuyl and Young 1982; Duffy 1990). It is expected that inertia–gravity waves arise spontaneously in primitive equation models (Warn and Menard 1986). Since IGWs are problematic to weather prediction, the emphasis has been mainly on how to eliminate them, rather than explicit modeling of IGW propagation and absorption. This omission is surprising in view of the large body of literature on ageostrophic circulations within synoptic weather systems (e.g., Cammas and Ramond 1989; Keyser et al. 1989, and references therein) and their role in frontogenesis, tropopause folding, and stratosphere–troposphere exchange (Shapiro 1980; Keyser and Shapiro 1986; Shapiro and Keyser 1990). By analogy to tidal theory, the ageostrophic motion contains a “forced” evanescent component (the meridional circulation) and a “free” radiating component (the inertia–gravity waves).

Recently the authors performed numerical simulations to examine the evolution of IGWs in a lifecycle of midlatitude baroclinic instability. Pattered after the classic experiments of Simmons and Hoskins (1978, 1980), our simulations described the nonlinear lifecycle of an unstable baroclinic wave assuming wavenumber 6 symmetry on the sphere but incorporated much higher vertical resolution and a domain extending into the lower stratosphere in order to explicitly model the evolving inertia–gravity waves and their radiation into the middle atmosphere. This work was complementary to several other studies (beginning with Ley and Peltier 1978; see also Uccellini and Koch 1987; Gall et al. 1988; Garner 1989; Benard et al. 1992a,b; Blumen 1992; Snyder et al. 1993) concerned with gravity wave excitation in relation to frontogenesis. We attempted to simulate inertia–gravity waves excited in the upper troposphere by an evolving baroclinic instability, considering their vertical as well as horizontal propagation. Numerical results demonstrated the role of baroclinic instability in exciting inertia–gravity waves, due to rapid deformation of upper-tropospheric flow and strong parcel accelerations.

Although our integrations were designed to yield well-resolved inertia–gravity waves independent of model resolution, only the excitation and initial propagation of IGWs could be realistically simulated. Their asymptotic propagation and breakdown requires vertical and horizontal resolution well beyond that used here (e.g., Dunkerton and Robins 1992). Resolved patterns of horizontal velocity divergence were nonetheless quite complex, containing an assortment of gravity waves and IGWs due to surface frontogenesis as well as parcel accelerations in the upper-level jet. This paper emphasizes geostrophic adjustment as the numerically best resolved, and probably most fundamental, mechanism of IGW excitation. Section 2 describes the numerical model. Section 3 discusses the baroclinic lifecycle and the resulting IGWs, examining IGW wave parameters and Lagrangian parcel accelerations. Some IGW-related issues are discussed briefly in section 4.

2. Model description

Simulations were performed with a slightly modified version of the 3D, global, hydrostatic primitive equation model of Young and Villere (1985). This is a spectral transform model with finite differencing in the vertical. There is no topography or moisture. Subgrid-scale processes are represented by sixth-order horizontal hyperdiffusion $\nu_6 \nabla^6$ acting on the vorticity, divergence, and temperature fields. The spectral truncation in runs described here, unless stated otherwise, was triangular at total horizontal wavenumber 126 (denoted T126) — approximately equivalent to $1^\circ$ of horizontal grid spacing. Hyperdiffusion is needed to prevent nonphysical behavior as sharp horizontal gradients are generated at the smallest resolvable scales. In practice, $\nu_6$ corresponding to an $e$-folding time ($\tau$) of 1 h or less for the smallest scale is sufficient for stable integrations. At T126, $\nu_6 = 1.36 \times 10^{25}$ m$^6$ s$^{-1}$ was used, corresponding to a damping time of about 20 min for the smallest resolved scale.

The vertical coordinate was uniformly spaced in log pressure with $\Delta z = 700$ m, unless stated otherwise, and the domain extended vertically from $z = 0$ to 35 km (51 levels). Boundary conditions assumed zero geometric vertical velocity at the lowest level, while log pressure vertical velocity was zero at the uppermost level. There was no surface friction.

As discussed in section 3d, comparison simulations were performed at various combinations of horizontal
and vertical resolution, for example, T42, T84, T126, and T156, and \( \Delta z = 350 \) or 700 m. Our custom has been to maintain approximately the same hyperdiffusion damping time for the smallest resolved scale, regardless of horizontal truncation. Higher horizontal resolutions therefore imply smaller hyperdiffusion coefficients and vice versa, unless stated otherwise.

Our primary interest was the generation of IGWs by geostrophic adjustment of the tropospheric jetstream as it is distorted by a developing baroclinic wave. For this purpose, the evolution should be realistic but not overly complicated. Following Simmons and Hoskins (1978), the model was initialized with a small amplitude wave-number 6 normal mode superposed on a zonally uniform, baroclinically unstable flow. This procedure yields within several days of model integration a well-resolved and fairly realistic finite perturbation to the jetstream, followed by a barotropic decay phase as described in section 3a. There are of course six identical lifecycles around the Northern Hemisphere.

For time integration the model used a semi-implicit scheme (Hoskins and Simmons 1975) with time step of 600 s, as required by the Courant–Friedrichs–Levy condition for windspeeds attained in the simulation. As we were mainly interested in the inertia–gravity waves, and it is known that the semi-implicit scheme affects (high frequency) gravity wave phase speeds, two T84 simulations of the same lifecycle were compared using either the semi-implicit scheme or a fully explicit, centered-difference time integration scheme. These simulations used time steps of 1200 and 150 s, respectively (demonstrating the computational advantage of the semi-implicit scheme). Comparing simulations during the mature stage of the lifecycle showed identical lifecycles occurring. Development of the lifecycle’s mature stage was delayed by about 1 day in the semi-implicit case, probably reflecting the different treatment of initial condition by the two schemes. Once this stage was reached, subsequent evolution did not differ noticeably. In particular, the low-frequency inertia–gravity waves of interest were insensitive to the choice of time integration scheme and appeared nearly identical in the two cases.

