Royal Meteorological Society

The interaction between synoptic-scale balanced flow and a finite-amplitude mesoscale wave field throughout all atmospheric layers: weak and moderately strong stratification

U. Achatz,[a]* B. Ribstein,[a] F. Senf[b] and R. Klein[c]

[a]*Institut für Atmosphäre und Umwelt, Goethe-Universität Frankfurt, Germany*
[b]*Leibniz-Institut für Troposphärenforschung, Leipzig, Germany*
[c]*Institut für Mathematik, Freie Universität, Berlin, Germany*

*Correspondence to: U. Achatz, Institut für Atmosphäre und Umwelt, Goethe-Universität Frankfurt, Altenhöferallee 1, D-60438 Frankfurt am Main, Germany. E-mail: achatz@iau.uni-frankfurt.de

The interaction between locally monochromatic finite-amplitude mesoscale waves, their nonlinearly induced higher harmonics, and a synoptic-scale flow is reconsidered, both in the tropospheric regime of weak stratification and in the stratospheric regime of moderately strong stratification. A review of the basic assumptions of quasi-geostrophic theory on an f-plane yields all synoptic scales in terms of a minimal number of natural variables, i.e. two out of the speed of sound, gravitational acceleration and Coriolis parameter. The wave scaling is defined so that all spatial and temporal scales are shorter by one order in the Rossby number, and by assuming their buoyancy field to be close to static instability. WKB theory is applied, with the Rossby number as scale separation parameter, combined with a systematic Rossby-number expansion of all fields. Classic results for synoptic-scale-flow balances and inertia-gravity-wave (IGW) dynamics are recovered. These are supplemented by explicit expressions for the interaction between mesoscale geostrophic modes (GMs), a possibly somewhat overlooked agent of horizontal coupling in the atmosphere, and the synoptic-scale flow. It is shown that IGW higher harmonics are slaved to the basic IGW, and that their amplitude is one order of magnitude smaller than the basic-wave amplitude. GM higher harmonics are not that weak and they are in intense nonlinear interaction between themselves and the basic GM. Compressible dynamics plays a significant role in the stratospheric stratification regime, where anelastic theory would yield insufficient results. Supplementing classic derivations, it is moreover shown that, in the absence of mesoscale waves, quasi-geostrophic theory holds also in the stratospheric stratification regime.

Key Words: gravity waves; geostrophic flow; mesoscale; wave–mean flow interaction; parametrization; wave action; enstrophy

Received 5 June 2016; Revised 11 September 2016; Accepted 12 September 2016; Published online in Wiley Online Library

1. Introduction

Mesoscale waves and their interaction with large-scale flow are an important problem of atmospheric dynamics. The significant contribution of inertia-gravity waves (IGWs) to the mesoscale dynamics of the atmosphere is undisputed (e.g. Fritts and Alexander, 2003; Kim *et al.*, 2003; Alexander *et al.*, 2010). They are radiated by various processes, often propagate over large distances, and finally break, unless dissipated by molecular diffusion and viscosity. Thereby and by other nonlinear interactions, they exert an impact on the momentum and energy budget of the large-scale flow. Corresponding effects cannot be neglected, either in weather prediction or in climate simulations. Often they must be parametrized because time- and length-scales of most parts of the IGW spectrum are too small to be resolved explicitly. Gravity-wave parametrizations were proposed, e.g. by Lindzen (1981), Holton (1982), Medvedev and Klaassen (1995), Hines (1997), and Alexander and Dunkerton (1999). Many of them are based on Wentzel–Kramers–Brillouin (WKB) theory (Bretherton, 1966; Grimshaw, 1975a,b; Müller, 1976). This approach assumes a small variation of the wave properties frequency, wave number, and amplitude over a wavelength and a period. In its most general form, it leads to a closed system of equations describing the propagation of frequency and wave number along rays, usually the conservative transport of wave action, and the interaction with the large-scale flow. However, the above-mentioned IGW parametrizations all take a single-column perspective where horizontal inhomogeneities of the large-scale flow are neglected as well as horizontal IGW propagation. Moreover, they assume instantaneous steady-state

IGW amplitudes representing an equilibrium, as would result after some adjustment time from a steady lower-boundary IGW source and a steady-state large-scale flow. These approximations allow the derivation of a wave-dissipation or non-acceleration paradigm, stating that changes of the large-scale flow can occur only when the IGWs dissipate, e.g. by wave breaking. Crucial progress in our understanding of the IGW–mean-flow interaction was facilitated by the development of Generalized Lagrangian-Mean (GLM) theory by Andrews and McIntyre (1978a,b), holding at arbitrary IGW amplitudes. Based on this theory Bühler and McIntyre (1998, 2003, 2005) have analyzed the wave-dissipation paradigm in detail. They show that dropping the single-column and steady-state assumptions leads to significant modifications. The wave-dissipation paradigm does not hold under these conditions, as the horizontal refraction of IGWs from large-scale-flow inhomogeneities goes along with substantial large-scale-flow accelerations. Moreover, it is the large-scale potential vorticity (PV) that is affected primarily, and corresponding accelerations are determined from PV inversion. Therefore these accelerations typically do not occur most strongly directly where the wave refraction takes place, or rather where the IGW forcing of synoptic-scale PV is largest. This is the so-called remote-recoil effect (Bühler and McIntyre, 2003), first demonstrated by Bretherton (1969). Studies of the interaction between IGWs and solar tides using a general WKB IGW model without single-column and steady-state approximations (Senf and Achatz, 2011; Ribstein *et al.*, 2015) indeed show that these approximations lead to a significant overestimation of the IGW-flux convergences, and hence to an incorrect estimation of tidal amplitudes.

However, as general as the GLM results are, with respect to IGW impacts on the large-scale flow, it remains difficult to directly implement them into weather forecast and climate models. These are formulated in an Eulerian perspective, whereas GLM assumes the resolved flow to be a Lagrangian mean. Moreover, Langrangian-mean results often stress issues around PV conservation and the related prognostic equation, whereas the practitioner is rather interested in explicit terms by which the standard prognostic equations for momentum and thermodynamics can be supplemented. Assuming low wave amplitudes, typically in terms of the ratio of displacement amplitude over wavelength, Andrews and McIntyre (1978b); Bühler and McIntyre (1998, 2003, 2005) and Bühler (2009) transform the GLM results in numerous examples to Eulerian representations, in shallow-water or Boussinesq dynamics, or assuming the large-scale flow to satisfy, in the absence of IGWs, quasi-geostrophic dynamics. However, they do not give corresponding results for general compressible dynamics. Vertical displacement amplitudes below the vertical wavelength also imply IGWs significantly below the static instability or overturning threshold, while waves of finite amplitudes are worth consideration as well. As low-amplitude theories rely on being able to use the wave amplitude as a small expansion parameter, it is not clear that their results can readily be used at finite amplitudes as well. Another issue is that classic quasi-geostrophic theory (Charney, 1948; Pedlosky, 1987) assumes the atmosphere to be as weakly stratified as in the troposphere where the pressure scale height H_p is about an order of magnitude less than the potential-temperature scale height H_θ. As pointed out by Klein *et al.* (2010), the ratio H_p/H_θ decides how the highest possible internal-wave frequency relates to a typical acoustic frequency. The larger it is, the more care is advisable in the use of sound-proof models, e.g. Bousinesq or anelastic models that are popular in this field. Stronger stratification, e.g. as in the stratosphere, does not seem to be have been fully considered. Zeitlin *et al.* (2003) indicate a derivation of quasi-geostrophic theory with strong stratification, however within Boussinesq theory.

Therefore, supplementary approaches remain interesting. Multi-scale asymptotics of the general compressible equations is such an approach. To the best of our knowledge, Grimshaw (1975b) first used this technique in a classic paper to analyze the IGW–mean-flow interaction in a rotating atmosphere, focusing on non-hydrostatic IGWs with comparable horizontal and vertical scales. He assumes equal scale heights for pressure, density, and entropy, as occur in the stratosphere. However, in the treatment of the IGW impact on the synoptic-scale flow, he switches to a Lagrangian-mean approach, and derives a conservation equation for a total PV consisting of quasi-geostrophic PV and a wave contribution. A complete treatment within the Eulerian perspective, and a corresponding link to quasi-geostrophic theory in an atmosphere with moderately strong stratification is not given. Finally, he does assume non-hydrostatic scaling for the waves, so that an application of his results to hydrostatic IGWs is not obviously possible. As a consequence of his scaling, e.g. the Coriolis frequency is assumed to be of the same order as the Brunt–Väisälä frequency, whereas in the atmosphere they are two orders of magnitude apart. Another related study is the one by Plougonven and Zhang (2007), where the interaction between synoptic-scale flow and IGWs is investigated in similar scaling regimes. However, they assume small wave amplitudes, and they do not take the step towards an efficient wave representation by slowly varying wavenumbers and amplitudes. Just as Yasuda *et al.* (2015a,b), that work rather addresses the problem of IGW emission by balanced flow, as reviewed by Plougonven and Zhang (2014).

Moreover, though IGWs are one possibility for mesoscale waves, there are also vortical geostrophic modes (GMs) which can contribute. In addition to IGWs, they arise as natural modes in the analysis of linear dynamics (e.g. Borchert *et al.*, 2014) and together with IGWs they form a complete modal basis of the part of the flow not attributable to accoustic modes. As shown below, they also constitute the mesoscale part of a flow described by quasi-geostrophic theory. In the soundproof approximation one can thus see the total dynamics as an interaction between synoptic-scale flow and mesoscale IGWs and GMs. The latter have been argued to be generated, e.g. by convective events (Gage, 1979; Lilly, 1983; Vallis *et al.*, 1997) and to represent the development of fronts at the edge of synoptic-scale vortices at the top of the troposphere (Tulloch and Smith, 2006). Beyond this they play a fundamental role in geostrophic adjustment (Rossby, 1938) and spontaneous imbalance (e.g. Plougonven and Zhang, 2014) where they represent the mesoscale part of the flow not radiated away in the form of IGWs. The study by Callies *et al.* (2014) indicates that IGWs dominate the mesoscale spectrum in the upper troposphere. However, the respective role of GMs and IGWs in horizontal coupling of synoptic-scale flows is unclear, so that the former still might deserve some attention. We are not aware of an analysis that systematically analyzes the GM interaction with a large-scale flow within the general compressible framework, and that develops a model for subgrid-scale GMs that can be used as parametrization in simulations that do not resolve the mesoscales. There is an extensive literature on interactions between synoptic-scale Rossby waves and planetary-scale mean flows, using quasi-geostrophic theory. An overview is given, e.g. by Vallis (2006). Most of it deals with zonally symmetric mean flows, but zonally inhomogeneous flows have also been discussed (Plumb, 1986). However, as detailed below and summarized in Table 2, mesoscale GMs are not in the low-Rossby-number regime. Hence it is not clear whether quasi-geostrophic theories can be used for them. In addition, differences could possibly be due to the different scales involved. It is not obvious that planetary-synoptic versus snyoptic-mesoscale interactions follow the same dynamics, even within quasi-geostrophic theory.

In summary, there seems to be room for a reconsideration of the interaction between a synoptic-scale flow and a mesoscale wave field of a locally monochromatic basic wave and its nonlinearly induced higher harmonics, hydrostatic or non-hydrostatic,

(i) holding at finite wave amplitudes,
(ii) derived from the compressible equations,
(iii) holding in all interesting stratification regimes, and
(iv) including the mean-flow interaction with GMs.

This is the plan of the work described here. It is an extension of the work of Grimshaw (1975b), but also of previous steps by some authors (Achatz *et al.*, 2010; Klein, 2011), using multi-scale asymptotic theory, where a finite-amplitude WKB theory for a non-rotating atmosphere has been derived in a particular distinguished limit of the governing equations. The prediction by that theory of weak higher harmonics, predominantly forced by large-scale gradients in the gravity-wave fluxes, has been validated numerically by Rieper *et al.* (2013). For the sake of better readability, results previously obtained by others, especially Grimshaw (1975b), are not just stated but re-derived, so as to provide a complete picture.

The article is structured as follows. In section 2 we identify the appropriate scales for our problem. These are used to non-dimensionalize the equations of motion in section 3, where the WKB ansatz is introduced as well, allowing for a basic wave and all its nonlinearly induced higher harmonics. Leading-order results of the asymptotic analysis are derived in section 4. These include the relevant dispersion and polarization relations as well as the eikonal ray-tracing equations. The next-order equations are derived in section 5, which are used in section 6 for the derivation of the amplitude equations for both wave modes. These are an IGW wave-action conservation equation, and potential-enstrophy equations for all GM harmonics. It is also shown that, due to their dispersion relation IGWs are dominated to leading order by the basic wave, whereas in a GM solution all higher harmonics contribute to the same order. The leading-order IGW harmonics are found to be slaved to the basic wave. This is followed by an analysis of the wave impact on the large-scale flow in section 7. Effectively the PV of the synoptic-scale flow is found to satisfy a quasi-geostrophic conservation equation, supplemented by a forcing due to the vertical curl of an Eliassen–Palm flux convergence vector. The most essential results are summarized in dimensional form in section 8. We conclude with a discussion in section 9.

2. Scaling for synoptic-scale flow and for small-scale waves

We assume inviscid and continuously stratified dynamics on an f-plane (e.g. Durran, 1989), with Coriolis parameter f, without external sources or sinks:

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} + f\mathbf{e}_z \times \mathbf{u} = -c_p\theta\nabla_h\pi, \tag{1}$$

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -c_p\theta\frac{\partial\pi}{\partial z} - g, \tag{2}$$

$$\frac{\mathrm{D}\theta}{\mathrm{D}t} = 0, \tag{3}$$

$$\frac{\mathrm{D}\pi}{\mathrm{D}t} + \frac{R}{c_v}\pi\nabla\cdot\mathbf{v} = 0, \tag{4}$$

where $\mathbf{u}$ and w are the horizontal and vertical components of the total wind $\mathbf{v}$, respectively. c_p and $c_v = c_p - R$ are the specific heat capacities at constant pressure and volume, respectively, with R the ideal gas constant of dry air. θ is potential temperature, π the Exner pressure, g the gravitational acceleration, and $\mathbf{e}_z$ is the vertical unit vector. Within this setting we consider a superposition of an exclusively altitude-dependent hydrostatic reference atmosphere at rest (with tropospheric or stratospheric stratification), a rather general synoptic-scale flow, and a locally monochromatic small-scale wave field.

2.1. *Reference-atmosphere scaling and synoptic scaling within quasi-geostrophic theory*

As a first step we review the synoptic scaling which quasi-geostrophic theory is built on (e.g. Pedlosky, 1987). Synoptic-scale flow is assumed to have typical horizontal and vertical length scales L_s and H_s. Velocity scales for horizontal and vertical wind are U_s and W_s. Density fluctuations are assumed sufficiently small to allow the estimate

$$W_s = \frac{H_s}{L_s}U_s. \tag{5}$$

The synoptic time-scale T_s matches the advective time-scale

$$T_s = \frac{L_s}{U_s} = \frac{H_s}{W_s}, \tag{6}$$

and is assumed much longer than the inertial time-scale, so that the Rossby number, ε, is small:

$$\varepsilon = \frac{U_s}{fL_s} = \frac{1}{fT_s} = \mathcal{O}(10^{-1}) \ll 1. \tag{7}$$

It is assumed that the vertical synoptic length-scale is comparable to a typical pressure scale height H_p (with $R\overline{T}/g = \mathcal{O}(H_p)$), i.e.

$$\frac{H_s}{H_p} = \mathcal{O}(1) \tag{8}$$

where $\overline{T}(z)$ is the temperature of the reference atmosphere. The horizontal synoptic length-scale is set by the reference-atmosphere stratification. Based on observation and also consistent with baroclinic instability theory (Charney, 1947; Eady, 1949), it is assumed to be of the same order as the internal Rossby deformation radius $L_{di} = NH_p/f$, i.e.

$$\frac{L_s}{L_{di}} = \mathcal{O}(1), \tag{9}$$

where the Brunt–Väisälä frequency $N = \sqrt{g/H_\theta}$ is here determined by a typical reference-atmosphere potential-temperature scale height H_θ, i.e. $\overline{\theta}/(\mathrm{d}\overline{\theta}/\mathrm{d}z) = \mathcal{O}(H_\theta)$. The stratification depends on the atmospheric layer. Assuming a constant temperature lapse rate $\Gamma = -\mathrm{d}\overline{T}/\mathrm{d}z$, one finds for a hydrostatic reference atmosphere

$$\frac{H_p}{H_\theta} = \frac{R}{c_p}\left(1 - \frac{\Gamma}{g/c_p}\right). \tag{10}$$

In the weakly stratified troposphere, a characteristic lapse rate is $\Gamma \sim 6.5\,\mathrm{K\,km^{-1}}$ and hence $H_p/H_\theta \sim 0.1 = \mathcal{O}(\varepsilon)$ and the midlatitude ratio between squared inertia and stratification is $f^2/N^2 = \mathcal{O}(\varepsilon^4)$. However, here we are also interested in the more strongly stratified case, more characteristic for the stratosphere, where $\Gamma \sim -5\,\mathrm{K\,km^{-1}}$ and therefore $H_p/H_\theta \sim 0.4 = \mathcal{O}(1)$, so that $f^2/N^2 = f^2H_\theta/g = \mathcal{O}(\varepsilon^5)$, assuming an equal pressure scale height in troposphere and stratosphere. This can be summarized by

$$\frac{H_p}{H_\theta} = \mathcal{O}(\varepsilon^\alpha), \tag{11}$$

$$\frac{f^2}{N^2} = \mathcal{O}(\varepsilon^{5-\alpha}), \tag{12}$$

where α is either 1 (weak stratification) or 0 (moderately strong stratification). Hence, with the external Rosssby deformation radius $L_d = \sqrt{gH_p}/f$,

$$\frac{{L_s}^2}{{L_d}^2} = \mathcal{O}\left(\frac{{L_{di}}^2}{{L_d}^2}\right) = \mathcal{O}\left(\frac{H_p}{H_\theta}\right) = \mathcal{O}(\varepsilon^\alpha), \tag{13}$$

$$\frac{H_s}{L_s} = \mathcal{O}\left(\frac{H_p}{L_{di}}\right) = \mathcal{O}\left(\frac{f}{N}\right) = \mathcal{O}(\varepsilon^{(5-\alpha)/2}). \tag{14}$$

The classic derivation of quasi-geostrophic theory (e.g. Pedlosky, 1987) assumes weak stratification ($\alpha = 1$). A result below will be that it also holds for moderately strong stratification ($\alpha = 0$). The estimates above are thus consistent for both cases.

For the scaling of the dynamic variables, we then observe that geostrophic equilibrium implies for the synoptic-scale Exner-pressure fluctuations $\pi' = \pi - \overline{\pi}$, together with the order-of-magnitude equality of vertical length-scale and the pressure scale height, and (13), that

$$\frac{\pi'}{\overline{\pi}} = \mathcal{O}\left(\frac{fU_sL_s}{c_p\overline{\theta}\overline{\pi}}\right) = \mathcal{O}\left(\frac{fU_sL_s}{c_p\overline{T}}\right) = \mathcal{O}\left(\frac{R}{c_p}\frac{fU_sL_s}{gH_p}\right)$$
$$= \mathcal{O}\left(\frac{R}{c_p}\frac{L_s^2}{L_d^2}\varepsilon\right) = \mathcal{O}\left(\varepsilon^{1+\alpha}\right). \tag{15}$$

Likewise, hydrostatic equilibrium yields for the synoptic-scale potential-temperature fluctuations

$$\frac{\theta'}{\overline{\theta}} = \mathcal{O}\left(\frac{\partial\pi'/\partial z}{\partial\overline{\pi}/\partial z}\right) = \mathcal{O}\left(\frac{\pi'/H_p}{g/c_p\overline{\theta}}\right) = \mathcal{O}\left(\frac{c_p}{R}\frac{\pi'}{\overline{\pi}}\right) = \mathcal{O}(\varepsilon^{1+\alpha}). \tag{16}$$

Moreover, geostrophic equilibrium implies

$$U_s = \mathcal{O}\left(\frac{c_p\overline{\theta}\pi'}{L_sf}\right) = \mathcal{O}\left(\frac{c_p}{R}\frac{\pi'}{\overline{\pi}}\frac{R\overline{T}}{L_sf}\right) = \mathcal{O}\left(\frac{c_p}{R}\frac{\pi'}{\overline{\pi}}\frac{L_d}{L_s}\sqrt{R\overline{T}}\right), \tag{17}$$

so that the appropriate horizontal-velocity scale is

$$U_s = \varepsilon^{(2+\alpha)/2}\sqrt{RT_{00}}, \tag{18}$$

where T_{00} is a typical mid-altitude value of tropospheric or stratospheric temperature. In other words, the Mach number is $Ma = U_s/\sqrt{RT_{00}} = \mathcal{O}[\varepsilon^{(2+\alpha)/2}]$. Using the definition (7) of the Rossby number, the aspect ratio (14), and (5), we also obtain

$$W_s = \varepsilon^{7/2}\sqrt{RT_{00}}, \tag{19}$$
$$L_s = \varepsilon^{\alpha/2}\sqrt{RT_{00}}/f, \tag{20}$$
$$H_s = \varepsilon^{5/2}\sqrt{RT_{00}}/f, \tag{21}$$

while we remember from (7) that

$$T_s = \varepsilon^{-1}/f. \tag{22}$$

We remark that the order-of-magnitude equality of H_s and the pressure scale height H_p then also implies the scaling

$$g = \varepsilon^{-5/2}f\sqrt{RT_{00}}. \tag{23}$$

Certainly one could as well get from this $\sqrt{RT_{00}}$, roughly the speed of sound, in terms of g and f and express all scales in terms of those constants. This would not change the results below. A summary of relevant flow numbers is given in Table 1.

2.2. *Scaling of inertia-gravity waves close to breaking*

Now consider small-scale inertia-gravity waves (IGWs) such that a typical vertical wavenumber m and corresponding vertical scale $H_w = 1/m$ obey

$$H_w = \varepsilon H_s = \varepsilon^{7/2}\sqrt{RT_{00}}/f. \tag{24}$$

The horizontal length-scale is chosen so that both stratification and rotation affect the IGW intrinsic frequency $\widehat{\omega}$. With the Boussinesq dispersion relation as an indicator (e.g. Sutherland, 2010), i.e.

$$\widehat{\omega}^2 = \frac{f^2m^2 + N^2(k^2+l^2)}{k^2+l^2+m^2}, \tag{25}$$

Table 1. Relevant flow numbers for the reference atmosphere and the synoptic-scale flow.

Rossby number	$Ro = \frac{U_s}{fL_s} = \varepsilon$
Froude number	$Fr = \frac{U_s}{NH_s} = \mathcal{O}(\varepsilon)$
Internal Burger number	$Bu_i = \left(\frac{NH_s/f}{L_s}\right)^2 = \left(\frac{Ro}{Fr}\right)^2 = \mathcal{O}(1)$
External Burger number	$Bu = \left(\frac{\sqrt{gH_s}/f}{L_s}\right)^2 = \mathcal{O}(\varepsilon^{-\alpha})$
Mach number	$Ma = \frac{U_s}{\sqrt{RT_{00}}} = \mathcal{O}(\varepsilon^{(2+\alpha)/2})$
Aspect ratio	$a_B = \frac{H_s}{L_s} = \mathcal{O}(\varepsilon^{(5-\alpha)/2})$

where k and l are typical horizontal wave numbers in x- and y-directions, the corresponding horizontal length-scale is $L_w = 1/\sqrt{k^2+l^2}$ so that equal impact of rotation and stratification implies, using (14),

$$\frac{H_w^2}{L_w^2} = \mathcal{O}\left(\frac{f^2}{N^2}\right) = \mathcal{O}(\varepsilon^{5-\alpha}), \tag{26}$$

and thus

$$L_w = \varepsilon^{-(5-\alpha)/2}H_w = \varepsilon^{(2+\alpha)/2}\sqrt{RT_{00}}/f = \varepsilon L_s. \tag{27}$$

Likewise, we deduce that the IGW time-scale $T_w = 1/\widehat{\omega}$ is

$$T_w = 1/f = \varepsilon T_s. \tag{28}$$

This is consistent with the assumption of an IGW field influenced by the Coriolis force. As is shown below, it is also in agreement with the IGW time-scaling to be obtained from the Doppler term, and hence also the absolute frequency.

The scaling of the dynamic variables is chosen so that it represents an IGW close to breaking by overturning of potential-temperature surfaces. This point of static instability is reached as soon as the wave has an amplitude allowing local negative vertical derivatives of total potential temperature,

$$\frac{\partial}{\partial z}\left\{\overline{\theta} + \Re\left(\theta_w\,e^{i(\mathbf{k}\cdot\mathbf{x}-\omega\mathbf{t})}\right)\right\} < 0, \tag{29}$$

where θ_w is the wave's potential-temperature amplitude. At the point of marginal stability it satisfies

$$\frac{\theta_w}{\overline{\theta}} = \mathcal{O}\left(\frac{1}{m}\frac{1}{\overline{\theta}}\frac{d\overline{\theta}}{dz}\right) = \mathcal{O}\left(\frac{H_w}{H_\theta}\right) = \mathcal{O}(\varepsilon^{1+\alpha}), \tag{30}$$

as follows from (11), the order-of-magnitude equality of H_s and H_p, and (24). This also implies an IGW buoyancy scaling

$$B_w = g\frac{\theta_w}{\overline{\theta}} = \mathcal{O}(N^2H_w). \tag{31}$$

Referring to buoyancy dynamics, one can see that the wave amplitudes considered are finite, i.e. the vertical-displacement amplitude B_w/N^2 is of the same magnitude as the vertical length scale of the waves. To obtain from this the horizontal wind-scale of the IGWs, we use the polarization relation between their horizontal wind amplitude U_w in the x-direction, for example, and buoyancy, as also derived further below (163), to estimate

$$U_w = \mathcal{O}\left(i\frac{k\widehat{\omega} + ilf}{mN^2}\frac{\widehat{\omega}^2 - N^2}{\widehat{\omega}^2 - f^2}B_w\right), \tag{32}$$

which can be simplified by obtaining from the dispersion relation

$$\frac{\widehat{\omega}^2 - N^2}{\widehat{\omega}^2 - f^2} = -\frac{m^2}{k^2+l^2} = \mathcal{O}\left(\frac{L_w^2}{H_w^2}\right). \tag{33}$$

Since both $k\widehat{\omega}$ and lf are $O\left(f/L_w\right)$, this shows, together with (31), that we can assume for the IGW horizontal wind-scale

$$U_w = fL_w = \varepsilon^{(2+\alpha)/2}\sqrt{RT_{00}} = U_s. \tag{34}$$

The horizontal wind-scales of synoptic-scale flow and IGW are thus assumed to be identical. Moreover, the IGW time-scale turns out also to be the advective time-scale, i.e. $T_w = L_w/U_w$. This also implies that the Doppler term in the IGW dispersion relation is in agreement with the IGW time scaling as well, as noted above. Likewise the vertical wind-scale is derived from the Boussinesq polarization relation between IGW vertical wind amplitude W_w and B_w ((164) below),

$$W_w = \mathcal{O}\left(\frac{i\hat{\omega}}{N^2}B_w\right). \tag{35}$$

Hence the IGW vertical-wind scale can be assumed to be

$$W_w = f\,H_w = \varepsilon^{7/2}\sqrt{RT_{00}} = W_s. \tag{36}$$

Also the vertical-wind scale can be assumed identical with that of the synoptic-scale flow. For an estimate of the Exner-pressure scaling in a marginally stable IGW, we use the corresponding polarization relation for the Exner-pressure IGW amplitude ((165) below), to estimate

$$\Pi_w = \mathcal{O}\left(\frac{i}{m}\frac{\widehat{\omega}^2 - N^2}{N^2}\frac{B_w}{c_p\overline{\theta}}\right) \tag{37}$$

Since $\widehat{\omega}^2 = \mathcal{O}(f^2) \ll N^2$, this implies together with (31), $\overline{\theta} = \mathcal{O}(T_{00})$, $N^2 = g/H_\theta$, and $H_w/H_p = \mathcal{O}(\varepsilon H_s/H_p) = \mathcal{O}(\varepsilon)$ that an appropriate scaling is

$$\Pi_w = \mathcal{O}\left(\varepsilon^{2+\alpha}\frac{R}{c_p}\right) = \mathcal{O}(\varepsilon^{2+\alpha}). \tag{38}$$

The IGW Exner-pressure fluctuations scale with Ma^2 as in incompressible flow and are extremely weak. As a consequence, sound waves are suppressed in this scaling regime.

We conclude this section by the remark that an analogous analysis of the polarization relations (163)–(165) below for the dynamic scaling of a GM close to breaking, generated by the processes named above and brought to large amplitudes by various nonlinear interactions, e.g. wave–mean flow interactions or harmonic–harmonic interactions as described below, yields exactly the same scaling as for IGWs, with the exception that its vertical wind vanishes. A summary of relevant flow numbers for the small-scale waves is given in Table 2. Notably, not only IGWs but also mesoscale GMs have a Rossby number that is not small, since both its horizontal-wind amplitude and the advecting synoptic-scale wind are of the same magnitude while the length-scale is shorter than the synoptic length-scale!

Table 2. Relevant flow numbers for the small-scale wave component.

Wave Rossby number	$Ro_w = \frac{U_w}{fL_w} = \mathcal{O}(1)$
Wave Froude number	$Fr_w = \frac{U_w}{NH_w} = \mathcal{O}(1)$
Internal wave Burger number	$Bu_{iw} = \left(\frac{NH_w/f}{L_w}\right)^2 = \left(\frac{Ro_w}{Fr_w}\right)^2 = \mathcal{O}(1) = Bu_i$
External wave Burger number	$Bu_w = \left(\frac{\sqrt{gH_w}/f}{L_w}\right)^2 = \mathcal{O}(\varepsilon^{-1-\alpha})$
Wave Mach number	$Ma_w = \frac{U_w}{\sqrt{RT_{00}}} = \mathcal{O}(\varepsilon^{(2+\alpha)/2}) = Ma$
Wave aspect ratio	$a_{B,w} = \frac{H_w}{L_w} = \mathcal{O}(\varepsilon^{(5-\alpha)/2}) = a_B$

3. Non-dimensional equations and WKB ansatz

3.1. Non-dimensionalization of the equations of motion

The IGW scaling defined above is now used to non-dimensionalize the equations of motion without friction, heating and heat conduction. In the basic equations (1)–(4), we replace

$$(\mathbf{u}, \mathbf{w}) \rightarrow (U_w\mathbf{u}, \mathbf{W}_w\mathbf{w}), \tag{39}$$

$$(x, y, z, t) \rightarrow \left[L_w(x, y), H_w z, T_w t\right], \tag{40}$$

$$(\theta, \pi) \rightarrow (T_{00}\theta, \pi), \tag{41}$$

$$f \rightarrow f f_0, \tag{42}$$

yielding the non-dimensional equations

$$\varepsilon^{2+\alpha}\left(\frac{D\mathbf{u}}{Dt} + f_0\mathbf{e}_z\times\mathbf{u}\right) = -\frac{c_p}{R}\theta\nabla_h\pi, \tag{43}$$

$$\varepsilon^7\frac{Dw}{Dt} = -\frac{c_p}{R}\theta\frac{\partial\pi}{\partial z} - \varepsilon, \tag{44}$$

$$\frac{D\theta}{Dt} = 0, \tag{45}$$

$$\frac{D\pi}{Dt} + \frac{R}{c_v}\pi\nabla\cdot\mathbf{v} = 0. \tag{46}$$

For later reference we also remark that the equation of state

$$\rho = \frac{p_{00}}{R\theta}\pi^{c_v/R} \tag{47}$$

becomes

$$\rho = \frac{\pi^{c_v/R}}{\theta} \tag{48}$$

if $\rho_{00} = p_{00}/RT_{00}$ is used to non-dimensionalize the density.