3. Baroclinic lifecycle and IGW excitation

a. Baroclinic lifecycle simulation

As described above, a baroclinic lifecycle analogous to that of Simmons and Hoskins (1978) was simulated with resolution adequate to model inertia–gravity waves. The unstable baroclinic wave developed on an initially axisymmetric jet centered at 45°N as shown in Fig. 1. The initial state contained several realistic features, namely, a change of static stability across the tropopause \( N^2 = 1.5 \times 10^{-4} \text{ s}^{-2} \) in the troposphere,
\( N^2 = 3.5 \times 10^{-4} \text{ s}^{-2} \) in the stratosphere), a sloping tropopause, a stratospheric polar night jet, and wind-speed minimum between stratospheric and tropospheric jetstreams. The initial state, with its strong jet displaced north of the climatological jetstream, might represent a cross section through midoceanic (Pacific or Atlantic) storm tracks during winter. It was disturbed by a small wavenumber 6 perturbation, which subsequently grew via baroclinic instability to finite amplitude. After 7 days of linear growth the baroclinic wave appeared as undulations on the westerly current.

Discussion of the baroclinic wave and accompanying IGWs begins on day 9 of the model run, by which time the baroclinic wave at upper levels attained finite amplitude. At midtropospheric levels an amplifying ridge, and deepening trough and low pressure center, show that the wave is still in its growth phase. Days 10 and 11 may be considered the wave’s mature stage as growth ceases and the elongated SW-trailing trough sheds a cutoff low near 30°N (Fig. 2a). After formation of the cutoff low, the lifecycle enters its decay stage with the main westerly jetstream shifting northward on day 12 to 50°–60°N and becoming more zonally symmetric and intense (Fig. 2b). At this time the baroclinicity has been largely removed from midlatitudes where areas of weak horizontal temperature gradient predominate, flanked to the north and south by frontal zones.

At low levels the characteristic poleward heat flux of baroclinic instability is seen as the cold air, is advected southward and the warm air northward (Figs. 2c,d). Cyclogenesis and frontogenesis proceed rapidly with a surface low and warm/cold fronts well formed by day 9. By this time, the surface fronts have sharpened to the point that their width is limited by the model’s small-scale hyperdiffusion.

The low-level cyclogenesis and frontogenesis strongly resemble the observed development of synoptic systems as discussed, for example, by Shapiro and Keyser (1990). The near-surface temperature on day 9 or 10 shows the so-called T-bone frontal stage of the baroclinic lifecycle, a sharp front at 60°–65°N extending back into the low pressure center, while the cold front, advancing eastward of the low, forms the vertical stem of the T. A pocket of warm air has been secluded behind the cold front near 55°N, 15°W on day 9 and 10, as has been observed over oceans when cyclones reach the mature stage of their lifecycle (apparently because of the lower surface friction there).

The baroclinic lifecycle also resembles that of Simmons and Hoskins (1980) and Thornicroft and Hoskins (1990), even showing the secondary frontal development discussed by the latter authors. This occurs on days 11 and 12 as the upper-level circulation associated with the cutoff low drifts over the surfce cold front, initiating a small-scale frontal wave near 40°N, 20°W (Fig. 2d) — an example of Peterssen’s “Type B” cyclogenesis.

In the tropopause region the evolution is like that described for the midtroposphere. The jetstream is similarly distorted as the baroclinic wave radiates upward. After a period of linear growth, the wave saturates and a cutoff low forms by day 11. This is a period of rapid geostrophic adjustment throughout the troposphere, especially at upper levels where the jetstream is strongest.

We next examine evidence of gravity wave activity associated with surface frontogenesis and geostrophic adjustment in the upper troposphere. The field of horizontal velocity divergence is useful for isolating gravity waves from the large-scale quasigeostrophic flow because of the divergent nature of such waves. (By contrast, the signature of IGWs in geopotential is hardly visible.) The horizontal divergence field includes forced ageostrophic circulations accompanying the varying jetstream. As shown next, waves located away from this apparent source region show a radiating pattern of divergence distinct from that associated with the forced circulations.

### b. IGW generation during lifecycle

The divergent component of tropospheric flow evolves quickly and in a complicated manner during the baroclinic lifecycle. Analysis of horizontal velocity divergence (hereafter \( \nabla \cdot v_H \)) shows gravity wave activity associated with the surface fronts and cross-frontal ageostrophic circulations. These surface-related features weaken with increasing height because frontogenesis is strongest at ground level, and the radiating waves are dissipated as the mean flow turns with height. At the tropopause level and above (10–25 km), gravity wave activity increases dramatically during the lifecycle’s mature stage. Figures 3a–d show \( \nabla \cdot v_H \) at 130 mb (near 14-km height) on days 9–12. The geopotential field (Figs. 4a–d) shows the location of the jetstream at this level. The first appearance of significant gravity wave activity in the upper troposphere occurs on day 9. Over the next day this activity quickly increases as a gravity wavepacket reaches 130 mb. The wavepacket is strongest on day 11, having advanced eastward with a secondary wave train splitting and propagating southward around the west side of the cutoff low. By day 12, the main packet has advanced rapidly eastward due to acceleration of the jet. A third wave train appears on the east side of the cutoff low, unrelated to the first packet. Over subsequent days, fragments of \( \nabla \cdot v_H \) become more widely spread and less coherent (not shown).

Figures 3 and 4 show that the upper-level gravity waves tend to be confined to the vicinity of the jetstream, propagating with wavevector parallel to the jet axis at midstream but at an angle to the jet along its flanks. This is also observed on the east and west sides of the cutoff low. Interestingly, at late times the remnants of IGWs in the cutoff low are carried back into the jet as the low itself is absorbed (not shown).
Fig. 2. (a, b) Geopotential at 503 mb, and (c, d) temperature at 964 mb on days 10 (a, c) and 12 (b, d) of the T126 model simulation. In this and following figures, two identical 60° sectors are displayed side by side.
Fig. 3. Horizontal velocity divergence at 300 mb on consecutive days of the T12 model simulation: (a) day 6, (b) day 10, (c) day 16, (d) day 22. Contour interval is 10⁻⁴ s⁻¹, with thick (thin) contours denoting positive (negative) values.
FIG. 4. As in Fig. 3 but for 150-mb geopotential. Vectors show the horizontal wind, the largest vector corresponding to 62 m s$^{-1}$ in (c).
While model gravity waves in the mid- and lower stratosphere can be related to frontal activity, those in the upper troposphere and lower stratosphere generally cannot and appear to have a different origin. In order to provide a 3D view of these waves, and to clarify their origin, cross-sectional views of the wavepacket are presented in Figs. 5 and 6. Figures 5a–b show latitude–height sections of $\nabla \cdot \mathbf{v}_w$ on days 10 and 11 at longitudes that intersect the gravity wave packet on each day. In these cross sections the dominant feature is the enhanced gravity wave activity at upper levels, between the tropopause and ~20 km. Frontal features at low levels are also evident, largely unconnected to the activity aloft. The upper-level waves appear like a stack of pancakes, with phase fronts sloping down on either side. These surfaces also tend to parallel the isopleths of constant windspeed.