3.2. Multi-scale asymptotics and WKB ansatz

In the following we consider particular solutions of the compressible equations that are a superposition of a reference atmosphere at rest, a synoptic-scale flow, a locally monochromatic basic-wave field (IGW or GM), and its higher harmonics. The latter are added as they will inevitably be forced by nonlinear interactions. The length- and time-scales of the synoptic-scale flow are $(L_s, H_s, T_s) = (L_w/\varepsilon, H_w/\varepsilon, T_w/\varepsilon)$, which we express by letting the synoptic-scale fields depend on the compressed coordinates

$$(\mathbf{X}, T) = \varepsilon(\mathbf{x}, \mathbf{t})\,. \tag{49}$$

The reference atmosphere can be characterized as follows. Weak potential-temperature stratification, where $\alpha = 1$ and $H_w/H_\theta = \mathcal{O}(\varepsilon^2)$, can be encoded by letting $\overline{\theta} = \overline{\Theta}^{(0,1)} + \varepsilon\overline{\Theta}^{(1)}(Z)$, with $\overline{\Theta}^{(0,1)}$ a constant of order unity, and only $\overline{\Theta}^{(1)}$ depending on Z. Moderately strong stratification, where $\alpha = 0$ and $H_w/H_\theta = \mathcal{O}(\varepsilon)$, can be described by letting $\overline{\theta} = \overline{\Theta}^{(0,0)}(Z)$, where now $\overline{\Theta}^{(0,0)}$ depends on Z. In both cases, however, $H_w/H_p = \mathcal{O}(\varepsilon)$ so that the reference-atmosphere Exner-pressure field has a leading-order term depending on Z. We summarize this by letting

$$\overline{\theta} = \sum_{j=0}^{\alpha}\varepsilon^j\overline{\Theta}^{(j)}(Z)$$

$$\text{with}\quad \overline{\Theta}^{(0)}(Z) = \alpha\overline{\Theta}^{(0,1)} + (1-\alpha)\overline{\Theta}^{(0,0)}(Z), \tag{50}$$

$$\overline{\pi} = \sum_{j=0}^{\alpha}\varepsilon^j\overline{\Pi}^{(j)}(Z)\,. \tag{51}$$

This way the leading-order term of the potential temperature depends on Z only in the case with moderately strong stratification, where $\alpha = 0$, and the stratification is

$$\frac{\partial \overline{\theta}}{\partial Z} = \varepsilon^{\alpha} \frac{\partial \overline{\Theta}^{(\alpha)}}{\partial Z}. \tag{52}$$

The wave field is assumed to have the following scaling properties:
(i) Wavelengths and periods of the basic wave are characterized by the wave scales assumed above.
(ii) They vary in space and time, in response to the interaction with the synoptic-scale wind. The space- and time-scales of these variations are therefore the synoptic scales.
(iii) Also the wave amplitude has a corresponding weak spatial and temporal dependence. Close to but below the breaking amplitude, this is a realistic assumption, as the non-smoothness of wave amplitudes arises as a result of a turbulent breaking process, but not before (e.g. Achatz, 2007). However, even at the breaking amplitude, Bölöni *et al.* (2016) show for the non-rotating case that WKB theory can reproduce the fully nonlinear dynamics surprisingly well.
(iv) The basic wave is supplemented by all its higher harmonics. In the case of the horizontal wind in the x-direction, for example, this is expressed via

$$u(\mathbf{x}, \mathbf{t}) = \Re \sum_{\beta=1}^{\infty} \mathbf{U}_\beta(\mathbf{X}, \mathbf{T}) e^{i\beta\phi(\mathbf{X},\mathbf{T})/\varepsilon} \tag{53}$$

with amplitudes U_β and phases $\beta\phi/\varepsilon$. The basic wave is represented by $\beta = 1$, while $\beta \geq 2$ indicates the higher harmonics. The time derivative and spatial gradient of the basic-wave phase define the local frequency ω and local wave number $\mathbf{k}$, respectively.

$$\omega(\mathbf{X}, T) = -\frac{\partial}{\partial t}\left(\frac{\phi}{\varepsilon}\right) = -\frac{\partial \phi}{\partial T}$$
$$\mathbf{k}(\mathbf{X}, T) = \nabla_{\mathbf{x}}\left(\frac{\phi}{\varepsilon}\right) = \nabla_{\mathbf{X}}\phi$$

In accordance with the scaling analysis above, using (15), (16), (30), (34), (36), and (38), all fields are now expanded in terms of $\varepsilon \ll 1$, setting

$$\mathbf{v} = \sum_{j=0}^{\infty} \varepsilon^j \mathbf{V}_0^{(j)}(\mathbf{X}, T) + \Re \sum_{\beta=1}^{\infty}\sum_{j=0}^{\infty} \varepsilon^j \mathbf{V}_\beta^{(j)}(\mathbf{X}, T) e^{i\beta\phi(\mathbf{X},T)/\varepsilon}, \tag{54}$$

$$\theta = \sum_{j=0}^{\alpha} \varepsilon^j \overline{\Theta}^{(j)}(Z) + \varepsilon^{1+\alpha} \sum_{j=0}^{\infty} \varepsilon^j \Theta_0^{(j)}(\mathbf{X}, T) + \varepsilon^{1+\alpha} \Re \sum_{\beta=1}^{\infty}\sum_{j=0}^{\infty} \varepsilon^j \Theta_\beta^{(j)}(\mathbf{X}, T) e^{i\beta\phi(\mathbf{X},T)/\varepsilon}, \tag{55}$$

$$\pi = \sum_{j=0}^{\alpha} \varepsilon^j \overline{\Pi}^{(j)}(Z) + \varepsilon^{1+\alpha} \sum_{j=0}^{\infty} \varepsilon^j \Pi_0^{(j)}(\mathbf{X}, T) + \varepsilon^{2+\alpha} \Re \sum_{\beta=1}^{\infty}\sum_{j=0}^{\infty} \varepsilon^j \Pi_\beta^{(j)}(\mathbf{X}, T) e^{i\beta\phi(\mathbf{X},T)/\varepsilon}, \tag{56}$$

where $\overline{\Theta}^{(j)}$ and $\overline{\Pi}^{(j)}$ are due to the reference atmosphere; all terms proportional to the phase factors $\exp(i\beta\phi/\varepsilon)$ are contributions from the wave (subscript 1 for the basic wave, and $\beta \geq 2$ for its βth harmonic), and the rest constitute the synoptic-scale part (subscript 0). The equation of state then also implies

$$\rho = \sum_{j=0}^{\alpha} \varepsilon^j \overline{R}^{(j)}(Z) + \mathcal{O}(\varepsilon^{1+\alpha}), \tag{57}$$

where

$$\overline{R}^{(0)} = \frac{\overline{\Pi}^{(0)c_V/R}}{\overline{\Theta}^{(0)}} \equiv \frac{\overline{P}^{(0)}}{\overline{\Theta}^{(0)}}, \quad \overline{R}^{(1)} = \overline{R}^{(0)}\left(\frac{c_V}{R}\frac{\overline{\Pi}^{(1)}}{\overline{\Pi}^{(0)}} - \frac{\overline{\Theta}^{(1)}}{\overline{\Theta}^{(0)}}\right). \tag{58}$$

4. Leading-order results: equilibria, dispersion and polarization relations, eikonal equations

We now insert the expansions (54)–(56) into the non-dimensional equations (43)–(46), collect the leading-order terms and use these for first results. Hereby we discriminate between the respective wave parts, each proportional to the phase factor $\exp(i\beta\phi/\varepsilon)$, and mean-flow terms, where no phase factor appears.

4.1. *Leading orders of the equations of motion*

The leading-order terms of the Exner-pressure equation (46) are $\mathcal{O}(1)$. There is just a wave part

$$\overline{P}^{(0)} \Re \sum_{\beta=1}^{\infty} i\beta \mathbf{k} \cdot \mathbf{V}_\beta^{(0)} e^{i\beta\phi/\varepsilon} = 0, \tag{59}$$

yielding for all β

$$\mathbf{k} \cdot \mathbf{V}_\beta^{(0)} = 0. \tag{60}$$

The leading-order velocity amplitudes of the wave part are orthogonal to the local wavenumber vector. This solenoidality property of the wave velocity field helps to eliminate numerous nonlinear advection terms in the treatment below.

Next we turn to the entropy equation (45). One finds that the leading-order terms are $\mathcal{O}(\varepsilon^{1+\alpha})$. Within these, the mean-flow contributions yield

$$W_0^{(0)} \frac{d\overline{\Theta}^{(\alpha)}}{dZ} = 0, \tag{61}$$

or

$$W_0^{(0)} = 0. \tag{62}$$

This reproduces the well-known result that the leading-order synoptic-scale (geostrophically balanced, see below) wind has no vertical component. In the remaining wave contributions, we first eliminate the nonlinear term by (60), leaving us with the linear buoyancy equations

$$-i\beta\,\widehat{\omega} B_\beta^{(0)} + W_\beta^{(0)} {N_0}^2 = 0, \tag{63}$$

where

$$\widehat{\omega} = \omega - \mathbf{k} \cdot \mathbf{V}_0^{(0)} \tag{64}$$

is the non-dimensional intrinsic frequency,

$$B_\beta^{(0)} = \frac{\Theta_\beta^{(0)}}{\overline{\Theta}^{(0)}} \tag{65}$$

are the non-dimensional leading-order wave buoyancy amplitudes, and

$${N_0}^2 = \frac{1}{\overline{\Theta}^{(0)}} \frac{d\overline{\Theta}^{(\alpha)}}{dZ} \tag{66}$$

is the non-dimensional squared reference-atmosphere Brunt–Väisälä frequency.

The leading terms of $\mathcal{O}(\varepsilon)$ in the vertical momentum equation (44), and of $\mathcal{O}(\varepsilon^2)$ in the weakly stratified case ($\alpha = 1$), yield

$$\frac{d\overline{\Pi}^{(0)}}{dZ} = -\frac{R/c_p}{\overline{\Theta}^{(0)}}, \quad \text{and (if } \alpha = 1\text{)} \quad \frac{d\overline{\Pi}^{(1)}}{dZ} = \frac{R/c_p}{\overline{\Theta}^{(0)}} \frac{\overline{\Theta}^{(1)}}{\overline{\Theta}^{(0)}}. \tag{67}$$

This reflects hydrostatic equilibrium of the reference atmosphere at leading order.

The $\mathcal{O}(\varepsilon^{2+\alpha})$ terms of (44) again reflect hydrostatic balance in their wave part,

$$-B_\beta^{(0)} + \mathrm{i}\beta m \frac{c_p}{R}\overline{\Theta}^{(0)}\Pi_\beta^{(0)} = 0\,, \tag{68}$$

where m is the non-dimensional vertical wavenumber. The mean-flow part yields, using (62), and again (67),

$$\frac{\partial \Pi_0^{(0)}}{\partial Z} = \frac{R/c_p}{\overline{\Theta}^{(0)}}\left[\frac{\Theta_0^{(0)}}{\overline{\Theta}^{(0)}} - \alpha\left(\frac{\overline{\Theta}^{(\alpha)}}{\overline{\Theta}^{(0)}}\right)^2\right] \tag{69}$$

which expresses the hydrostatic equilibrium of the synoptic-scale flow. The second term in the brackets is neither horizontally nor time dependent, and hence does not have much relevance for the following.

Finally we analyze the horizontal momentum equation (43). There the leading order is $\mathcal{O}(\varepsilon^{2+\alpha})$. Once more the nonlinear term vanishes due to (60), leaving the wave contributions

$$-\mathrm{i}\beta\,\widehat{\omega}\mathbf{U}_\beta^{(0)} + f_0\mathbf{e}_z\times\mathbf{U}_\beta^{(0)} + \mathrm{i}\beta\mathbf{k}_h\frac{c_p}{R}\overline{\Theta}^{(0)}\Pi_\beta^{(0)} = 0, \tag{70}$$

where $\mathbf{U}_\beta^{(0)}$ is the horizontal component of $\mathbf{V}_\beta^{(0)}$, and $\mathbf{k}_h$ the horizontal component of the non-dimensional wavenumber. The mean-flow part reads

$$f_0\mathbf{e}_z\times\mathbf{U}_0^{(0)} = -\frac{c_p}{R}\overline{\Theta}^{(0)}\nabla_{X,h}\Pi_0^{(0)} \tag{71}$$

The latter expresses the geostrophic equilibrium of the synoptic-scale flow. Clearly the hydrostatic and geostrophic equilibrium are in agreement with the original expectations. We also point out that for none of the leading-order results we had to resort to weak wave amplitudes. The latter are indeed allowed to be close to the level of static instability, and it is *exclusively the scale separation, combined with the Boussinesq-type solenoidality of the wave velocities, that sorts out all the nonlinear terms.*

4.2. *Dispersion relation, leading-order wave amplitudes, and polarization relations*

The leading-order wave contributions (70), (68), (63)/N_0, and (60) can be summarized as

$$0 = \underbrace{\begin{pmatrix} -\mathrm{i}\beta\,\widehat{\omega} & -f_0 & 0 & 0 & \mathrm{i}\beta k \\ f_0 & -\mathrm{i}\beta\,\widehat{\omega} & 0 & 0 & \mathrm{i}\beta\,l \\ 0 & 0 & 0 & -N_0 & \mathrm{i}\beta m \\ 0 & 0 & N_0 & -\mathrm{i}\beta\,\widehat{\omega} & 0 \\ \mathrm{i}\beta k & \mathrm{i}\beta\,l & \mathrm{i}\beta m & 0 & 0 \end{pmatrix}}_{M_\beta(\beta\mathbf{k},\,\beta\widehat{\omega})}\mathbf{Z}_\beta^{(0)} \tag{72}$$

with

$$\mathbf{Z}_\beta^{(0)^t} = \left(U_\beta^{(0)}, V_\beta^{(0)}, W_\beta^{(0)}, B_\beta^{(0)}/N_0, \frac{c_p}{R}\overline{\Theta}^{(0)}\Pi_\beta^{(0)}\right). \tag{73}$$

Non-trivial basic-wave amplitudes require $\det(M_1) = 0$, yielding either

$$\widehat{\omega} = 0, \tag{74}$$

which is the GM solution, or

$$\widehat{\omega}^2 = \frac{{N_0}^2(k^2+l^2) + f_0^2 m^2}{m^2}, \tag{75}$$

which is the dispersion relation for hydrostatic IGWs. It might be worthwhile stressing that the GM is only balanced in the local and non-inertial reference frame of the synoptic-scale flow. In the global reference frame at rest, it oscillates at a high frequency, due to advection by the spatially and time-dependent synoptic-scale flow.

The structure of basic wave and higher harmonics of either the GM or the IGW is given by the null vector of M_β, using $\widehat{\omega}$ from (74) or (75). One obtains

$$\mathbf{U}_\beta^{(0)} = \frac{\beta^2\mathbf{k}_h\widehat{\omega} - \mathrm{i}f_0\mathbf{e}_z\times\beta\mathbf{k}_h}{\beta^2\widehat{\omega}^2 - f_0^2}\frac{B_\beta^{(0)}}{\mathrm{i}\beta m}, \tag{76}$$

$$W_\beta^{(0)} = \frac{\mathrm{i}\beta\widehat{\omega}}{N_0^2}B_\beta^{(0)}, \tag{77}$$

$$\frac{c_p}{R}\overline{\Theta}^{(0)}\Pi_\beta^{(0)} = \frac{B_\beta^{(0)}}{\mathrm{i}\beta m}. \tag{78}$$

However, a notable difference arises in the higher-harmonics wave amplitudes. In the IGW case, due to the dispersiveness of the dispersion relation, M_β is non-singular for $\beta \geq 2$. Hence the leading-order higher harmonics of an IGW basic wave vanish:

$$\mathbf{Z}_\beta^{(0)} = 0 \quad \text{for IGWs and } \beta \geq 2. \tag{79}$$

However, the GM dispersion relation satisfies $\widehat{\omega}(\beta\mathbf{k}) = 0 = \beta\,\widehat{\omega}$, so that $\det(M_\beta) = 0$ for all β. Thus the leading-order GM higher harmonics do not vanish:

$$\mathbf{Z}_\beta^{(0)} \neq 0 \quad \text{for GMs and all } \beta, \tag{80}$$

and they satisfy the polarization relations (76)–(78). The difference between the two cases lies in the fact that GMs can force higher harmonics which are GMs as well, whereas IGW higher harmonics cannot be IGWs so that the basic-wave interaction with its higher harmonics is non-resonant and leads to a response at the next order in ε, as discussed below in section 6.2.4. While the GM results are thus less trivial than in the IGW case, they still provide valuable information in the form of the polarization relations above and the amplitude equations derived below.

4.3. *Eikonal equations*

From (75) follows the IGW dispersion relation:

$$\omega = \Omega(\mathbf{X}, T, \mathbf{k}) = \mathbf{k}\cdot\mathbf{U}_0^{(0)}(\mathbf{X}, T) \pm \sqrt{\frac{{N_0}^2(Z)(k^2+l^2)+f_0^2 m^2}{m^2}}. \tag{81}$$

Both ω and $\mathbf{k}$ depend on $(\mathbf{X}, T)$, and by definition they satisfy

$$\frac{\partial\mathbf{k}}{\partial T} = \frac{\partial}{\partial T}\nabla_X\phi = \nabla_X\frac{\partial\phi}{\partial T} = -\nabla_X\omega. \tag{82}$$

From (81) and (82) follow the eikonal equations, with $\mathbf{c}_\mathrm{g} = \nabla_\mathbf{k}\Omega$ the local group velocity,

$$\left(\frac{\partial}{\partial T} + \mathbf{c}_\mathrm{g}\cdot\nabla_X\right)\omega = \frac{\partial\Omega}{\partial T}, \tag{83}$$

$$\left(\frac{\partial}{\partial T} + \mathbf{c}_\mathrm{g}\cdot\nabla_X\right)\mathbf{k} = -\nabla_X\Omega, \tag{84}$$

which can be used for predicting frequency and wavenumber. In the present context they are

$$\left(\frac{\partial}{\partial T} + \mathbf{c}_\mathrm{g}\cdot\nabla_X\right)\omega = \mathbf{k}\cdot\frac{\partial\mathbf{U}_0^{(0)}}{\partial T}, \tag{85}$$

$$\left(\frac{\partial}{\partial T} + \mathbf{c}_\mathrm{g}\cdot\nabla_X\right)\mathbf{k} = -\left(\nabla_X\mathbf{U}_0^{(0)}\right)\cdot\mathbf{k} + \frac{k^2+l^2}{2\widehat{\omega}m^2}\frac{\mathrm{d}N_0^2}{\mathrm{d}Z}\mathbf{e}_z. \tag{86}$$

With the last term in (86) removed, these eikonal equations also hold for the GMs. For these waves the group velocity equals the leading-order synoptic-scale horizontal wind, i.e. $\mathbf{c}_g = \mathbf{U}_0^{(0)}$ for the GMs.

5. Next-order equations

The leading-order equations have established well-known equilibria for the synoptic-scale flow, and the Boussinesq dispersion and polarization relations for hydrostatic waves. Since the vertical scale of the waves was assumed to be smaller, by one order in ε, than the density scale height, it is not surprising that the waves are found to locally follow Boussinesq dynamics. We stress again that all these results hold at finite wave amplitudes, close to the level of static instability. What has not been touched on so far is whether and how the wave amplitude responds to the synoptic-scale flow, and whether and how waves can influence the latter. This can be settled by considering the respective next orders of the basic equations. In this section the next-order terms will be identified. They will be used in section 6 for the derivation of wave-amplitude equations, and in section 7 for analyzing the wave impact on the synoptic-scale flow.

Using (58), the wave part of the $\mathcal{O}(\varepsilon)$ of the *Exner-pressure equation* (46) yields

$$\mathrm{i}\beta\mathbf{k}\cdot\mathbf{V}_\beta^{(1)} = -\frac{1}{\overline{P}^{(0)}}\nabla_X\cdot\left(\overline{P}^{(0)}\mathbf{V}_\beta^{(0)}\right), \tag{87}$$

while the mean-flow part is

$$\nabla_X\cdot\mathbf{V}_0^{(0)} = 0. \tag{88}$$

The latter is expected since (62) establishes $W_0^{(0)} = 0$, and since the leading-order horizontal synoptic-scale flow is in geostrophic equilibrium according to (71), and therefore non-divergent. Finally, from the next $\mathcal{O}(\varepsilon^2)$ terms in (46), only the mean-flow part is needed, i.e.

$$(1-\alpha)\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\Pi_0^{(0)}+\frac{R}{c_v}\frac{\overline{\Pi}^{(0)}}{\overline{P}^{(0)}}\nabla_X\cdot\left(\overline{P}^{(0)}\mathbf{V}_0^{(1)}\right)=0. \tag{89}$$

In the weakly stratified case ($\alpha = 1$), where $\overline{\Theta}^{(0)}$ is a constant, this amounts via (58) to $\nabla_X\cdot\left(\overline{R}^{(0)}\mathbf{V}_0^{(1)}\right) = 0$, i.e. the leading-order ageostrophic mass flux is non-divergent. In the case with moderately strong stratification ($\alpha = 0$), the leading-order synoptic-scale flow exhibits *elastic compressibility effects that would not be reproduced by a before-hand Boussinesq or anelastic ansatz*.

Likewise, from the wave part of the $\mathcal{O}(\varepsilon^{2+\alpha})$ of the entropy equation (45) one obtains, after division by $N_0\overline{\Theta}^{(0)}$, and using (87),

$$\begin{aligned}
-\mathrm{i}\beta\,\widehat{\omega}\frac{B_\beta^{(1)}}{N_0}+N_0W_\beta^{(1)} &= -\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\frac{B_\beta^{(0)}}{N_0}\\
&-\frac{\mathbf{V}_\beta^{(0)}}{N_0\overline{\Theta}^{(0)}}\cdot\nabla_X\Theta_0^{(0)}-\mathrm{i}\mathbf{k}\cdot\mathbf{V}_0^{(1)}\frac{B_\beta^{(0)}}{N_0}\\
&+\frac{1}{2N_0\overline{\Theta}^{(0)}}\sum_{\beta'=1}^{\infty}\sum_{\beta''=1}^{\infty}\left[D\left(\beta''/\beta',\mathbf{V}_{\beta'}^{(0)}\right)\Theta_{\beta''}^{(0)}\delta\left(\beta'+\beta''-\beta\right)\right.\\
&+D\left(-\beta''/\beta',\mathbf{V}_{\beta'}^{(0)}\right)\Theta_{\beta''}^{(0)*}\delta\left(\beta'-\beta''-\beta\right)\\
&\left.+D\left(-\beta''/\beta',\mathbf{V}_{\beta'}^{(0)*}\right)\Theta_{\beta''}^{(0)}\delta\left(-\beta'+\beta''-\beta\right)\right],
\end{aligned} \tag{90}$$

where $B_\beta^{(1)} = \Theta_\beta^{(1)}/\overline{\Theta}^{(0)}$ are the first-order non-dimensional wave-buoyancy amplitudes, and where we use the operator

$$D(\lambda,\mathbf{V}) = \frac{\lambda}{\overline{P}^{(0)}}\left[\nabla_X\cdot\left(\overline{P}^{(0)}\mathbf{V}\right)\right]-\mathbf{V}\cdot\nabla_X. \tag{91}$$

Due to (79), the nonlinear terms vanish in the IGW case. However, some nonlinearities remain even there, in the mean-flow part, yielding with the help of (87)

$$\begin{aligned}
&\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\Theta_0^{(0)}+W_0^{(1)}\overline{\Theta}^{(0)}N_0^2\\
&=-\frac{1}{2}\Re\frac{1}{\overline{P}^{(0)}}\nabla_X\cdot\left(\overline{P}^{(0)}\sum_{\beta=1}^{\infty}\mathbf{V}_\beta^{(0)}\Theta_\beta^{(0)*}\right).
\end{aligned} \tag{92}$$

The synoptic-scale potential temperature is forced by the wave-entropy flux convergence. In the IGW case only the basic wave contributes to the latter.

In the $\mathcal{O}(\varepsilon^{3+\alpha})$ terms in the *vertical-momentum equation* (44), we make use of $W_0^{(0)} = 0$, and of the hydrostatic equilibrium (67). The remaining wave parts then are

$$\begin{aligned}
&-B_\beta^{(1)}+\mathrm{i}\beta m\frac{c_p}{R}\overline{\Theta}^{(0)}\Pi_\beta^{(1)}\\
&=-\frac{c_p}{R}\overline{\Theta}^{(0)}\frac{\partial\Pi_\beta^{(0)}}{\partial Z}-\frac{c_p}{R}\Theta_\beta^{(0)}\frac{\partial}{\partial Z}\left[\alpha\overline{\Pi}^{(\alpha)}+(1-\alpha)\,\Pi_0^{(0)}\right]\\
&\quad-\frac{c_p}{R}im\left[\alpha\overline{\Theta}^{(\alpha)}+(1-\alpha)\,\Theta_0^{(0)}\right]\Pi_\beta^{(0)}\\
&\quad-\frac{1-\alpha}{2}\frac{c_p}{R}\sum_{\beta'=1}^{\infty}\sum_{\beta''=1}^{\infty}\left[\mathrm{i}\beta''m\Theta_{\beta'}^{(0)}\Pi_{\beta''}^{(0)}\delta\left(\beta'+\beta''-\beta\right)\right.\\
&\qquad+\mathrm{i}\beta''m\Theta_{\beta'}^{(0)*}\Pi_{\beta''}^{(0)}\delta\left(\beta'-\beta''-\beta\right)\\
&\qquad\left.-\mathrm{i}\beta''m\Theta_{\beta'}^{(0)}\Pi_{\beta''}^{(0)*}\delta\left(-\beta'+\beta''-\beta\right)\right].
\end{aligned} \tag{93}$$

Again the nonlinear terms vanish in the IGW case. They also do not appear in the weakly stratified case. The corresponding mean-flow part is not needed below.

The terms of $\mathcal{O}(\varepsilon^{3+\alpha})$ in the *horizontal-momentum equation* (43) have the wave parts

$$\begin{aligned}
&-\mathrm{i}\beta\widehat{\omega}\mathbf{U}_\beta^{(1)}+f\mathbf{e}_z\times\mathrm{i}\beta\mathbf{k}_h\mathbf{U}_\beta^{(1)}+\mathrm{i}\beta\mathbf{k}_h\frac{c_p}{R}\overline{\Theta}^{(0)}\Pi_\beta^{(1)}\\
&=-\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\mathbf{U}_\beta^{(0)}-\mathrm{i}\beta\mathbf{k}_h\cdot\mathbf{V}_0^{(1)}\mathbf{U}_\beta^{(0)}\\
&\quad-\mathbf{V}_\beta^{(0)}\cdot\nabla_X\mathbf{U}_0^{(0)}-\frac{c_p}{R}\overline{\Theta}^{(0)}\nabla_{X,h}\Pi_\beta^{(0)}-\frac{c_p}{R}\Theta_\beta^{(0)}(1-\alpha)\,\nabla_{X,h}\Pi_0^{(0)}\\
&\quad-\frac{c_p}{R}\mathrm{i}\beta\mathbf{k}_h\left[\alpha\overline{\Theta}^{(\alpha)}+(1-\alpha)\,\Theta_0^{(0)}\right]\Pi_\beta^{(0)}\\
&\quad+\frac{1}{2}\sum_{\beta'=1}^{\infty}\sum_{\beta''=1}^{\infty}\left\{\vphantom{\frac{c_p}{R}}\right.\\
&\qquad\left[D\left(\beta''/\beta',\mathbf{V}_{\beta'}^{(0)}\right)\mathbf{U}_{\beta''}^{(0)}-(1-\alpha)\mathrm{i}\beta''\mathbf{k}_h\frac{c_p}{R}\Theta_{\beta'}^{(0)}\Pi_{\beta''}^{(0)}\right]\\
&\qquad\times\delta\left(\beta'+\beta''-\beta\right)\\
&\quad+\left[D\left(-\beta''/\beta',\mathbf{V}_{\beta'}^{(0)}\right)\mathbf{U}_{\beta''}^{(0)*}+(1-\alpha)\mathrm{i}\beta''\mathbf{k}_h\frac{c_p}{R}\Theta_{\beta'}^{(0)}\Pi_{\beta''}^{(0)*}\right]\\
&\qquad\times\delta\left(\beta'-\beta''-\beta\right)\\
&\quad+\left[D\left(-\beta''/\beta',\mathbf{V}_{\beta'}^{(0)*}\right)\mathbf{U}_{\beta''}^{(0)}-(1-\alpha)\mathrm{i}\beta''\mathbf{k}_h\frac{c_p}{R}\Theta_{\beta'}^{(0)*}\Pi_{\beta''}^{(0)}\right]\\
&\qquad\left.\times\delta\left(-\beta'+\beta''-\beta\right)\right\}.
\end{aligned} \tag{94}$$

Here as well the nonlinear terms vanish in the IGW case. The corresponding mean-flow part finally yields, again using (60) and (87),

$$\left(\frac{\partial}{\partial T} + \mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\mathbf{U}_0^{(0)} + f_0\mathbf{e}_z\times\mathbf{U}_0^{(1)} = -\frac{c_p}{R}\overline{\Theta}^{(0)}\nabla_{X,h}\Pi_0^{(1)} - \frac{c_p}{R}\left[\alpha\overline{\Theta}^{(\alpha)} + (1-\alpha)\,\Theta_0^{(0)}\right]\nabla_{X,h}\Pi_0^{(0)} - \frac{1}{2}\Re\frac{1}{\overline{P}^{(0)}}\nabla_X \cdot\left(\overline{P}^{(0)}\sum_{\beta=1}^{\infty}\mathbf{V}_\beta^{(0)}\mathbf{U}_\beta^{(0)*}\right) + \frac{1-\alpha}{2}\frac{c_p}{R}\Re\sum_{\beta=1}^{\infty} i\beta\mathbf{k}_h\Theta_\beta^{(0)}\Pi_\beta^{(0)*}. \tag{95}$$

This describes the impact of the wave-momentum flux convergences, but in the moderately strongly stratified case ($\alpha = 0$) also of contributions of potential-temperature and Exner-pressure fluctuations to an elastic mean pressure-gradient force, on the synoptic-scale horizontal flow. It seems thus important to take the route from the fully compressible equations in order not to miss potentially essential aspects. Also here the wave impact is in the IGW case only due to the basic wave. The net wave impact on the mean flow will be discussed further below. However, first we address the mean-flow impact on the wave amplitude.

6. Wave-action conservation and potential-enstrophy equations

The wave equations can lead us to the prediction of the IGW amplitude via the concept of wave-action conservation. Likewise the GM amplitudes can be predicted from coupled potential-enstrophy equations. One first derives the wave-energy theorem, then reformulates pressure flux and the various production terms, using the dispersion and polarization relations as well as the mean-flow balance conditions, and finally combines all, also using the eikonal equations.

6.1. Wave-energy theorem

The wave equations (94), (93), (90), and (87) can be summarized as

$$M_\beta\mathbf{Z}_\beta^{(1)} = \mathbf{R}_\beta, \tag{96}$$

with

$$\mathbf{Z}_\beta^{(1)\mathrm{T}} = \left(U_\beta^{(1)}, V_\beta^{(1)}, W_\beta^{(1)}, \frac{1}{N_0}B_\beta^{(1)}, \frac{c_p}{R}\overline{\Theta}^{(0)}\Pi_\beta^{(1)}\right), \tag{97}$$

the transposed vector of the first-order wave amplitudes, and where $\mathbf{R}_\beta$ can be read from the right-hand side of the equations. Let us now set aside the IGW higher harmonics – which are zero to leading order – and focus on either

(i) the IGW basic wave alone ($\beta = 1$) or

(ii) the GM basic wave and all its higher harmonics ($\beta \geq 1$).

As we have seen before, in all of these cases M_β is singular so that it has a non-vanishing null space. Therefore $\mathbf{R}_\beta$ may not project onto this null space. Up to a constant factor, the latter is given by the null vector $\mathbf{Z}_\beta^{(0)}$ satisfying the polarization relations (76)–(78). By definition $M_\beta\mathbf{Z}_\beta^{(0)} = 0$, and thus also, with $\mathbf{Z}_\beta^{(0)+}$ the complex conjugate transpose of $\mathbf{Z}_\beta^{(0)}$,

$$\mathbf{Z}_\beta^{(0)+}M_\beta = 0, \tag{98}$$

since M is anti-Hermitian. Therefore, multiplying (96) by $\mathbf{N}_\beta^+$ yields

$$0 = \mathbf{Z}_\beta^{(0)+}\mathbf{R}_\beta. \tag{99}$$

In evaluating this we note that, due to the polarization relation (77), there is no vertical wave-buoyancy flux,

$$\Re\left(B_\beta^{(0)*}W_\beta^{(0)}\right) = 0. \tag{100}$$

Moreover, there is no leading-order synoptic-scale vertical flow, $W_0^{(0)} = 0$, so that one obtains from the real part of (99) the prognostic equation

$$\left(\frac{\partial}{\partial T} + \mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)E_{\beta,w} + \underbrace{\frac{1}{2}\Re\nabla_X\cdot\left(\overline{P}^{(0)}\frac{c_p}{R}\Pi_\beta^{(0)*}\mathbf{V}_\beta^{(0)}\right)}_{\text{pressure (energy) flux}} = \underbrace{-\frac{1}{2}\Re\left(\overline{R}^{(0)}\mathbf{U}_\beta^{(0)*}\mathbf{V}_\beta^{(0)}\right)\cdot\cdot\nabla_X\mathbf{U}_0^{(0)}}_{\text{shear production}} \underbrace{-\frac{1}{2}\Re\left(\frac{\overline{R}^{(0)}}{{N_0}^2}B_\beta^{(0)*}\mathbf{U}_\beta^{(0)}\right)\cdot\nabla_{X,h}\frac{\Theta_0^{(0)}}{\overline{\Theta}^{(0)}}}_{\text{buoyancy production}} \underbrace{-\frac{1-\alpha}{2}\Re\left(\overline{P}^{(0)}\frac{c_p}{R}B_\beta^{(0)*}\mathbf{U}_\beta^{(0)}\right)\cdot\nabla_{X,h}\Pi_0^{(0)}}_{\text{elastic term}} + \underbrace{T_\beta}_{\text{triad term}} \tag{101}$$

for the energy density

$$E_{\beta,w} = \frac{\overline{R}^{(0)}}{2}\left(\frac{|\mathbf{U}_\beta^{(0)}|^2}{2} + \frac{1}{{N_0}^2}\frac{|B_\beta^{(0)}|^2}{2}\right) \tag{102}$$

of either the basic wave ($\beta = 1$) or, in the GM case, any of its higher harmonics ($\beta \geq 2$). Both advection by the mean flow and pressure or energy flux redistribute wave energy, while shear production and buoyancy production act as sources or sinks for the latter. In the case with moderately strong stratification ($\alpha = 0$), the latter are supplemented by an elastic term, arising from the potential-temperature fluctuations in the horizontal pressure-gradient force. This term would not occur in an analysis based on the Boussinesq or anelastic equations. The nonlinear triad term

$$T_\beta = \frac{\overline{R}^{(0)}}{4}\Re\sum_{\beta'=1}^{\infty}\sum_{\beta''=1}^{\infty} \times\left\{\left[\mathbf{U}_\beta^{(0)+}D\left(\frac{\beta''}{\beta'},\mathbf{V}_{\beta'}^{(0)}\right)\mathbf{U}_{\beta''}^{(0)} + \frac{B_\beta^{(0)*}}{N_0^2\overline{\Theta}^{(0)}}D\left(\frac{\beta''}{\beta'},\mathbf{V}_{\beta'}^{(0)}\right)\Theta_{\beta''}^{(0)}\right]\times\delta\left(\beta'+\beta''-\beta\right) + \left[\mathbf{U}_\beta^{(0)+}D\left(-\frac{\beta''}{\beta'},\mathbf{V}_{\beta'}^{(0)}\right)\mathbf{U}_{\beta''}^{(0)*} + \frac{B_\beta^{(0)*}}{N_0^2\overline{\Theta}^{(0)}}D\left(-\frac{\beta''}{\beta'},\mathbf{V}_{\beta'}^{(0)}\right)\Theta_{\beta''}^{(0)*}\right]\times\delta\left(\beta'-\beta''-\beta\right) + \left[\mathbf{U}_\beta^{(0)+}D\left(-\frac{\beta''}{\beta'},\mathbf{V}_{\beta'}^{(0)*}\right)\mathbf{U}_{\beta''}^{(0)} + \frac{B_\beta^{(0)*}}{N_0^2\overline{\Theta}^{(0)}}D\left(-\frac{\beta''}{\beta'},\mathbf{V}_{\beta'}^{(0)*}\right)\Theta_{\beta''}^{(0)}\right]\times\delta\left(-\beta'+\beta''-\beta\right)\right\} \tag{103}$$