Figures 6a–d show longitude–height cross sections of $\nabla \cdot \mathbf{v}_w$ on days 9–12 at 55°N, a latitude intersecting the main wavepacket on those days. This again shows the upper-level wavepacket to be the dominant feature. The waves have phase surfaces sloping uniformly upward to the west. It is convenient here that the jetstream exit region, located near the crest of the geopotential ridge, is propagating eastward almost parallel to a latitude circle during this stage of the lifecycle. Thus the longitude–height cross section at 55°N contains the exit region on these days. Figure 6 illustrates an important feature of the upper-level gravity waves, namely that their phase and group speed are small (or zero) relative to the jetstream exit region. This is best seen in animations of the cross-sectional view, but it can also be discerned from the plots for day 9–11, where the first three distinct wave crests (counting upward from the bottom of the wavepacket) are advected eastward with the jetstream exit region. In animation it appears that the jetstream pattern and attendant gravity waves are jointly advected eastward as a coherent structure. While the waves are approximately stationary relative to the jetstream pattern, they are, of course, in a strong westerly flow, so their intrinsic phase velocity is rapidly westward. The intrinsic group velocity is also westward as discussed below. (Likewise, the intrinsic propagation of the jetstream pattern is westward with respect to the fluid in the upper troposphere.) An upward group velocity can be discerned for these waves over the 3-day period of Fig. 6, as the wavepacket expands upward with time.

The impression gained from Fig. 6 is of gravity wave generation in the upper troposphere near the level of maximum wind in the vicinity of the jetstream exit region. There is a striking vertical asymmetry with most of the IGWs on the upper side of the jet, where static stability increases dramatically with height. Over the 3 days of the "mature lifecycle" stage the gravity waves radiate upward into the lower stratosphere where, except for one or two small packets, most of them approach a critical surface and are absorbed. (Note that critical layer absorption can occur when the mean flow turns with height—the windspeed need not approach the wave's phase speed.) There seems to be continual generation of gravity wave activity near the level of maximum wind during this time, since the lower edge of the wavepacket does not move upward. Van Tuyl and Young (1982) noted similar stationarity of unbalanced motions (IGW) relative to the jet pattern in their two-level mechanistic experiments of geostrophic adjustment. As noted by those authors, IGWs are a sub-synoptic motion that can have significant amplitude in the exit region of strong jetstreams. In our high-resolution simulations IGWs explain a substantial part of the divergent wind component and at maximum amplitude (3–5 m s$^{-1}$) constitute about 10% of the total wind near the tropopause.

To test the conclusion that IGWs are due to geostrophic adjustment of the upper level jet and are not caused by gravity wave generation near the surface frontal regions, the simulation was repeated with strong damping of horizontal velocity divergence in the lowest 5 km. Applying additional sixth-order hyperdiffusion to divergence only—500 times stronger at the surface than the standard value and decreasing linearly to zero at 5 km—gave results similar to the standard case but with much smoother horizontal velocity divergence at low levels, effectively eliminating any gravity wave radiation from the surface fronts. At upper levels the flow was very similar except for a slight reduction of jet core speed (~5%). Most importantly, IGWs were again found, unchanged other than a slight reduction of amplitude (~10%).

The IGWs are well resolved in this T126 simulation, at least near the wavepacket center where their horizontal wavelength is 600–1000 km. This was borne out by comparison to an identical T84 simulation showing a very similar IGW packet with identical phase structure on day 11 at 130 mb. Differences were noticeable at the jet flanks where the IGW horizontal wavelength undergoes contraction as discussed in sections 3c,d.

c. Description of IGWs at upper levels

The gravity waves found in the simulated baroclinic lifecycle are of inertia–gravity type; that is, $N \gg |\hat{\omega}| \gg |f|$, where $N$ is the Brunt–Väisälä frequency, $\hat{\omega}$ is the intrinsic wave frequency, and $f$ is the Coriolis parameter. Examining the center of the main IGW packet between days 10 and 12 indicates that the IGWs have a period with respect to the ground of about 12–24 h and propagate eastward at a slower rate than the speed of the (westerly) jet. Following Dunkerton (1984), we

---

1 The term "exit region" is used to describe the relatively weak, tightly curved flow downstream of the jet maximum. Parcels in this area do not decelerate to zero but are deflected southward into the west side of the developing cutoff low.
Fig. 5. Latitude–height cross section of horizontal velocity divergence at two longitudes following the main IGW packet on consecutive days: (a) day 10, 20°E; (b) day 11, 20°W. Contour interval is $10^{-3}$ s$^{-1}$. Windspeed superposed with shading increment 10 m s$^{-1}$ as in Fig. 1.
adopt the convention that the wavevector points opposite the mean flow, reflecting the fact that phase propagation is westward with respect to parcels in the jet. Thus the intrinsic frequency is positive by convention, and the frequency relative to the ground is mostly in the range \(-2\pi/\text{day} < \omega \approx -\pi/\text{day}\). For a zonally oriented wavevector with zonal wavenumber \(k\) and a zonal flow \(\bar{u}\), the intrinsic frequency \(\hat{\omega} = \omega - \bar{u}k\). Taking \(k \approx -2\pi/(600 \text{ km})\), \(\omega \approx -2\pi/(14 \text{ h})\), and \(\bar{u} \approx 30 \text{ m s}^{-1}\), representative of the zonal wind through the wavepacket at 130 mb on day 11, yields \(\hat{\omega} \approx 1.9 \times 10^{-4} \text{ s}^{-1}\). Thus, \(\hat{\omega} > f(\approx 1.2 \times 10^{-4} \text{ s}^{-1} \text{ at 55°N})\). The waves propagate westward relative to the wind with an intrinsic period of about 9 h, which is shorter than the inertial period of 14.6 h at this latitude.

This quick estimate of intrinsic frequency may be compared to that derived from the dispersion relation for plane hydrostatic IGWs of the form \(\phi = \phi_0 \exp(\mathbf{k} \cdot \mathbf{x} - \omega t)\); namely,

\[
\hat{\omega}^2 = -f^2 + \frac{N^2(k^2 + l^2)}{m^2 + 1/(4H^2)}
\]  

(3.1)

with \(l\) and \(m\) being the meridional and vertical wave-numbers, respectively, and \(H\) is the density scale height. (The second term in the denominator is negligible compared to the first.) On day 11 the wavepacket at 130 mb is centered near 55°N, and the horizontal mean wind and wavevector are both directed zonally. The horizontal wavelength is approximately 600 km with \(k^2 \gg l^2\), and, from Fig. 6c, the vertical wavelength is about 4 km. Substitution into the dispersion relation yields \(\hat{\omega} \approx 1.7 \times 10^{-4} \text{ s}^{-1}\), which agrees reasonably well with the other estimate. The two values of wave frequency at 130 mb indicate that gravity wave parameters are consistent with an IGW interpretation provided that estimates are made near the center of the wavepacket. Exact agreement is not expected, since the waves are not plane and the dispersion relation ignores latitudinal and vertical shear. A more detailed analysis of \(\hat{\omega}\) is presented below.