only contributes in the GM case. The whole can be further simplified by noting that geostrophy (71) and hydrostaticity (69) lead together to the thermal-wind relations

$$\nabla_{X,h}\frac{\Theta_0^{(0)}}{\overline{\Theta}^{(0)}} = -f_0\mathbf{e}_z\times\frac{\partial\mathbf{U}_0^{(0)}}{\partial Z} - (1-\alpha)N_0^2\frac{c_p}{R}\overline{\Theta}^{(0)}\nabla_{X,h}\Pi_0^{(0)} \tag{104}$$

for the synoptic-scale flow, so that

$$\begin{aligned}
&\frac{1}{2}\Re\left(\overline{R}^{(0)}\mathbf{U}_\beta^{(0)*}W_\beta^{(0)}\right)\cdot\frac{\partial \mathbf{U}_0^{(0)}}{\partial Z}\\
&+\frac{1}{2}\Re\left(\frac{\overline{R}^{(0)}}{N_0{}^2}B_\beta^{(0)*}\mathbf{U}_\beta^{(0)}\right)\cdot\nabla_{X,h}\frac{\Theta_0^{(0)}}{\overline{\Theta}^{(0)}}\\
&+\frac{1-\alpha}{2}\Re\left(\overline{P}^{(0)}\frac{c_p}{R}B_\beta^{(0)*}\mathbf{U}_\beta^{(0)}\right)\cdot\nabla_{X,h}\Pi_0^{(0)}\\
=&\frac{1}{2}\left[\Re\left(\overline{R}^{(0)}\mathbf{U}_\beta^{(0)*}W_\beta^{(0)}\right)+f_0\mathbf{e}_z\times\Re\left(\frac{\overline{R}^{(0)}}{N_0{}^2}\mathbf{U}_\beta^{(0)*}B_\beta^{(0)}\right)\right]\cdot\frac{\partial \mathbf{U}_0^{(0)}}{\partial Z}. \qquad (105)
\end{aligned}$$

With this, the wave-energy theorem finally becomes

$$\begin{aligned}
&\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)E_{\beta,w}+\frac{1}{2}\Re\nabla_X\cdot\left(\overline{P}^{(0)}\frac{c_p}{R}\Pi_\beta^{(0)*}\mathbf{V}_\beta^{(0)}\right)\\
&=-\frac{1}{2}\Re\left(\overline{R}^{(0)}\mathbf{U}_\beta^{(0)*}\mathbf{U}_\beta^{(0)}\right)\cdot\cdot\nabla_{X,h}\mathbf{U}_0^{(0)}\\
&\quad-\frac{1}{2}\left[\Re\left(\overline{R}^{(0)}\mathbf{U}_\beta^{(0)*}W_\beta^{(0)}\right)\right.\\
&\qquad\left.+f_0\mathbf{e}_z\times\Re\left(\frac{\overline{R}^{(0)}}{N_0{}^2}\mathbf{U}_\beta^{(0)*}B_\beta^{(0)}\right)\right]\cdot\frac{\partial \mathbf{U}_0^{(0)}}{\partial Z}+T_\beta\\
&\equiv\mathbf{E}_\beta\cdot\cdot\nabla_X\mathbf{U}_0^{(0)}+T_\beta, \qquad (106)
\end{aligned}$$

where the vertical-shear production, buoyancy production and the non-Boussinesq terms have been combined to effective vertical shear-production terms so that all involved fluxes can be written in terms of the celebrated Eliassen–Palm flux tensor $\mathbf{E}_\beta$ (Eliassen and Palm, 1961; Andrews and McIntyre, 1976).

6.2. *Wave-action conservation for inertia-gravity waves*

6.2.1. *Reformulation of the energy flux*

In the IGW case only the basic wave ($\beta = 1$) appears to leading order. Hence the IGW-energy density $E_{\mathrm{gw}} = E_{1,\mathrm{w}}$ is the energy density of the basic wave. Due to the polarization relations (76), and the dispersion relation (75) it is

$$E_{\mathrm{gw}} = \overline{R}^{(0)}\frac{|B_1^{(0)}|^2}{2N_0{}^2}\frac{\widehat{\omega}^2m^2}{N_0{}^2(k^2+l^2)}. \qquad (107)$$

Via the dispersion relation (83), the horizontal and vertical parts of the intrinsic group velocity are found to be

$$\widehat{\mathbf{c}}_{g,h} = \nabla_{K,h}\widehat{\omega} = \mathbf{k}_h\frac{N_0{}^2}{\widehat{\omega}m^2}, \qquad (108)$$

$$\widehat{c}_{g,z} = \frac{\partial\widehat{\omega}}{\partial m} = -\frac{N_0{}^2(k^2+l^2)}{\widehat{\omega}m^3}. \qquad (109)$$

On the other hand, by the polarization relations the horizontal and vertical pressure fluxes are

$$\frac{1}{2}\Re\left(\overline{P}^{(0)}\frac{c_p}{R}\Pi_1^{(0)*}\mathbf{U}_1^{(0)}\right) = \frac{\mathbf{k}_h\widehat{\omega}}{k^2+l^2}\overline{R}^{(0)}\frac{|B_1^{(0)}|^2}{2N_0{}^2}, \qquad (110)$$

$$\frac{1}{2}\Re\left(\overline{P}^{(0)}\frac{c_p}{R}\Pi_1^{(0)*}W_1^{(0)}\right) = -\frac{\widehat{\omega}}{m}\overline{R}^{(0)}\frac{|B_1^{(0)}|^2}{2N_0{}^2}. \qquad (111)$$

Comparison with the above then shows that the pressure and energy flux can be written as product between wave-energy density and intrinsic group velocity

$$\frac{1}{2}\Re\left(\overline{P}^{(0)}\frac{c_p}{R}\Pi_1^{(0)*}\mathbf{V}_1^{(0)}\right) = \widehat{\mathbf{c}}_g E_{\mathrm{gw}}. \qquad (112)$$

6.2.2. *Reformulation of the production terms*

We now use *geostrophy and hydrostaticity of the synoptic-scale flow* to convert the combined production terms. First we consider the contributions due to the horizontal synoptic-scale-flow wind gradients. Due to its geostrophy (71), the synoptic-scale horizontal wind is non-divergent,

$$\frac{\partial U_0^{(0)}}{\partial X}+\frac{\partial V_0^{(0)}}{\partial Y} = 0. \qquad (113)$$

By this, the polarization relations (76), the dispersion relation (83), and (107)–(109) we obtain

$$\begin{aligned}
&\frac{1}{2}\Re\left\{\overline{R}^{(0)}U_1^{(0)*}U_1^{(0)}\frac{\partial U_0^{(0)}}{\partial X}+\overline{R}^{(0)}V_1^{(0)*}V_1^{(0)}\frac{\partial V_0^{(0)}}{\partial Y}\right\}\\
&=\overline{R}^{(0)}\frac{|B_1^{(0)}|^2}{2N_0{}^2}\frac{1}{k^2+l^2}\left(k^2\frac{\partial U_0^{(0)}}{\partial X}+l^2\frac{\partial V_0^{(0)}}{\partial Y}\right)\\
&=\frac{E_{\mathrm{gw}}}{\widehat{\omega}}\widehat{c}_{gx}k\frac{\partial U_0^{(0)}}{\partial X}+\frac{E_{\mathrm{gw}}}{\widehat{\omega}}\widehat{c}_{gy}l\frac{\partial V_0^{(0)}}{\partial Y}. \qquad (114)
\end{aligned}$$

Likewise, one gets

$$\frac{1}{2}\Re\left(\overline{R}^{(0)}U_1^{(0)*}V_1^{(0)}\right) = \frac{E_{\mathrm{gw}}}{\widehat{\omega}}\widehat{c}_{gy}k = \frac{E_{\mathrm{gw}}}{\widehat{\omega}}\widehat{c}_{gx}l \qquad (115)$$

and

$$\frac{1}{2}\left[\Re\left(\overline{R}^{(0)}\mathbf{U}_1^{(0)*}W_1^{(0)}\right)+f_0\mathbf{e}_z\times\Re\left(\frac{\overline{R}^{(0)}}{N_0{}^2}\mathbf{U}_1^{(0)*}B_1^{(0)}\right)\right] = \frac{E_{\mathrm{gw}}}{\widehat{\omega}}\widehat{c}_{gz}\mathbf{k}_h. \qquad (116)$$

6.2.3. *Wave-action equation*

In summary, inserting (112) and (114)–(116) into (106) yields

$$\begin{aligned}
0=&\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)E_{\mathrm{gw}}+\frac{E_{\mathrm{gw}}}{\widehat{\omega}}\mathbf{k}\widehat{\mathbf{c}}_g\cdot\cdot\nabla_X\mathbf{U}_0^{(0)}\\
&+\nabla_X\cdot\left(\widehat{\mathbf{c}}_g E_{\mathrm{gw}}\right). \qquad (117)
\end{aligned}$$

We have $\widehat{\mathbf{c}}_g = \mathbf{c}_g - \mathbf{U}_0^{(0)}$ and $\nabla_X\cdot\mathbf{U}_0^{(0)} = 0$, so that

$$\begin{aligned}
0=&\,\widehat{\omega}\left[\frac{\partial}{\partial T}\left(\frac{E_{\mathrm{gw}}}{\widehat{\omega}}\right)+\nabla_X\cdot\left(\mathbf{c}_g\frac{E_{\mathrm{gw}}}{\widehat{\omega}}\right)\right]\\
&+\frac{E_{\mathrm{gw}}}{\widehat{\omega}}\left(\frac{\partial}{\partial T}+\mathbf{c}_g\cdot\nabla_X\right)\widehat{\omega}+\frac{E_{\mathrm{gw}}}{\widehat{\omega}}\mathbf{k}\widehat{\mathbf{c}}_g\cdot\cdot\nabla_X\mathbf{U}_0^{(0)}. \qquad (118)
\end{aligned}$$

Application of the eikonal equations (85) and (86), and of (62), leads to

$$\left(\frac{\partial}{\partial T}+\mathbf{c}_g\cdot\nabla_X\right)\widehat{\omega} = -\mathbf{k}\widehat{\mathbf{c}}_g\cdot\cdot\nabla_X\mathbf{U}_0^{(0)}. \qquad (119)$$

With this we finally obtain the conservation law

$$\frac{\partial\mathcal{A}}{\partial T}+\nabla_X\cdot\left(\mathbf{c}_g\mathcal{A}\right) = 0 \qquad (120)$$

for the IGW wave-action density $\mathcal{A} = E_{\mathrm{gw}}/\widehat{\omega}$. This facilitates the prediction of the wave amplitude.

As is important for the discussion below, we note that wave action has several 'relatives' which have important applications in wave–mean flow interaction theory (Andrews and McIntyre, 1978b; Bühler and McIntyre, 1998, 2005; Bühler, 2009, 2010), one of them the horizontal pseudo-momentum vector $\mathbf{p}_h = \mathbf{k}_h\mathcal{A}$.

With the help of (120) and (86), it can be shown that the horizontal pseudo-momentum vector obeys

$$\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\mathbf{p}_h=-\nabla_X\cdot\left(\widehat{\mathbf{c}}_g\mathbf{p}_h\right)-\nabla_X\mathbf{U}_0^{(0)}\cdot\mathbf{p}_h, \qquad (121)$$

where ∇_X contracts with $\widehat{\mathbf{c}}_g$. The divergence of the momentum flux tensor $\widehat{\mathbf{c}}_g\mathbf{p}_h$ defines a typical mean forcing $\mathbf{F}_h=-\nabla_X\cdot\left(\widehat{\mathbf{c}}_g\mathbf{p}_h\right)$ due to waves, also of relevance for the angular-pseudo-momentum equation. Taking the vertical component of the curl of (121) and using the leading-order non-divergence (113) of the background flow results in

$$\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\left[\mathbf{e}_z\cdot(\nabla_X\times\mathbf{p}_h)\right]=\mathbf{e}_z\cdot(\nabla_X\times\mathbf{F}_h). \qquad (122)$$

6.2.4. IGW higher harmonics

While the leading-order higher harmonics of a basic IGW vanish, the next-order higher harmonics can be determined directly from (96). Since M_β is non-singular for $\beta\geq 2$, one can solve this equation by

$$\mathbf{Z}_\beta^{(1)}=M_\beta^{-1}\mathbf{R}_\beta \qquad (\beta\geq 2); \qquad (123)$$

i.e. the first-order higher harmonics are slaved to the leading-order basic wave. In a further step we note that $\mathbf{R}_\beta$, i.e. the right-hand sides of the wave equations (94), (93), (90), and (87), vanishes for $\beta\geq 3$. This follows because the leading-order higher-harmonic amplitudes are all zero (79). Thus only the nonlinear triad terms can contribute to $\mathbf{R}_\beta$. However, those are only due to the non-zero basic wave, yielding only triad contributions to the second harmonic $\beta=2$. Therefore the only non-zero first-order higher harmonic is the second harmonic, i.e.

$$\mathbf{Z}_\beta^{(1)}=0 \qquad (\beta\geq 3). \qquad (124)$$

6.3. Potential-enstrophy dynamics for the geostrophic mode

6.3.1. Reformulation of the wave energy and the pressure flux

In the GM case we have separate energy equations for the basic wave ($\beta=1$) and all higher harmonics ($\beta\geq 2$). Due to the polarization relations (76), and the dispersion relation (74), the energy density $E_{\beta,\mathrm{gm}}$ of the βth GM harmonic is

$$E_{\beta,\mathrm{gm}}=\overline{R}^{(0)}\frac{|B_\beta^{(0)}|^2}{4{N_0}^2}\left(1+\frac{N_0^2}{f_0^2}\frac{k^2+l^2}{m^2}\right), \qquad (125)$$

while one finds from the polarization relations that there is no pressure flux, i.e.

$$\frac{1}{2}\Re\left(\overline{P}^{(0)}\frac{c_p}{R}\Pi_\beta^{(0)*}\mathbf{V}_\beta^{(0)}\right)=0. \qquad (126)$$

6.3.2. Reformulation of the production terms

Again we use geostrophy and hydrostaticity of the synoptic-scale flow to convert the combined production terms. First we consider the contributions due to the horizontal synoptic-scale-flow wind gradients. Due to the horizontal non-divergence (113) of the synoptic-scale flow, and due to the polarization relations (76), and the GM dispersion relation (74), we obtain, using (125),

$$\begin{aligned}\frac{1}{2}\Re&\left\{\overline{R}^{(0)}U_\beta^{(0)*}U_\beta^{(0)}\frac{\partial U_0^{(0)}}{\partial X}+\overline{R}^{(0)}V_\beta^{(0)*}V_\beta^{(0)}\frac{\partial V_0^{(0)}}{\partial Y}\right\}\\ &=-\overline{R}^{(0)}\frac{|B_\beta^{(0)}|^2}{2{N_0}^2}\frac{N_0^2}{f_0^2\beta^2m^2}\left(\beta^2k^2\frac{\partial U_0^{(0)}}{\partial X}+\beta^2l^2\frac{\partial V_0^{(0)}}{\partial Y}\right)\\ &=\frac{E_{\beta,\mathrm{gm}}}{\widehat{\gamma}_\beta}\widehat{c}_{\beta,\gamma x}\beta k\frac{\partial U_0^{(0)}}{\partial X}+\frac{E_{\beta,\mathrm{gm}}}{\widehat{\gamma}_\beta}\widehat{c}_{\beta,\gamma y}\,l\frac{\partial V_0^{(0)}}{\partial Y},\end{aligned} \qquad (127)$$

where

$$\widehat{\gamma}_\beta(\beta\mathbf{k})=\frac{N_0^2}{N_0^2(\beta^2k^2+\beta^2l^2)+f_0^2\beta^2m^2} \qquad (128)$$

is a wavenumber-dependent function, and

$$\widehat{\mathbf{c}}_{\beta,\gamma}=\nabla_{\beta\mathbf{k}}\widehat{\gamma}_\beta=-2\frac{N_0^2\left(N_0^2\beta k,N_0^2\beta\,l,f_0^2\beta m\right)^t}{\left\{N_0^2(\beta^2k^2+\beta^2l^2)+f_0^2\beta^2m^2\right\}^2} \qquad (129)$$

the corresponding 'group velocity'. Likewise we obtain

$$\frac{1}{2}\Re\left(\overline{R}^{(0)}U_\beta^{(0)*}V_\beta^{(0)}\right)=\frac{E_{\beta,\mathrm{gm}}}{\widehat{\gamma}_\beta}\widehat{c}_{\beta,\gamma y}\beta k=\frac{E_{\beta,\mathrm{gm}}}{\widehat{\gamma}_\beta}\widehat{c}_{\beta,\gamma x}\beta\,l. \qquad (130)$$

Due to the polarization relation (77) the vertical-wind amplitude of the GM vanishes. Hence vertical momentum fluxes do not contribute to the Eliassen–Palm production. To further simplify the latter, we again use the polarization relations and the dispersion relation to show that

$$\begin{aligned}\frac{1}{2}&\left[\Re\left(\overline{R}^{(0)}\mathbf{U}_\beta^{(0)*}W_\beta^{(0)}\right)+f_0\mathbf{e}_z\times\Re\left(\frac{\overline{R}^{(0)}}{{N_0}^2}\mathbf{U}_\beta^{(0)*}B_\beta^{(0)}\right)\right]\cdot\frac{\partial\mathbf{U}_0^{(0)}}{\partial Z}\\ &=\frac{E_{\beta,\mathrm{gm}}}{\widehat{\gamma}_\beta}\widehat{c}_{\beta,\gamma z}\beta\mathbf{k}_h\cdot\frac{\partial\mathbf{U}_0^{(0)}}{\partial Z}.\end{aligned} \qquad (131)$$

6.3.3. Reformulation of the nonlinear triad term

The reformulation of the nonlinear triad term (103) is detailed in Appendix A. One obtains after some algebra

$$\begin{aligned}T_\beta=&\widehat{\gamma}_\beta\frac{\overline{R}^{(0)}}{4}\Re\sum_{\beta'=1}^{\infty}\sum_{\beta''=1}^{\infty}\\ &\times\left\{P_\beta^{(0)*}D\left(\frac{\beta''}{\beta'},\mathbf{U}_{\beta'}^{(0)}\right)P_{\beta''}^{(0)}\delta\left(\beta'+\beta''-\beta\right)\right.\\ &+P_\beta^{(0)*}D\left(-\frac{\beta''}{\beta'},\mathbf{U}_{\beta'}^{(0)}\right)P_{\beta''}^{(0)*}\delta\left(\beta'-\beta''-\beta\right)\\ &\left.+P_\beta^{(0)*}D\left(-\frac{\beta''}{\beta'},\mathbf{U}_{\beta'}^{(0)*}\right)P_{\beta''}^{(0)}\delta\left(-\beta'+\beta''-\beta\right)\right\}.\end{aligned} \qquad (132)$$

6.3.4. Potential-enstrophy equation

In summary, inserting (126), (127), and (130)–(131) into (106) yields

$$0=\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)E_{\beta,\mathrm{gm}}+\frac{E_{\beta,\mathrm{gm}}}{\widehat{\gamma}_\beta}\beta\mathbf{k}\widehat{\mathbf{c}}_{\beta,\gamma}\cdot\cdot\nabla\mathbf{U}_0^{(0)}+T_\beta. \qquad (133)$$

We have $\mathbf{c}_g=\mathbf{U}_0^{(0)}$ and $\nabla_X\cdot\mathbf{U}_0^{(0)}=0$, so that

$$\begin{aligned}0=&\widehat{\gamma}_\beta\left[\frac{\partial}{\partial T}\left(\frac{E_{\beta,\mathrm{gm}}}{\widehat{\gamma}_\beta}\right)+\nabla_X\cdot\left(\mathbf{c}_g\frac{E_{\beta,\mathrm{gm}}}{\widehat{\gamma}_\beta}\right)\right]\\ &+\frac{E_{\beta,\mathrm{gm}}}{\widehat{\gamma}_\beta}\left[\left(\frac{\partial}{\partial T}+\mathbf{c}_g\cdot\nabla_X\right)\widehat{\gamma}_\beta+\beta\mathbf{k}\widehat{\mathbf{c}}_{\beta,\gamma}\cdot\cdot\nabla\mathbf{U}_0^{(0)}\right]\\ &+T_\beta.\end{aligned} \qquad (134)$$

Application of the eikonal equations (86) leads to

$$\left(\frac{\partial}{\partial T}+\mathbf{c}_g\cdot\nabla_X\right)\widehat{\gamma}_\beta=-\beta\,\mathbf{k}\widehat{\mathbf{c}}_{\beta,\gamma}\cdot\cdot\nabla_X\mathbf{U}_0^{(0)}. \qquad (135)$$

With this and (132) we finally obtain the prognostic equation

$$\frac{\partial \mathcal{P}_\beta}{\partial T}+\nabla_X\cdot\left(\mathbf{c}_g\mathcal{P}_\beta\right)=\frac{\overline{R}^{(0)}}{4}\Re\sum_{\beta'=1}^{\infty}\sum_{\beta''=1}^{\infty}\times\left\{P_\beta^{(0)*}D\left(\frac{\beta''}{\beta'},\mathbf{U}_{\beta'}^{(0)}\right)P_{\beta''}^{(0)}\delta\left(\beta'+\beta''-\beta\right)+P_\beta^{(0)*}D\left(-\frac{\beta''}{\beta'},\mathbf{U}_{\beta'}^{(0)}\right)P_{\beta''}^{(0)*}\delta\left(\beta'-\beta''-\beta\right)+P_\beta^{(0)*}D\left(-\frac{\beta''}{\beta'},\mathbf{U}_{\beta'}^{(0)*}\right)P_{\beta''}^{(0)}\delta\left(-\beta'+\beta''-\beta\right)\right\} \quad (136)$$

for the leading-order potential enstrophy

$$\mathcal{P}_\beta=\frac{E_{\beta,\mathrm{gm}}}{\widehat{\gamma}_\beta}=\overline{R}^{(0)}\frac{|P_\beta^{(0)}|^2}{4} \quad (137)$$

of the βth GM harmonic. The latter identity can be verified using (125), the polarization relations (A1)–(A4), and the definition (A5). We note that the potential-enstrophy equations are not fully closed, as they would also need in products like $P_\beta^{(0)*}P_{\beta''}^{(0)}$ additional information about the phase of the GM harmonics, not provided by the potential enstrophies, here $\mathcal{P}_\beta$ and $\mathcal{P}_{\beta''}$. Closer inspection shows that one would need to know the next-order synoptic-scale-flow wind $\mathbf{V}_0^{(1)}$ to obtain these. Alternatively one could think about a random-phase approach. However, for the time being, we do not pursue this further.

7. Wave impact on the synoptic-scale flow

The synoptic-scale flow is governed by the horizontal momentum equation (95), the entropy equation (92), the Exner-pressure equation (89), geostrophic equilibrium (71) and hydrostatic equilibrium (69). In the following these will be used to derive a prognostic equation for the synoptic-scale PV. First, the vertical component of the curl of the horizontal momentum equation (95) yields a quasi-geostrophic vorticity equation with wave impact, and in the case with moderately strong stratification ($\alpha=0$) with the contributions from a baroclinic term and an elastic wave term which will need special attention,

$$\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\nabla_{X,h}^2\left(\frac{c_p}{R}\frac{\overline{\Theta}^{(0)}\Pi_0^{(0)}}{f_0}\right)+f_0\nabla_{X,h}\cdot\mathbf{U}_0^{(1)}=\frac{-1}{\overline{P}^{(0)}}\nabla_X\cdot\frac{\overline{P}^{(0)}}{2}\left[\frac{\partial}{\partial X}\Re\sum_{\beta=1}^{\infty}\mathbf{V}_\beta^{(0)}V_\beta^{(0)*}-\frac{\partial}{\partial Y}\Re\sum_{\beta=1}^{\infty}\mathbf{V}_\beta^{(0)}U_\beta^{(0)*}\right]+(1-\alpha)\frac{c_p}{R}\mathbf{e}_z\cdot\left[\nabla_X\Pi_0^{(0)}\times\nabla_X\Theta_0^{(0)}+\nabla_X\times\frac{1}{2}\Re\sum_{\beta=1}^{\infty}\mathrm{i}\beta\mathbf{k}_h\Theta_\beta^{(0)}\Pi_\beta^{(0)*}\right] \quad (138)$$

In the IGW case, only the basic wave ($\beta=1$) contributes to the fluxes involved. Due to (89), we find for the synoptic-scale horizontal divergence

$$\nabla_{X,h}\cdot\mathbf{U}_0^{(1)}=-(1-\alpha)\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\left(\frac{c_v}{R}\frac{\Pi_0^{(0)}}{\overline{\Pi}^{(0)}}\right)-\frac{1}{\overline{P}^{(0)}}\frac{\partial}{\partial Z}\left(\overline{P}^{(0)}W_0^{(1)}\right), \quad (139)$$

so that it is affected by compressibility effects in the case of moderately strong stratification. Using (92), $W_0^{(0)}=0$, and $\Re\left(W_\beta^{(0)}B_\beta^{(0)*}\right)=0$, we can re-express herein the vertical wind and then reinsert the whole into (138), yielding

$$\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\left[\nabla_{X,h}^2\left(\frac{c_p}{R}\frac{\overline{\Theta}^{(0)}\Pi_0^{(0)}}{f_0}\right)-(1-\alpha)f_0\frac{c_v}{R}\frac{\Pi_0^{(0)}}{\overline{\Pi}^{(0)}}\right]+\frac{f_0}{\overline{P}^{(0)}}\frac{\partial}{\partial Z}\left[\overline{P}^{(0)}\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\left(\frac{1}{N_0^2}\frac{\Theta_0^{(0)}}{\overline{\Theta}^{(0)}}\right)\right]-(1-\alpha)\frac{c_p}{R}\mathbf{e}_z\cdot\left(\nabla_X\Pi_0^{(0)}\times\nabla_X\Theta_0^{(0)}\right)=-\frac{f_0}{\overline{P}^{(0)}}\frac{\partial}{\partial Z}\left[\frac{1}{2}\Re\,\nabla_{X,h}\cdot\left(\frac{\overline{P}^{(0)}}{N_0^2}\sum_{\beta=1}^{\infty}\mathbf{U}_\beta^{(0)}B_\beta^{(0)*}\right)\right]-\frac{1}{\overline{P}^{(0)}}\nabla_X\cdot\frac{\overline{P}^{(0)}}{2}\left[\frac{\partial}{\partial X}\Re\sum_{\beta=1}^{\infty}\mathbf{V}_\beta^{(0)}V_\beta^{(0)*}-\frac{\partial}{\partial Y}\Re\sum_{\beta=1}^{\infty}\mathbf{V}_\beta^{(0)}U_\beta^{(0)*}\right]+(1-\alpha)\frac{c_p}{R}\mathbf{e}_z\cdot\nabla_X\times\frac{1}{2}\Re\sum_{\beta=1}^{\infty}\mathrm{i}\beta\mathbf{k}_h\Theta_\beta^{(0)}\Pi_\beta^{(0)*}. \quad (140)$$

Again, in the IGW case only the basic wave ($\beta=1$) contributes to the fluxes. Due to geostrophy and hydrostaticity the thermal-wind relations (104) hold, so that the terms on the left-hand side can be combined to yield

$$\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\left[\nabla_{X,h}^2\left(\frac{c_p}{R}\frac{\overline{\Theta}^{(0)}\Pi_0^{(0)}}{f_0}\right)-(1-\alpha)f_0\frac{c_v}{R}\frac{\Pi_0^{(0)}}{\overline{\Pi}^{(0)}}\right]+\frac{f_0}{\overline{P}^{(0)}}\frac{\partial}{\partial Z}\left[\overline{P}^{(0)}\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\left(\frac{1}{N_0^2}\frac{\Theta_0^{(0)}}{\overline{\Theta}^{(0)}}\right)\right]-(1-\alpha)\frac{c_p}{R}\mathbf{e}_z\cdot\left(\nabla_X\Pi_0^{(0)}\times\nabla_X\Theta_0^{(0)}\right)=\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\left[\nabla_{X,h}^2\left(\frac{c_p}{R}\frac{\overline{\Theta}^{(0)}\Pi_0^{(0)}}{f_0}\right)-(1-\alpha)f_0\frac{c_v}{R}\frac{\Pi_0^{(0)}}{\overline{\Pi}^{(0)}}+\frac{1}{\overline{P}^{(0)}}\frac{\partial}{\partial Z}\left(\overline{P}^{(0)}\frac{f_0}{N_0^2}\frac{\Theta_0^{(0)}}{\overline{\Theta}^{(0)}}\right)\right]. \quad (141)$$

Herein one has, by repeated use of the hydrostatic relations (67) and (69), and of the equation of state (58),

$$\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\frac{1}{\overline{P}^{(0)}}\frac{\partial}{\partial Z}\left(\overline{P}^{(0)}\frac{f_0}{N_0^2}\frac{\Theta_0^{(0)}}{\overline{\Theta}^{(0)}}\right)=\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)\left\{\frac{1}{\overline{R}^{(0)}}\frac{\partial}{\partial Z}\left[\overline{R}^{(0)}\frac{f_0}{N_0^2}\frac{\partial}{\partial Z}\left(\frac{c_p}{R}\overline{\Theta}^{(0)}\Pi_0^{(0)}\right)\right]+(1-\alpha)f_0\frac{c_v}{R}\frac{\Pi_0^{(0)}}{\overline{\Pi}^{(0)}}\right\}. \quad (142)$$

Hence one obtains a prognostic equation with wave impact

$$\left(\frac{\partial}{\partial T}+\mathbf{U}_0^{(0)}\cdot\nabla_{X,h}\right)P_0^{(0)}=-\frac{f_0}{\overline{P}^{(0)}}\frac{\partial}{\partial Z}\left[\frac{1}{2}\Re\,\nabla_{X,h}\cdot\left(\frac{\overline{P}^{(0)}}{N_0^2}\sum_{\beta=1}^{\infty}\mathbf{U}_\beta^{(0)}B_\beta^{(0)*}\right)\right]-\frac{1}{\overline{P}^{(0)}}\nabla_X\cdot\left\{\frac{\overline{P}^{(0)}}{2}\left[\frac{\partial}{\partial X}\Re\sum_{\beta=1}^{\infty}\mathbf{V}_\beta^{(0)}V_\beta^{(0)*}-\frac{\partial}{\partial Y}\Re\sum_{\beta=1}^{\infty}\mathbf{V}_\beta^{(0)}U_\beta^{(0)*}\right]\right\}+(1-\alpha)\frac{c_p}{R}\mathbf{e}_z\cdot\nabla_X\times\frac{1}{2}\Re\sum_{\beta=1}^{\infty}\mathrm{i}\beta\mathbf{k}_h\Theta_\beta^{(0)}\Pi_\beta^{(0)*} \quad (143)$$

for the leading-order synoptic-scale quasi-geostrophic PV

$$P_0^{(0)} = \nabla_{X,h}{}^2 \left(\frac{c_p}{R} \frac{\overline{\Theta}^{(0)} \Pi_0^{(0)}}{f_0} \right) + \frac{f_0}{\overline{R}^{(0)}} \frac{\partial}{\partial Z} \left[\frac{\overline{R}^{(0)}}{N_0{}^2} \frac{\partial}{\partial Z} \left(\frac{c_p}{R} \overline{\Theta}^{(0)} \Pi_0^{(0)} \right) \right]. \tag{144}$$

Supplementing classic derivations (Pedlosky, 1987), it is thus shown here that without wave impact the latter is conserved, even in the case of moderately strong stratification ($\alpha = 0$), not just in the weakly stratified case ($\alpha = 1$).