The group velocity of the IGW packet may be derived from the dispersion relation:

\[
e_g = \frac{N^2}{m^2 \hat{\omega}} \left[ k, l, \frac{(k^2 + l^2)}{-m} \right] + \bar{u}.
\]  

(3.2)

For wave parameters near the 130-mb jet axis (where \(|k| \gg |l|\)):

\[
e_g = (-7.1 \times 10^2, -0, 4.7) \text{ km day}^{-1} + \bar{u}.
\]  

(3.3)

The vertical group velocity is positive since \(m\) is negative. Upward group velocity is also indicated by a wind hodograph taken at day 11 through the center of the wavepacket at 53°N, 23°W (Fig. 7). The hodograph of perturbation winds (a five-point running mean having been removed) shows anticyclonic (clockwise) rotation of wind with increasing height, implying that the group velocity is directed upward. The major axis of the hodograph ellipse, which is parallel to the direction of wave propagation, is oriented from slightly south of east to north of west. The perturbation horizontal winds associated with the IGWs have magnitudes of about \((5, 3) \text{ m s}^{-1}\) for components parallel and perpendicular to the direction of wave propagation. At other locations we find similarly that the major axis of the hodograph ellipse is oriented in the direction of wave propagation suggested by Fig. 3 (not shown).

The horizontal group velocity is directed upstream relative to the wind and has a magnitude of about \(-8 \text{ m s}^{-1}\). The wavepacket is seen to propagate eastward at about \(15-25 \text{ m s}^{-1}\), accelerating to the higher value after day 11. Although this agrees with predictions, the observed distribution of the wavepacket does not necessarily indicate the group velocity since it is affected by mechanisms of excitation and absorption, as well as propagation.

Upward radiating IGWs such as these will encounter a "Jones critical level" when \(|\hat{\omega}| \rightarrow |f|\). Using wave parameters typical of the center of the packet, oriented in the direction of the mean flow, we estimate \(\bar{u}_{cl} \approx 25 \text{ m s}^{-1}\) as the critical mean wind speed. In this simulation IGWs are seen to penetrate slightly higher than the Jones critical level, to levels where the westerly windspeed has decreased into the range \(\bar{u} \approx 10-25 \text{ m s}^{-1}\). This penetration above the Jones critical level is likely due to transience in the forcing of the IGWs, the derived \(\bar{u}_{cl}\) being appropriate for steady-state conditions. IGW breakdown is thought to occur via Kelvin–Helmholtz (KH) instability due to vertical shear of the wave's transverse wind component. Our large-scale model cannot resolve the KH instabilities, of course, so the IGWs are absorbed by hyperdiffusion as they approach the critical level. This could happen either through contraction of the vertical wavelength or rotation of the horizontal wavevector away from the direction of the mean flow, both of which would reduce the intrinsic frequency and vertical group velocity. We believe in this instance (as explained further in section 3d) that diffusion of model IGWs is due primarily to contraction of horizontal wavelength and rotation of the horizontal wavevector away from the direction of the mean flow. Comparing a T84 version of this simulation (with \(\Delta z = 700 \text{ m}\)) to an identical T84 run with double vertical resolution (\(\Delta z = 350 \text{ m}\)) gave strikingly similar IGW and vertical wavelengths. Propagation of IGWs in the model is therefore not limited by vertical resolution. Furthermore, by repeating the simulation at higher horizontal resolution while keeping the hyperdiffusion coefficient unchanged gave almost identical IGWs. Thus, the simulated IGWs are well resolved and are not significantly affected by inconsistent vertical or horizontal resolution as discussed by Lindzen and Fox-Rabinovitz (1989).

Along the flanks of the jet the horizontal wavevector is rotated inward to response to differential advection
by the jet. As seen in Fig. 3, the horizontal wavelength along the flanks of the jetstream has contracted somewhat relative to that along the jet axis. This is consistent with waves refracting into the jetstream as discussed by Dunkerton (1984); for example,

$$\frac{dl}{dt} \approx -k\vec{u},$$

(3.4)

etc. On the southern flank of the jet \(l > 0\) so \(c_w > 0\), while \(l < 0\) and \(c_w < 0\) on the northern flank.

A consequence of rotation and contraction is that wave activity along the jet flanks is rapidly dissipated by the model’s hyperdiffusion. This mainly accounts for the sensitivity of simulated IGWs to the model’s horizontal resolution: at lower resolution (T84), IGW activity away from the jet axis is lost even more readily to diffusion. We expect that with finer horizontal resolution well beyond T126, the trailing IGWs along the flanks of the jet would be resolved for a longer time, evolving to even shorter horizontal wavelengths. Some evidence in support of this claim is provided by results of a T156 run as discussed in section 3d.

That the waves are being absorbed along the lateral boundaries of the wavepacket is also suggested by calculations of intrinsic frequency, determined objectively from the spatial distribution and temporal behavior of \(\nabla \cdot \vec{v}_w\). Wavevector orientations and magnitudes were calculated at approximately 1° intervals within the wavepacket (at 130 mb) by applying a high-pass spatial filter to \(\nabla \cdot \vec{v}_W\) (to eliminate large-scale ageostrophic circulations) and evaluating the horizontal gradient of this quantity at the zero crossing. Although the phase function is ill behaved in a complicated wave field such as this, a local wavevector direction may be determined uniquely as parallel (or antiparallel) to \(\nabla (\nabla \cdot \vec{v}_W)\) at the zero crossing. Half-wavelengths were estimated as the distance between adjacent zero crossings in the direction of the wavevector. To get reasonably smooth results, individual estimates within a 4 × 4 matrix of grid points were binned together and interpolated back to the original grid. A similar binning procedure was used for frequency \(\omega\), crudely estimating the local half-wave period as the time interval between adjacent zero crossings in a bandpass-filtered time series of \(\nabla \cdot \vec{v}_W\) at each grid point. The wavevector direction and sign of \(\omega\) were carefully determined by examining the phase propagation around the time of interest. Wavevectors were aligned in the direction of intrinsic phase propagation, that is, with a component antiparallel to the mean flow. The sign of \(\omega\) was negative in all but a few places; this quantity was typically smaller than the advective term \(k\vec{u}\) by a factor of 3–4. To evaluate the intrinsic frequency \(\hat{\omega} = \omega - k\cdot\vec{u}\), a small amount of spatial smoothing was applied to the horizontal wind, since IGWs are not part of the “mean flow.”