7.1. *Gravity-wave impact*

To reformulate the IGW impact, we decompose

$$\begin{aligned} \nabla_{X,h} \cdot &\left[\frac{\partial}{\partial X} \frac{1}{2} \Re \left(\mathbf{U}_1^{(0)} V_1^{(0)*} \right) - \frac{\partial}{\partial Y} \frac{1}{2} \Re \left(\mathbf{U}_1^{(0)} U_1^{(0)*} \right) \right] \\ &= \frac{1}{2} \Re \left\{ \frac{\partial^2}{\partial X^2} \left(U_1^{(0)} V_1^{(0)*} \right) + \frac{\partial^2}{\partial X \partial Y} \left(|V_1^{(0)}|^2 - |U_1^{(0)}|^2 \right) \right. \\ &\quad \left. - \frac{\partial^2}{\partial Y^2} \left(V_1^{(0)} U_1^{(0)*} \right) \right\}. \end{aligned} \tag{145}$$

The first and last term on the right-hand side can be reformulated using (115). Moreover, due to the polarization relations, the dispersion relation, (107), and (108), the middle term is

$$|U_1^{(0)}|^2 - |V_1^{(0)}|^2 = 2 \left(\widehat{c}_{gx} k \mathcal{A} - \widehat{c}_{gy} l \mathcal{A} \right) / \overline{R}^{(0)}, \tag{146}$$

so that

$$\begin{aligned} \nabla_{X,h} \cdot &\left[\frac{\partial}{\partial X} \frac{1}{2} \Re \left(\mathbf{U}_1^{(0)} V_1^{(0)*} \right) - \frac{\partial}{\partial Y} \frac{1}{2} \Re \left(\mathbf{U}_1^{(0)} U_1^{(0)*} \right) \right] \\ &= \frac{\partial}{\partial X} \left[\frac{1}{\overline{P}^{(0)}} \nabla_{X,h} \cdot \left(\overline{\Theta}^{(0)} \widehat{\mathbf{c}}_g l \mathcal{A} \right) \right] \\ &\quad - \frac{\partial}{\partial Y} \left[\frac{1}{\overline{P}^{(0)}} \nabla_{X,h} \cdot \left(\overline{\Theta}^{(0)} \widehat{\mathbf{c}}_g k \mathcal{A} \right) \right]. \end{aligned} \tag{147}$$

Taking also (116) into account, one finds that

$$\begin{aligned} &- \frac{f_0}{\overline{P}^{(0)}} \frac{\partial}{\partial Z} \left[\frac{1}{2} \Re \nabla_{X,h} \cdot \left(\frac{\overline{P}^{(0)}}{N_0^2} \mathbf{U}_1^{(0)} B_1^{(0)*} \right) \right] \\ &- \frac{1}{\overline{P}^{(0)}} \frac{\partial}{\partial Z} \left\{ \frac{\overline{P}^{(0)}}{2} \left[\frac{\partial}{\partial X} \Re \left(W_1^{(0)} V_1^{(0)*} \right) - \frac{\partial}{\partial Y} \Re \left(W_1^{(0)} U_1^{(0)*} \right) \right] \right\} \\ &= - \frac{\partial}{\partial X} \left[\frac{1}{\overline{P}^{(0)}} \frac{\partial}{\partial Z} \left(\overline{\Theta}^{(0)} c_{gz} l \mathcal{A} \right) \right] \\ &\quad + \frac{\partial}{\partial Y} \left[\frac{1}{\overline{P}^{(0)}} \frac{\partial}{\partial Z} \left(\overline{\Theta}^{(0)} c_{gz} k \mathcal{A} \right) \right]. \end{aligned} \tag{148}$$

Finally we use again the polarization relations to reformulate the elastic wave-impact term that appears in the case of moderately strong stratification ($\alpha = 0$). One obtains

$$\frac{1}{2} \frac{c_p}{R} \Re \left(\mathrm{i} \, \mathbf{k}_h \Theta_1^{(0)} \Pi_1^{(0)*} \right) = c_{gz} N_0^2 \mathbf{k}_h \mathcal{A} / \overline{R}^{(0)}, \tag{149}$$

and thus, using (66) with $\alpha = 0$,

$$\begin{aligned} &\frac{c_p}{R} \mathbf{e}_z \cdot \nabla_X \times \frac{1}{2} \Re \left(\mathrm{i} \, \mathbf{k}_h \Theta_1^{(0)} \Pi_1^{(0)*} \right) \\ &= \frac{\partial}{\partial X} \left(\frac{1}{\overline{P}^{(0)}} \frac{\mathrm{d} \overline{\Theta}^{(0)}}{\mathrm{d}Z} c_{gz} l \mathcal{A} \right) - \frac{\partial}{\partial Y} \left(\frac{1}{\overline{P}^{(0)}} \frac{\mathrm{d} \overline{\Theta}^{(0)}}{\mathrm{d}Z} c_{gz} k \mathcal{A} \right). \end{aligned} \tag{150}$$

Substitution of (148) and (150) into (143) finally leads to the prognostic equation

$$\begin{aligned} \left(\frac{\partial}{\partial T} + \mathbf{U}_0^{(0)} \cdot \nabla_{X,h} \right) P_0^{(0)} &= - \frac{\partial}{\partial X} \left[\frac{1}{\overline{R}^{(0)}} \nabla_X \cdot \left(\widehat{\mathbf{c}}_g l \mathcal{A} \right) \right] \\ &\quad + \frac{\partial}{\partial Y} \left[\frac{1}{\overline{R}^{(0)}} \nabla_X \cdot \left(\widehat{\mathbf{c}}_g k \mathcal{A} \right) \right] = \frac{\mathbf{e}_z}{\overline{R}^{(0)}} \cdot \left(\nabla_X \times \mathbf{F}_h \right). \end{aligned} \tag{151}$$

More light can be shed on this by using (122) for the divergence of the vector of IGW angular pseudo-momentum, yielding

$$\left(\frac{\partial}{\partial T} + \mathbf{U}_0^{(0)} \cdot \nabla_{X,h} \right) \left(P_0^{(0)} - \frac{\mathbf{e}_z}{\overline{R}^{(0)}} \cdot \nabla_X \times \mathbf{p}_h \right) = 0. \tag{152}$$

Consequently, changes in the vertical curl of pseudomomentum translate to changes in the background PV within a medium without friction, heating and heat conduction, as has been discussed within GLM theory by Bühler and McIntyre (1998, 2003, 2005). In agreement with these studies and Wagner and Young (2015), this result also implies that the theory respects the conservation of total PV, here

$$P_{\mathrm{tot}}^{(0)} = P_0^{(0)} - \frac{\mathbf{e}_z}{\overline{R}^{(0)}} \cdot \nabla_X \times \mathbf{p}_h, \tag{153}$$

consisting of the synoptic-scale-flow part $P_0^{(0)}$ and a wave contribution from the vertical curl of pseudo-momentum. An interesting difference to the studies referred to above and also Grimshaw (1975b) is that in the latter this result is formulated in terms of a Lagrangian-mean synoptic-scale flow, whereas our study does not use this kind of average. However, Grimshaw (1975b) shows that the difference between Eulerian-mean and Lagrangian-mean flow, the Stokes drift, is $\mathcal{O}(\varepsilon)$ so that it does not appear in our leading-order results. It would show up to the next order, however that is not of prior importance here.

7.2. *Geostrophic-mode impact*

The calculations for the GM case are analogous to the IGW case. One obtains the prognostic equation

$$\begin{aligned} \left(\frac{\partial}{\partial T} + \mathbf{U}_0^{(0)} \cdot \nabla_{X,h} \right) P_0^{(0)} &= - \frac{\partial}{\partial X} \left(\frac{1}{\overline{R}^{(0)}} \nabla_X \cdot \sum_{\beta=1}^{\infty} \widehat{\mathbf{c}}_{\beta,\gamma} \beta l \mathcal{P}_\beta \right) \\ &\quad + \frac{\partial}{\partial Y} \left(\frac{1}{\overline{R}^{(0)}} \nabla_X \cdot \sum_{\beta=1}^{\infty} \widehat{\mathbf{c}}_{\beta,\gamma} \beta k \mathcal{P}_\beta \right). \end{aligned} \tag{154}$$

However, for the moment, we do not see that there is an equivalent to the relation (152) between changes in angular pseudo-momentum and balanced PV. Notably, our results *cannot* be brought into agreement with those from the quasi-geostrophic theory of the interaction between synoptic-scale Rossby waves and planetary-scale mean flows, as summarized by Vallis (2006). Those theories show that Rossby waves have a conserved wave action that is potential enstrophy divided by the planetary-scale-flow PV gradient, while potential enstrophy itself is not conserved. GM potential enstrophy, however, is a conserved quantity itself. A first guess might be that quasi-geostrophic theory does not hold in this context since mesoscale GMs do not have a small Rossby number, as discussed above and summarized in Table 2. However, as is shown in Appendix B, both the potential-enstrophy equations (136) for the GM harmonics and the PV equation for the synoptic-scale flow with mesoscale GM impact (154) can be derived from quasi-geostrophic theory. Therefore it seems that it is rather the difference in scale between planetary-synoptic versus synoptic-mesoscale interactions that causes these discrepancies. In fact the planetary-vorticity gradient, excluded in our treatment by the f-plane assumption, is not felt significantly by mesoscale motions. Inclusion of the β-effect would supplement the planetary vorticity by a correction at $\mathcal{O}(\varepsilon^2)$ which we would also not expect to essentially influence the results from the calculations in Appendix B. Therefore it is not too surprising that the GM dynamics identified here differ from the dynamics of the interaction between planetary-scale motions and synoptic-scale vortices.

In conclusion, we have not derived an extension of quasi-geostrophic theory to describe the dynamics of mesoscale GMs in interaction with a synoptic-scale flow. All essentials seem to be imbedded in that theory. But even within this framework, the issue of the parametrization of unresolved mesoscale modes arises for simulations at sufficiently coarse resolution. This is the merit and purpose of respective results in this study.

8. Summary of the most essential equations in dimensional form

8.1. *Dispersion relation and polarization relations*

A re-dimensionalization of (75) by the substitutions

$$\widehat{\omega} \rightarrow \widehat{\omega} T_w = \widehat{\omega}/f, \tag{155}$$

$$\overline{\Theta}^{(\alpha)} \rightarrow \begin{cases} \overline{\theta}/T_{00} & \text{if } \alpha = 0, \\ \left(\overline{\theta}/T_{00} - \overline{\Theta}^{(0)}\right)/\varepsilon & \text{if } \alpha = 1, \end{cases} \tag{156}$$

$$Z \rightarrow \varepsilon z/H_w, \tag{157}$$

$$(k, l, m) \rightarrow [L_w(k, l), H_w m], \tag{158}$$

$$f_0 \rightarrow f/f, \tag{159}$$

$$\mathbf{U}_0^{(0)} \rightarrow \mathbf{U}/U_w, \tag{160}$$

and the identities (23) and (27) lead to the dimensional IGW dispersion relation

$$\widehat{\omega}^2 = (\omega - \mathbf{k}\cdot\mathbf{U})^2 = f^2 + N^2 \frac{k^2 + l^2}{m^2}, \tag{161}$$

with $N^2 = g(\mathrm{d}\overline{\theta}/\mathrm{d}z)/\overline{\theta}$.

Re-dimensionalizing (76)–(78) by the substitutions

$$\left(\mathbf{U}_\beta^{(0)}, W_\beta^{(0)}, B_\beta^{(0)}, \Pi_\beta^{(0)}\right) \rightarrow \left(\frac{\mathbf{u}'_\beta}{U_w}, \frac{w'_\beta}{W_w}, \frac{\theta'_\beta}{\varepsilon^{1+\alpha}\overline{\theta}}, \frac{\pi'_\beta}{\varepsilon^{2+\alpha}}\right) \tag{162}$$

results in the dimensional polarization relations

$$\mathbf{u}'_\beta = \frac{\beta^2 \mathbf{k}_h \widehat{\omega} - \mathrm{i} f \mathbf{e}_z \times \beta \mathbf{k}_h}{\beta^2 \widehat{\omega}^2 - f^2} \frac{b'_\beta}{\mathrm{i}\beta m}, \tag{163}$$

$$w'_\beta = \frac{\mathrm{i}\beta\widehat{\omega}}{N^2} b'_\beta, \tag{164}$$

$$c_p \overline{\theta} \pi'_\beta = \frac{b'_\beta}{\mathrm{i}\beta m}, \tag{165}$$

where $b'_\beta = g\,\theta'_\beta/\overline{\theta}$ is the dimensional buoyancy of the leading-order βth wave harmonic. In the IGW case, only the basic wave ($\beta = 1$) has non-zero leading-order amplitudes.

8.2. *Gravity-wave dynamics*

Likewise we obtain the dimensional IGW wave-action equation

$$\frac{\partial \mathcal{A}}{\partial t} + \nabla \cdot (\mathbf{c}_g \mathcal{A}) = 0, \tag{166}$$

where $\mathbf{c}_g = \nabla_\mathbf{k} \omega$ is the IGW group velocity, and $\mathcal{A} = E_w/\widehat{\omega}$ the IGW wave action, with

$$E_w = \frac{\overline{\rho}}{2} \left(\frac{|\mathbf{u}'_1|^2}{2} + \frac{|b'_1|^2}{2N^2} \right) \tag{167}$$

the wave energy. Here $\overline{\rho}$ is the reference-atmosphere density. This equation, together with the ray-tracing equations

$$\left(\frac{\partial}{\partial t} + \mathbf{c}_g\cdot\nabla\right)\mathbf{k} = -(\nabla\mathbf{U})\cdot\mathbf{k} + \frac{k^2 + l^2}{2\widehat{\omega} m^2}\frac{\mathrm{d}N^2}{\mathrm{d}z}\mathbf{e}_z \tag{168}$$

describes the mean-flow impact on the IGW amplitudes and wave numbers. The IGW impact on the mean flow is given by the potential-vorticity equation

$$\left(\frac{\partial}{\partial t} + \mathbf{U}\cdot\nabla_h\right) P = -\frac{\partial}{\partial x}\left[\frac{1}{\overline{\rho}}\nabla\cdot(\widehat{\mathbf{c}}_g l \mathcal{A})\right] + \frac{\partial}{\partial y}\left[\frac{1}{\overline{\rho}}\nabla\cdot(\widehat{\mathbf{c}}_g k \mathcal{A})\right], \tag{169}$$

with $\widehat{\mathbf{c}}_g = \nabla_\mathbf{k}\widehat{\omega}$ the intrinsic group velocity. The leading-order synoptic-scale PV is

$$P = \nabla_h^2 \Psi + \frac{1}{\overline{\rho}}\frac{\partial}{\partial z}\left(\overline{\rho}\frac{f^2}{N^2}\frac{\partial\Psi}{\partial z}\right), \tag{170}$$

with $\Psi = c_p\overline{\theta}_0\Pi/f$ the streamfunction, where $\Pi = \varepsilon^{1+\alpha}\Pi_0^{(0)}$ is the leading-order synoptic-scale Exner pressure, and $\overline{\theta}_0 = T_{00}\overline{\Theta}^{(0)}$ is the leading-order reference-atmosphere potential temperature. The latter depends on z only in the case with moderately strong stratification ($\alpha = 0$), while it is a constant in the weakly stratified case ($\alpha = 1$). The streamfunction also yields the leading-order synoptic-scale horizontal wind $\mathbf{U}$, via geostrophic equilibrium,

$$\mathbf{U} = \mathbf{e}_z \times \nabla\Psi, \tag{171}$$

and the leading-order synoptic-scale potential temperature fluctuations $\Theta = \varepsilon^{1+\alpha} T_{00}\Theta_0^{(0)}$, via hydrostatic equilibrium,

$$g\frac{\Theta}{\overline{\theta}_0} = f\frac{\partial\Psi}{\partial z} + \begin{cases} g\left(\dfrac{\overline{\theta} - \overline{\theta}_0}{\overline{\theta}_0}\right)^2 & \text{if } \alpha = 1, \\ -N^2 f\Psi/g & \text{if } \alpha = 0. \end{cases} \tag{172}$$

Moreover, re-dimensionalization of (152) leads to

$$\left(\frac{\partial}{\partial t} + \mathbf{U}\cdot\nabla_h\right)\left(P - \frac{\mathbf{e}_z}{\overline{\rho}}\cdot\nabla\times\mathbf{p}_h\right) = 0, \tag{173}$$

where $\mathbf{p}_h = \mathbf{k}_h\mathcal{A}$ is the wave pseudo-momentum. In agreement with Bühler and McIntyre (1998, 2003, 2005) and Wagner and Young (2015), this shows that the theory respects the conservation of total PV,

$$P_{tot} = P - \frac{\mathbf{e}_z}{\overline{\rho}}\cdot\nabla\times\mathbf{p}_h, \tag{174}$$

consisting of the synoptic-scale-flow part P and a wave contribution from the vertical curl of pseudo-momentum. An interesting difference to the studies referred to above and also Grimshaw (1975b) is that this result is formulated there in terms of a Lagrangian-mean synoptic-scale flow, whereas our study does not use this kind of average. However, as Grimshaw (1975b) shows, the difference between Eulerian-mean and Lagrangian-mean flow, the Stokes drift, is $\mathcal{O}(\varepsilon)$, and corresponding differences would show up only in higher-order terms that are not of prior relevance here.