Figures 8a–c show \(\nabla \cdot \vec{v}_W\) (with high-pass spatial filter) at 130 mb on days 9.5, 11, and 12.5, illustrating the early, middle, and late stages of IGW evolution at this level, respectively. In the early stage (Fig. 8a) there is a single wavepacket in the jet stream exit region, with most wavevectors pointing backward and to the right with respect to the local mean flow (for an observer facing downstream). By day 11 (Fig. 8b), the packet has begun to split, part of it going south into the cutoff low. Wavevectors in both halves converge into the jet, and a few waves manage to sneak across the saddle of the cutoff low into a region where they do not belong. By day 12.5 (Fig. 8c), the main part has advected rapidly eastward to a point immediately northwest of the cutoff low (reentering from the left side of the diagram), while the wavepacket trapped within the cutoff low remains largely stationary. Waves crossing the saddle rapidly disappear, replaced by a third packet (unrelated to the others), on the east side of the cutoff low, on a collision course with the first.

The intrinsic frequency on day 9.5 (Fig. 8d) shows most values lying in the range \(f–4f\), with a flat ledge of smaller values (slightly below \(f\)) along the northeastern boundary of the wavepacket. The variation of \(\hat{\omega}\) is due mainly to the angle of wavevector with respect to the mean flow and not the wavelength (which is rather uniform on this day). Although penetration beyond \(\hat{\omega} = f\) would require that a steady-state wave be evanescent, in this case the phenomenon does not persist (as noted above, in connection with the vertical propagation) and is probably due to transients excited by geostrophic adjustment. On day 11 (Fig. 8e) the region of subinertial frequency has disappeared, although by this time there is another such region lying within the saddle of the cutoff low, which is likewise temporary and quickly damped. Most values of \(\hat{\omega}\) are again in the range \(f–3f\). By day 12.5, a larger percentage of wavevectors are becoming nearly orthogonal to the mean flow, and values of \(\hat{\omega}\) are somewhat smaller, mostly \(f–2f\) (Fig. 8f).

It must be noted that the absorbing action of the sidewalls of the jet is an artifact of the model, being sensitive to the model’s horizontal resolution and hyperdiffusion. These results cannot address the issue of whether the surface \(\chi = f\) acts as a critical layer to absorb IGWs (as in vertical shear) or as a turning point to reflect them (as in tidal theory). This matter was discussed by Kitchen and McIntyre (1980) among others but deserves further study. There is little, if any, evidence for reflection of model IGWs, although such an effect might be masked by the rapid evolution of the flow. It was suggested by Dunkerton (1984) that, due to horizontal and temporal variations of the basic state, any natural waveguides for IGWs are likely to “leak” wave activity rather than contain it. In the simulations some wave activity manages to escape to the stratosphere, but scale-dependent damping is the dominant effect.

Radiation to the middle stratosphere occurs only in the trailing part of the main wavepacket, where IGWs
Fig. 6. Longitude–height cross section of horizontal velocity divergence at 55° latitude on consecutive days: (a) day 9; (b) day 10; (c) day 11; (d) day 12. Contour interval is 10⁻³ s⁻¹. Windspeed superposed as in Fig. 5 with contour interval 10 m s⁻¹.
Fig. 7. Hodograph of the vertical profile of perturbation horizontal wind components on day 11 at 53°N, 23°W. Perturbation denotes deviation from winds smoothed with a five-point running mean vertically. Altitude of model levels indicated in km.

tilt northward into the polar night jet (not shown). Propagation at this location is possible on account of increasing mean windspeed despite the wavevector becoming nearly orthogonal to the mean flow. Another packet approaches 20 km on the west side of the cutoff low (not shown). These results demonstrate that vertical propagation of model IGWs is limited to the region of strong mean windspeed where tropospheric and stratospheric flows have approximately the same direction. The baroclinic wave distorts the tropospheric flow from a zonal orientation but is itself evanescent in height (i.e., does not similarly distort the middle stratosphere). Only a few tropospheric IGWs can therefore propagate into the midstratospheric circumpolar vortex.

d. Effects of horizontal and vertical resolution

Model IGWs are unaffected by vertical resolution if the grid spacing is sufficiently fine to describe the rotation of perturbation velocity with height, as shown in Fig. 7. Smaller Δz did not affect the evolution of model IGWs, indicating that the IGWs are primarily sensitive to horizontal hyperdiffusion.

Integrations were performed at horizontal resolution T42, T84, T126, and T156 with hyperdiffusivity coefficients \( \nu_s = 1.37 \times 10^{27}, 1.02 \times 10^{26}, 1.36 \times 10^{25}, 2.52 \times 10^{24} \text{ m}^6 \text{ s}^{-1} \), respectively, such that the damping time at the smallest resolved scale was nearly the same in each run (~20–30 min). These integrations illustrate quite dramatically the sensitivity of model IGWs, as shown in Figs. 9a–d. Baroclinic lifecycles in each run were at virtually the same stage at day 10, but major differences are seen in model IGWs. Contour intervals are the same in each panel, indicating a large increase of IGW amplitude between T42 (when IGWs were largely absent) and T126. Figure 9 also demonstrates that certain parts of the divergence pattern are adequately resolved at lower resolution, while others are not. Increasing horizontal resolution beyond T84 did not significantly affect the horizontal scale of IGWs at the rear of the wavepacket nor the structure of forced ageostrophic circulations evident at T42. IGW structure at the center of the wavepacket was largely unchanged going from T126 to T156. Differences were observed mainly on the fringes of IGW packets, for reasons discussed in the previous subsection.

Interpretation of these results is complicated by possible changes of the baroclinic wave as resolution is increased, for example, sharper curvature of the jet and stronger gradients of velocity. An example of windspeed at day 10 for T84 is shown in Fig. 9e. Windspeeds and velocity gradients were slightly stronger at higher resolution, by a few percent. The basic pattern of horizontal velocity divergence within the source region (Fig. 9f) was similar, except for substantially more finestructure at higher resolution due to IGWs (not shown). Therefore, changes of the jetstream pattern were minor compared to the observed changes of IGWs, leading us to believe that the primary effect of increasing resolution in this range (T42–T156) is to allow the region of imbalance to project better onto the IGW manifold, irrespective of any change in IGW sources. We cannot rule out the possibility that such sources may contract further in horizontal scale or become more intense locally, as resolution is increased beyond T156, with a resulting contraction of IGWs and further increase of amplitude.\(^2\) Indeed, upper-level frontogenesis is expected to excite mesoscale gravity waves. Model experience suggests, however, that such waves would add to the spectrum of IGWs already resolved. Important to this argument is the assumption that model IGWs are attributable to an adequately resolved region of imbalance and do not arise entirely from frontal collapse or from instabilities lacking a preferred horizontal scale.

e. Comparison with observations

The upper-level IGWs seen in this model study are broadly consistent with observations of IGWs near midlatitude jetstreams, except possibly regarding their direction of horizontal energy propagation. Typically, the lower stratospheric gravity wave field is dominated by quasi-monochromatic waves with horizontal wave-
Fig. 8. (a–c) Pattern of horizontal velocity divergence (with high-pass spatial filter, each plot normalized by its maximum amplitude) and (d–f) normalized intrinsic frequency $\omega/\Omega$ of IGWs at 130 mb on day 9.5 (a, d), day 11 (b, e), and day 12.5 (c, f). Isopleths of geopotential superposed as thin contours, indicating the location of jetstream and cutoff low. Contours of filtered, normalized divergence should not be confused with wave amplitude, which is much smaller at early times.
Fig. 9. Horizontal velocity divergence at day 10, 130 mb from various integrations: (a) T42; (b) T84; (c) T126; (d) T156. Contour interval as in Fig. 3. Thin contours display geopotential height at this level. (e, f) Windspeed and horizontal velocity divergence at day 10, 238 mb, with geopotential at this level superposed, from T84 run.
lengths of several hundred kilometers and vertical wavelengths of 1–4 km with upward energy propagation. These IGWs appear to be related to the tropospheric jet stream, though it has proven difficult to directly relate them to a specific source mechanism. Thomas et al. (1992) reported a case study based on VHF radar observations made at Aberystwyth, Wales (52°N, 4°W). They found southwestern IGW energy propagation away from the jet stream, which was passing to the north. Prichard and Thomas (1993) examined stratospheric IGWs over 23 days and also noted the dominant southward IGW propagation direction. In apparent contrast to this are the results of Sato (1994), who statistically analyzed IGWs observed by radar at Shigaraki, Japan (35°N, 136°E). She found the lower stratospheric IGWs to be most energetic in wintertime and to propagate southward toward the strong subtropical jet, which is persistently located just south of there during winter. Finally, in situ observations of IGWs made with an ER-2 aircraft off the California coast in April 1986 found large-amplitude IGWs in a jet stream exit region with propagation roughly parallel to the jet axis (L. Pfister 1994, personal communication). At 13–15-km height the vertical wavelength of these waves was 1–2 km, and the horizontal wavelength was deduced to be about 250 km.

Thus, while lower stratospheric IGWs primarily tend to propagate meridionally, there is some uncertainty about their relation to the jet stream. There may be large geographical variability in the relation of IGWs to nearby jetstream. Similarly, it is likely that there are several different types of jet stream developments that result in significant IGW generation. For zonally elongated imbalances to the jet stream, meridionally radiating (away from the jet) IGWs should result. However, when the zonal extent of the imbalance region is reduced relative to its meridional extent, then zonally propagating IGWs should become significant (Luo and Fritts 1993).

Although not discussed in section 3c, the simulation at a later time shows evidence of zonally aligned IGW phase fronts propagating sideways into the tropical lower stratosphere. The waves are quite weak compared to those near the jet, but their (N–S) wavevector orientation is consistent with IGW observations at Arecibo Puerto Rico, described by Sato and Woodman (1982) and later authors. This would suggest that not all observed IGWs in the Tropics are due to underlying convection but that some of them may originate in the midlatitude jet.

f. Generation of IGWs in the upper troposphere

Given that IGWs are difficult to observe in atmospheric data lacking fine vertical and temporal resolution, it would be desirable to relate the excitation of IGWs to observable parameters in the large-scale flow field. The following thoughts are based partly on earlier literature and partly on our modeling experience, suggesting how this might be done.

As discussed in section 1, IGW generation is expected during periods of geostrophic adjustment of the tropospheric jet stream accompanying synoptic-scale developments. As the flow evolves, imbalances develop between the mass, temperature, and velocity fields causing ageostrophic motions that cannot be supported in a balanced system (quasi-geostrophic, semi-geostrophic, nonlinear balance, etc.). In this event the balance breaks down, and IGWs are radiated away from the region of imbalance. There is no completely reliable way to diagnose the conditions required for IGW generation however, largely for lack of knowledge of the balanced state to which the flow is relaxing (Koch and Dorian 1988).

The horizontal wind divergence field near the level of maximum windspeed (~250 mb) shows the signature of slowly varying ageostrophic circulations that accompany the jet stream (Fig. 9f). These patterns are similar to those analyzed by Cammas and Raymond (1989) for jet streams where flow curvature was strong, for example, their Fig. 6. In such cases the familiar quadrupole pattern of upper-level divergence associated with the entrance and exit regions of a localized zonal jet stream are obscured by the divergence associated with the flow curvature. The appearance of a strong divergence center in the jet exit region at 250 mb, just upstream of the ridge, implies that air parcels rapidly advected through this region will experience a large local rate of change of divergence, suggestive of flow imbalance. In analogy to wave generation by flow over an obstacle, it is to be expected that the dominant waves generated by flow through this jet exit region should propagate with a wavevector approximately parallel to the jet axis, as observed in the model near the center of the jet. Also consistent with this idea, we find that the IGW pattern is approximately phase locked to the jet exit region over a period of several days (Fig. 6). Similarly, the IGW structure in the jet exit region is fairly insensitive to improvements in resolution.

Typical measures of imbalance rely on estimating the magnitude of those terms that are assumed to be negligible in any particular balanced-flow state. Thus, one might have a small Rossby number \( R_e = U/fL \ll 1 \) for geostrophic balance, a small Lagrangian Rossby number \( R_e^{(L)} = \|d\mathbf{V}/dt/f|V| \ll 1 \) for semi-geostrophic balance (since it implies \( d\mathbf{V}/dt \sim d\mathbf{V}/dt \)), and a small material derivative of horizontal divergence \( d\nabla \cdot \mathbf{v}_i/dt \ll 10^{-9} \text{s}^{-2} \) for the nonlinear balance equation.

In this study the plots of \( d\nabla \cdot \mathbf{v}_i/dt \) tended to mimic those of \( \nabla \cdot \mathbf{v}_i \) (e.g., Fig. 3), which we use to show the IGWs. This quantity is therefore an accurate but redundant estimate of imbalance. Instead, one would like a measure of flow imbalance that does not explicitly involve \( \nabla \cdot \mathbf{v}_i \) and may be diagnosed from the observ-
able flow. The Lagrangian Rossby number appears suitable for this purpose.

When applied to synoptic data, Koch and Dorian suggest an indicator of unbalanced flow is the Lagrangian Rossby number $R_o^{(L)} \gtrsim 0.5$ in a region where $|V| > 10 \text{ m s}^{-1}$ (to avoid spuriously large values of $R_o^{(L)}$ from the denominator being small). In practice, when calculating $dV/dt$ from synoptic data, the local wind tendency $\partial V / \partial t$ must often be neglected relative to the advective tendency $V \cdot \nabla V$ because of the crude time resolution of the observations. For this simulation we calculated both terms of $dV/dt$ and found that, indeed, the local wind tendency is usually small compared with the advective tendency $V \cdot \nabla V$, although locally it can amount to $\sim 30\%$ of the latter.

Figure 10 shows $R_o^{(L)}$ at 195 mb on day 11 of the baroclinic lifecycle with horizontal velocity divergence overplotted at 130 mb, showing the main IGW packet at $50^\circ - 60^\circ$N and a second wavepacket west of the cutoff low near $30^\circ$N. Nearer to the jet core, $R_o^{(L)}$ was calculated at a slightly lower level, since the wave source lies below where the waves are prominently seen. In this simulation large values of $R_o^{(L)}$ are found extending down to the level of the jet maximum, near 250 mb. In Fig. 10, these large values (heavy lines) are seen to coincide with the location of IGW activity. Two regions of large $R_o^{(L)}$ occur: one between $45^\circ$ and $60^\circ$N that overlies the main IGW packet in the jetstream exit region and a second region lying southwest of the cutoff low, suggesting that two separate regions of imbalance exist. To be sure, large $R_o^{(L)}$ suggests where an imbalance exists but says nothing about the intensity of resulting IGWs. Since IGWs radiate unbalanced energy away from the jet, their intensity should be approximately proportional to $|V|^2$. This is consistent with results of mechanistic numerical experiments by Van Tuy1 and Young (1982), who found that the strength of unbalanced motions (IGWs) increases roughly as the square of jet core strength. Thus, stronger IGW radiation is to be expected from regions of imbalance (large $R_o^{(L)}$) where winds speeds are weaker. This may explain why the IGWs are much stronger in the high-latitude region of imbalance than at low latitudes where winds speeds are weaker.

A combination of Lagrangian Rossby number and mean flow kinetic energy, then, provides a useful diagnostic of IGW excitation in the model. It remains to be seen whether it is systematically useful in observations.

4. Discussion

The ubiquitous nature of IGWs in the atmosphere has long been recognized, yet their effects are only beginning to be considered in the context of the general circulation. This situation is partly due to their subsynoptic scale and preferred occurrence at upper tropospheric levels and above. IGWs are nevertheless likely to play an important role in several aspects of atmospheric circulation.

Our simulations show that IGWs are generated in the upper troposphere, propagate energy upward, and are absorbed approaching the Jones critical level (either in weak winds or as the wavevector rotates away from the mean flow). As discussed in section 3c, IGW break-down is thought to occur via KH instabilities (unresolved in the model), which develop as amplitude increases and vertical wavelength decreases. These instabilities are expected in the region of strong transverse vertical shear, at which point the flow is statically stable (Dunkerton 1984; Fritts and Rastogi 1985). IGW breaking is, therefore, a potentially efficient means of cross-isentropic transport and may be responsible for most of the vertical diffusion found in the lower stratosphere, a region where systematic cross-isentropic transport due to diabatic heating is quite weak. Vertical mixing due to IGW breaking must vary geographically and seasonally, just as IGW generation does.

In addition to cross-isentropic transport induced by IGWs, there is an isentropic component due to advection by the horizontal velocity. One normally thinks of gravity wave motions as fast, reversible undulations superposed on the "slow manifold" of rotational motions (e.g., nonlinear baroclinic instabilities and breaking planetary Rossby waves), the latter being mainly responsible for isentropic mixing on a global scale. Parcel displacements due to IGWs, nevertheless, can in certain cases be large, due to their large horizontal scale and low intrinsic frequency—particularly as the critical level is approached and amplitude increases. In this case an air parcel advected through the IGWs will undergo a considerable horizontal excursion during a half wave period before the transverse velocity reverses sign.
Under such circumstances, any background tracer field with a nonzero gradient orthogonal to the IGW wavevector will experience considerable deformation. This is true, in particular, for potential vorticity (PV) near the tropopause. Such deformations of PV on an isentropic surface are in fact seen in these simulations. Figure 11 shows PV on the 430-K surface, illustrating large meanders of the (approximately conserved) PV contours across the jet in a region of strong IGWs. The meanders are rapidly strained to smaller scales by the background flow and lost because of limited model resolution. It is certain that they will lead to irreversible isentropic transport and that this will be especially strong in the cross-jet direction if real IGWs behave like those simulated here. In ongoing work we are investigating the effect of IGWs on isentropic tracer transport in the lower stratosphere by comparing transport and mixing in simulations with and without IGWs. This is relevant in light of recent contour advection studies (Plumb et al. 1994; Waugh et al. 1994; Waugh and Plumb 1994; Norton 1994) in which the isentropic tracer redistribution is computed from advecting winds known only at low resolution and as infrequently as daily. Recognizing the importance of unbalanced motions, Pierce and Fairlie (1993) speculate on the role of IGWs in enhancing cross-jet tracer dispersion, perhaps enhancing the permeability of the lower stratospheric vortex at its edge.

Figure 12 shows a latitude–height cross section through the IGW-induced meanders shown in Fig. 11. This illustrates the effect of quasi-horizontal isentropic IGW motion on the PV field, which is representative of a conserved, vertically stratified tracer. An important implication of such isentropic transport in general is that IGWs may directly cause stratosphere–troposphere exchange. If the IGW breaking level occurred a few kilometers lower than seen in Fig. 12, the quasi-horizontal motions would involve the tropopause [commonly defined by the 2-PVU (potential vorticity units) contour]. For instance, if the IGW motions of Fig. 11 were to occur 100 K lower at approximately 330 K where the isentropic surface intersected a sloping tropopause, then the 2-PVU contour would be involved in lateral meanders. The corresponding version of Fig. 11 would then show filaments of stratospheric air being advected into the higher PV stratosphere and filaments of high PV stratospheric air similarly being advected into the troposphere. The alternating filaments of stratospheric and tropospheric air appear like interleaved fingers on an isentropic surface. Much of this exchange would be irreversible, given the rapid straining of features by the background wind field. There is observational evidence for stratosphere–troposphere exchange by waves of short vertical wavelength that produce interleaving laminae of tropospheric and stratospheric air at and above the tropospheric jet (Shapiro et al. 1980). Danielsen et al. (1991) and Kritz et al. (1991) show evidence of such filaments mixing air along isentropes in the stratosphere above the jet stream.

Aside from tracer transport, IGWs are important in that they constitute a subsynoptic signal having large amplitude in the vicinity of jet streaks and jet stream exit regions. Van Tuyl and Young (1982) emphasized this aspect of IGWs and the importance of not suppressing them during the process of assimilating new data into numerical forecast models. IGWs can also play a role in organizing convection and precipitation. Koch and Dorian (1988) observed IGWs from a jet streak exit region, organizing a sequence of thunderstorms as they propagated through a region of weak conditional convective stability.

![Fig. 11. Isentropic potential vorticity on the 430-K surface on day 11. Contour interval is 1 PVU (≡ 10⁻⁶ m²s⁻¹K kg⁻¹). This isentropic surface is near 90 mb at the location of the IGW packet.](image-url)
Lastly, the upward momentum flux due to IGWs is considered. Given the widespread nature of IGWs and their preferred location in jetstreams, it might be thought that their upward flux of momentum could be important. However, the integrated momentum flux is proportional to energy times a factor $\delta^{1/2} = \sqrt{1 - (f/\bar{\omega})^2}$, showing that smaller scale high-frequency gravity waves are more efficient in fluxing momentum vertically than IGWs (Fritts and van Zandt 1993). (The factor $\delta$, in their formula is extraneous to the momentum flux and may be safely omitted.) It is of interest to note that the Eliassen–Palm flux approaches zero more rapidly as $\bar{\omega} \to f$, by an additional factor of $\delta$. At the center of the wavepacket on day 11, $u'w' \approx -0.08 \text{ m}^2 \text{ s}^{-2}$; a time- or area-average flux would be even smaller. Both are small numbers compared to the momentum flux near the tropopause estimated by Fritts and van Zandt.

IGWs could, nevertheless, have some impact on momentum transport by modulation of higher-frequency gravity waves (Broutman and Young 1986). Such waves might be absorbed at IGW-induced critical levels or, conversely, escape to higher altitudes. Partial reflection due to steep gradients created by IGW mixing is another possibility. Neither would be very important in the simulation where IGWs account for a small fraction of the total wind.

5. Conclusions

The excitation and propagation of inertia–gravity waves (IGWs) generated by an unstable baroclinic wave was examined with a high-resolution 3D nonlinear numerical model. IGWs arose spontaneously as the tropospheric jet stream was distorted by baroclinic instability and strong parcel accelerations took place, primarily in the jet stream exit region of the upper troposphere. Subsequent propagation of IGWs occurred in regions of strong windspeed, in the tropospheric and stratospheric jets, and in a cutoff low formed during the baroclinic lifecycle. IGWs on the flanks of these jets were rotated inward by differential advection and subsequently absorbed by the model’s hyperdiffusion. This was the principal mechanism of IGW absorption in the model, the results being insensitive to vertical resolution. IGWs in reality are thought to break via Kelvin–Helmholtz (KH) instability, a process unresolved in the 3D model. Although the model’s absorption of IGWs at the sidewalls of the jet is artificial, IGW propagation was for the most part confined to regions of $\bar{\omega} > f$. Only a few IGWs were able to penetrate the middle stratosphere, due to weak winds or an unfavorable alignment of wavevector with respect to the mean flow (except beneath the polar night jet).

IGWs are important both as a synoptic signal in the jet stream that may influence subsequent tropospheric
developments and as a source of isentropic or cross-isentropic mixing in the lower stratosphere. Our results demonstrated for the first time numerically a significant isentropic displacement of potential vorticity isopleths due to IGWs above the tropopause. Since conditions for IGW propagation are favorable within a jet, a region of strong isentropic PV gradient, it is likely that inertia–gravity waves affect the permeability of the lower stratospheric vortex and may in some instances lead to stratosphere–troposphere exchange, particularly if IGW-induced parcel displacements are large and irreversible. Although IGW breaking may help ensure that the mixing is permanent, it is interesting to note from the simulations that straining by the large-scale flow, without IGW breaking, was sufficient to guarantee irreversible mixing of PV.

Observations of tracer laminae in the lower stratosphere sometimes suggest a concurrent IGW signal, while others do not; it is possible that “fossil” laminae were due to an earlier IGW event or, perhaps, the baroclinic instability itself. The large-scale shear distends tracer isopleths into many thin filaments (e.g., Plumb et al. 1994). It will be interesting to compare mixing rates with and without IGWs to assess their role in isentropic transport.

Considerably higher vertical (and horizontal) resolution is required to determine the mechanisms of IGW breaking and the role of IGWs in cross-isentropic transport. This issue is important in the lower stratosphere where other mechanisms of cross-isentropic transport are quite weak. The chemistry of this region is sensitive to the rate at which constituents are diffused vertically and exchanged between the troposphere and stratosphere.

IGW-induced mixing is expected to vary as a function of location and season. The climatology of IGWs and their relation to tropospheric sources is essentially unknown. If progress is to made, it will begin with an analysis of IGWs in relation to the synoptic-scale flow at locations where continuous radar or lidar observations of high-frequency motions are routinely available. Relationships established locally may then, as a first approximation, be extrapolated to global models.

Acknowledgments. Conversations with Kaoru Sato and Dick Reed are gratefully acknowledged. This research was supported by National Aeronautics and Space Administration, Contract NAS1-19563, and by Air Force Office of Scientific Research, Contract F49620-92-C-0033.

REFERENCES


Sato, K., 1993: Small-scale wind disturbances observed by the MU radar during the passage of Typhoon Kelly. J. Atmos. Sci., 50, 518–537.