Finally we also note an energy-conservation theorem. From the dimensional variant of (119),

$$\left(\frac{\partial}{\partial t} + \mathbf{c}_g\cdot\nabla\right)\widehat{\omega} = -\mathbf{k}\widehat{\mathbf{c}}_g\cdot\cdot\nabla\mathbf{U} \tag{175}$$

and the wave-action equation (166), one obtains

$$\frac{\partial E_w}{\partial t} + \nabla \cdot (\mathbf{c}_g E_w) = -\mathcal{A}\mathbf{k}\widehat{\mathbf{c}}_g \cdot \cdot \nabla \mathbf{U}. \tag{176}$$

Multiplying (169) by $-\overline{\rho}\Psi$ yields, with repeated use of (171),

$$\frac{\partial E_s}{\partial t} - \nabla_h \cdot \left[\overline{\rho}\Psi\left(\frac{\partial}{\partial t}\nabla_h\Psi + \mathbf{U}P\right)\right] - \frac{\partial}{\partial z}\left(\overline{\rho}\Psi\frac{f^2}{N^2}\frac{\partial}{\partial t}\frac{\partial\Psi}{\partial z}\right)$$
$$= \frac{\partial}{\partial x}\left[\Psi\nabla\cdot\left(\widehat{\mathbf{c}}_g l\mathcal{A}\right)\right] - \frac{\partial}{\partial y}\left[\Psi\nabla\cdot\left(\widehat{\mathbf{c}}_g k\mathcal{A}\right)\right]$$
$$- \nabla\cdot\left(\widehat{\mathbf{c}}_g\mathbf{k}\cdot\mathbf{U}\mathcal{A}\right) + \mathcal{A}\mathbf{k}\widehat{\mathbf{c}}_g\cdot\cdot\nabla\mathbf{U}, \tag{177}$$

where

$$E_s = \frac{\overline{\rho}}{2}\left[|\nabla_h\Psi|^2 + \frac{f^2}{N^2}\left(\frac{\partial\Psi}{\partial z}\right)^2\right] \tag{178}$$

is the energy density of the synoptic-scale flow. Hence

$$\frac{\partial}{\partial t}(E_s + E_w) = \nabla_h\cdot\left[\overline{\rho}\Psi\left(\frac{\partial}{\partial t}\nabla_h\Psi + \mathbf{U}P\right)\right] + \frac{\partial}{\partial z}\left(\overline{\rho}\Psi\frac{f^2}{N^2}\frac{\partial}{\partial t}\frac{\partial\Psi}{\partial z}\right)$$
$$- \nabla\cdot\left(\widehat{\mathbf{c}}_g\mathbf{k}\cdot\mathbf{U}\mathcal{A} + \mathbf{c}_g E_w\right)$$
$$+ \frac{\partial}{\partial x}\left[\Psi\nabla\cdot\left(\widehat{\mathbf{c}}_g l\mathcal{A}\right)\right] - \frac{\partial}{\partial y}\left[\Psi\nabla\cdot\left(\widehat{\mathbf{c}}_g k\mathcal{A}\right)\right], \tag{179}$$

so that the total of synoptic-scale-flow energy and wave energy is conserved under usual boundary conditions.

8.3. *Geostrophic-mode dynamics*

The dimensional GM potential-enstrophy equations are

$$\left(\frac{\partial}{\partial t} + \mathbf{U}\cdot\nabla_h\right)\mathcal{P}_\beta = \frac{\partial\mathcal{P}_\beta}{\partial t} + \nabla\cdot(\mathbf{U}\mathcal{P}_\beta)$$
$$= \frac{\overline{\rho}}{4}\Re\sum_{\beta'=1}^{\infty}\sum_{\beta''=1}^{\infty}\left\{P_\beta{}^* D\left(\frac{\beta''}{\beta'}, \mathbf{u}'_{\beta'}\right)P_{\beta''}\delta\left(\beta' + \beta'' - \beta\right)\right.$$
$$+ P_\beta{}^* D\left(-\frac{\beta''}{\beta'}, \mathbf{u}'_{\beta'}\right)P_{\beta''}{}^*\delta\left(\beta' - \beta'' - \beta\right)$$
$$\left.+ P_\beta{}^* D\left(-\frac{\beta''}{\beta'}, \mathbf{u}'_{\beta'}{}^*\right)P_{\beta''}\delta\left(-\beta' + \beta'' - \beta\right)\right\}, \tag{180}$$

where

$$D\left(\lambda, \mathbf{u}'_\beta\right) = \lambda\left(\nabla\cdot\mathbf{u}'_\beta\right) - \mathbf{u}'_\beta\cdot\nabla, \tag{181}$$

and where $\mathcal{P}_\beta = \overline{\rho}|P_\beta|^2/4$ is the leading-order potential enstrophy of the βth GM harmonic, with $P_\beta = (\beta^2|\mathbf{k}_h|^2 + (f^2/N^2)\beta^2 m^2)\psi_\beta$ the corresponding leading-order PV amplitude, and $\psi_\beta = c_p\overline{\theta}_0\pi'_\beta/f$ the corresponding streamfunction. This is supplemented by the ray-tracing equations

$$\left(\frac{\partial}{\partial t} + \mathbf{U}\cdot\nabla\right)\mathbf{k} = -(\nabla\mathbf{U})\cdot\mathbf{k}$$

to describe the mean-flow impact on the GM. The GM impact on the mean flow is determined by the synoptic-scale potential-vorticity equation

$$\left(\frac{\partial}{\partial t} + \mathbf{U}\cdot\nabla_h\right)P = -\frac{\partial}{\partial x}\left(\frac{1}{\overline{\rho}}\nabla\cdot\sum_{\beta=1}^{\infty}\widehat{\mathbf{c}}_{\beta,\gamma}\beta\, l\mathcal{P}_\beta\right)$$
$$+ \frac{\partial}{\partial y}\left(\frac{1}{\overline{\rho}}\nabla\cdot\sum_{\beta=1}^{\infty}\widehat{\mathbf{c}}_{\beta,\gamma}\beta k\mathcal{P}_\beta\right), \tag{182}$$

with $\widehat{\mathbf{c}}_{\beta,\gamma} = \nabla_{\beta\mathbf{k}}\widehat{\gamma}_\beta$, where

$$\widehat{\gamma}_\beta(\beta\mathbf{k}) = \frac{N^2}{N^2\beta^2|\mathbf{k}_h|^2 + f^2\beta^2 m^2}, \tag{183}$$

so that $P_\beta = \psi_\beta/\widehat{\gamma}_\beta$.

8.4. *Synopsis and implications for subgrid-scale modelling*

To summarize, in the absence of IGW fluxes and GM fluxes, the synoptic-scale PV is conserved, and quasi-geostrophic theory holds, both in the weakly stratified case ($\alpha = 1$) for which this result is standard knowledge from text books (Pedlosky, 1987), and the case of moderately strong stratification ($\alpha = 0$). Otherwise it is controlled by the curl of the vector of divergences of the fluxes $\widehat{\mathbf{c}}_g k\mathcal{A}$, $\widehat{\mathbf{c}}_g l\mathcal{A}$, $\sum_\beta\widehat{\mathbf{c}}_{\beta,\gamma}\beta k\mathcal{P}_\beta$ and $\sum_\beta\widehat{\mathbf{c}}_{\beta,\gamma}\beta l\mathcal{P}_\beta$. In the case of IGW–mean-flow interactions, this forcing vector results from changes in the vertical curl of the IGW pseudo-momentum. Inverting the PV to obtain the streamfunction yields all information necessary to obtain the synoptic-scale fields by geostrophy and hydrostaticity. In the IGW case we thus have a fully coupled system where the wave properties can be predicted by the eikonal equations, wave-action conservation, and the polarization relations, and where the synoptic-scale flow is controlled by the potential-vorticity equation obtained above. In the GM case, all upper harmonics contribute to the fluxes controlling the synoptic-scale PV. Their amplitude is predicted by respective potential-enstrophy equations.

The practitioner will typically not consider the synoptic-scale-flow PV, but rather want to insert the relevant mesoscale-wave fluxes directly into the prognostic equations of an NWP code or climate model. How this can be done is discernible from the synoptic-scale-flow momentum, entropy and pressure equation in section 5. Wave fluxes actually only appear in the entropy equation (92) and in the horizontal momentum equation (95). Only horizontal entropy fluxes arise so that the dimensional synoptic-scale-flow entropy equation becomes

$$\left(\frac{\partial}{\partial t} + \mathbf{U}\cdot\nabla_h\right)\theta + N^2 W = -\frac{1}{2}\nabla_h\cdot\Re\sum_{\beta=1}^{\infty}\mathbf{u}'_\beta\theta'^*_\beta. \tag{184}$$

The entropy-flux convergence on the right-hand side is standard, but takes also GM impacts into account. From the polarization relations and the definition of wave action and potential enstrophy, respectively, one can re-express the relevant fluxes in terms of the predicted fields as

$$\frac{1}{2}\Re\left(\mathbf{u}'_1\theta'^*_1\right) = -\mathbf{e}_z\times\frac{f}{\widehat{\omega}}\frac{\overline{\theta}}{g}N^2\frac{\mathbf{k}_h}{m}\frac{\mathcal{A}}{\overline{\rho}} \tag{185}$$

in the IGW case, and

$$\frac{1}{2}\Re\left(\mathbf{u}'_\beta\theta'^*_\beta\right) = \mathbf{e}_z\times 2\frac{\overline{\theta}}{g}N^4\frac{\mathbf{k}_h mf}{\beta^2\left(N^2|\mathbf{k}_h|^2 + f^2m^2\right)^2}\frac{\mathcal{P}_\beta}{\overline{\rho}} \tag{186}$$

for GMs.

More interesting is the horizontal momentum equation (95). The two flux terms appearing are a *pseudo-incompressible* momentum-flux convergence and an *elastic term arising from the potential-temperature fluctuations in the pressure gradient term* (or equivalently the density fluctuations, if one prefers $-(\nabla p)/\rho$). Splitting in the former $\overline{P}^{(0)} = \overline{R}^{(0)}\overline{\Theta}^{(0)}$, and using in the IGW case (116) and (149), or using in the GM case $W^{(0)}_\beta = 0$ and the polarization relations (76)–(77), one can rewrite these, and finally re-dimensionalize them, as

$$
\begin{aligned}
&-\frac{1}{2}\Re\frac{1}{\overline{P}^{(0)}}\nabla_X\cdot\left(\overline{P}^{(0)}\sum_{\beta=1}^{\infty}\mathbf{V}_\beta^{(0)}\mathbf{U}_\beta^{(0)*}\right)\\
&\quad+\frac{1-\alpha}{2}\frac{c_p}{R}\Re\sum_{\beta=1}^{\infty}\mathrm{i}\beta\mathbf{k}_h\Theta_\beta^{(0)}\Pi_\beta^{(0)*}\\
&=-\frac{1}{2}\Re\frac{1}{\overline{R}^{(0)}}\nabla_X\cdot\left(\overline{R}^{(0)}\sum_{\beta=1}^{\infty}\mathbf{V}_\beta^{(0)}\mathbf{U}_\beta^{(0)*}\right)\\
&\quad+\frac{1-\alpha}{2}f_0\mathbf{e}_z\times\Re\sum_{\beta=1}^{\infty}\mathbf{U}_\beta^{(0)*}B_\beta^{(0)}\\
&\to-\frac{1}{\overline{\rho}}\nabla\cdot\left(\frac{\overline{\rho}}{2}\Re\sum_{\beta=1}^{\infty}\mathbf{v}'_\beta\mathbf{u}'^*_\beta\right)+\frac{f}{2g}\mathbf{e}_z\times\Re\sum_{\beta=1}^{\infty}\mathbf{u}'^*_\beta b'_\beta. \qquad (187)
\end{aligned}
$$

The former term is the classic anelastic momentum-flux convergence, supplemented by an additional elastic term that cannot be derived from anelastic theory. Using the polarization relations, we again re-express everything required in terms of the explicitly predicted fields as

$$\frac{\overline{\rho}}{2}\Re\left(\mathbf{u}'_1\mathbf{u}'^*_1\right)=\frac{\mathbf{k}_h\mathbf{k}_h\widehat{\omega}^2+(\mathbf{e}_z\times\mathbf{k}_h)(\mathbf{e}_z\times\mathbf{k}_h)f^2}{|\mathbf{k}_h|^2\widehat{\omega}}\mathcal{A}, \qquad (188)$$

$$\frac{\overline{\rho}}{2}\Re\left(\mathbf{u}'_1 w'^*_1\right)=-\frac{\mathbf{k}_h}{m}\widehat{\omega}\mathcal{A}=\frac{c_{gz}\mathbf{k}_h\mathcal{A}}{1-\widehat{\omega}^2/f^2}, \qquad (189)$$

$$\frac{f}{2g}\mathbf{e}_z\times\Re\left(\mathbf{u}'^*_1 b'_1\right)=\frac{f}{g}\frac{f}{\widehat{\omega}}\frac{\mathbf{k}_h}{m}N^2\frac{\mathcal{A}}{\overline{\rho}}, \qquad (190)$$

for IGWs and

$$\frac{\overline{\rho}}{2}\Re\left(\mathbf{u}'_\beta\mathbf{u}'^*_\beta\right)=2N^4\frac{(\mathbf{e}_z\times\mathbf{k}_h)\,(\mathbf{e}_z\times\mathbf{k}_h)}{\beta^2\left(N^2|\mathbf{k}_h|^2+f^2m^2\right)^2}\mathcal{P}_\beta, \qquad (191)$$

$$\frac{\overline{\rho}}{2}\Re\left(\mathbf{u}'_\beta w'^*_\beta\right)=0, \qquad (192)$$

$$\frac{f}{2g}\mathbf{e}_z\times\Re\left(\mathbf{u}'^*_\beta b'_\beta\right)=-\frac{2f}{g}N^4\frac{\mathbf{k}_h mf}{\beta^2\left(N^2|\mathbf{k}_h|^2+f^2m^2\right)^2}\frac{\mathcal{P}_\beta}{\overline{\rho}} \qquad (193)$$

for GMs. Grimshaw (1975b) also shows the elastic term in his equations for the IGW case, but then moves to Lagrangian-mean theory. Nonetheless, it should be stressed that models with a Eulerian formulation should take it into account, together with the anelastic momentum-flux convergence, as wave forcing of the synoptic-scale flow. Using the expressions above, one can illustrate the relevance of the elastic term by the ratios

$$\frac{(f/g)\mathbf{e}_z\times\Re(\mathbf{u}'^*_1 b'_1)}{(1/\overline{\rho})(\partial/\partial z)\left[\overline{\rho}\Re(w'_1\mathbf{u}'^*_1)\right]}=\mathcal{O}\left(\frac{f^2}{\widehat{\omega}^2}\frac{H_s}{H_\theta}\right) \qquad (194)$$

in the IGW case, and

$$\frac{(f/g)\mathbf{e}_z\times\Re(\mathbf{u}'^*_1 b'_1)}{\nabla_h\cdot\Re\left(\mathbf{u}'_1\mathbf{u}'^*_1\right)}=\mathcal{O}\left(\frac{f^2}{N^2}\frac{m}{|\mathbf{k}_h|}\frac{L_s}{H_\theta}\right) \qquad (195)$$

in the GM case. In the scaling regime considered here, these are both $\mathcal{O}(1)$, but one also sees that the elastic term loses relevance in the weakly stratified regime, i.e. where H_θ is larger, and that it is most relevant for low-frequency IGWs in a strongly stratified regime, and for GMs with small vertical wavelength as compared to the horizontal wavelength, and large horizontal synoptic scale as compared to the potential-temperature scale height.

9. Discussion

Our reconsideration and review of the interaction between synoptic-scale flow and mesoscale wave packets is based on a detailed scale analysis. A review of the basic assumptions of quasi-geostrophic theory for synoptic-scale flow on an f-plane shows that all relevant scales can be determined from the Rossby number ε and two out of the three following variables: gravitational acceleration g, inertial frequency f_0 and sound-related speed $c_s=\sqrt{RT_{00}}$, with T_{00} a typical temperature value. The wave scaling is then defined by requiring the spatial- and time-scales to be shorter by $\mathcal{O}(\varepsilon)$, and by assuming their buoyancy field to be close to static instability. In the latter considerations, we have applied Boussinesq polarization relations, as justified by the results of the following analysis. Two stratification regimes are considered, by assuming that the potential-temperature scale height either is larger, by $\mathcal{O}(\varepsilon^{-1})$, than both the density and the pressure scale height (tropospheric regime of weak stratification), or is of the same order as those two (stratospheric regime of moderately strong stratification). After a non-dimensionalization of the equations of motion, a WKB ansatz is introduced for the wave fields, allowing a basic-wave field and all nonlinearly induced higher harmonics. All synoptic-scale fields and wave amplitudes are then expanded in terms of the Rossby number. Ordering by powers of the latter and the WKB phase factor then yields all results.

These re-establish the geostrophic and hydrostatic balance of the synoptic-scale flow. They also lead to eikonal equations for wavenumber and frequency, both for inertia-gravity waves (IGW) and geostrophic modes (GM). These results hold at finite wave amplitudes, i.e. close to the threshold of static instability. No explicit linearization of the equations is necessary that would require weak wave amplitudes. It is the scale separation between the wave phases on the one hand and large-scale flow and wave amplitudes on the other, combined with a hence derived solenoidality of the wave velocity fields, that removes the nonlinearities from the leading-order equations.

To next order one finds that, due to their dispersive nature, the IGW higher harmonics must be one order of magnitude weaker than the IGW basic wave. They are slaved to the basic wave, and their amplitude can be determined directly from the basic-wave dynamics. However, the GM higher harmonics are found to be as strong as the basic wave. Amplitude equations are derived for IGWs and GMs that describe, together with the eikonal equations, the mean-flow impact on the waves. The dynamics of the higher harmonics, both for IGWs and GMs, is an aspect of the present finite-amplitude theory that had not been derived before from weak-amplitude theories.

The IGW amplitude equation is the well-known wave-action conservation equation, with wave action only due to the basic wave, while one obtains in the GM case a potential-enstrophy equation for each harmonic, with a nonlinear triad term describing the interaction between different harmonics. This is found for both stratification regimes, and it implies a lower degree of stability for GM wave packets than IGW wave packets. Potentially this might contribute to the lower energy in mesoscale GMs, as compared to IGWs, in the upper troposphere, as reported by Callies *et al.* (2014). However, recent work by Lindborg (2015) and Bierdel *et al.* (2016) indicates that the GM contribution to mesoscale energy might be significant in various atmospheric regions. This would further support the relevance of investigations of GM dynamics.

The analysis of the wave-impact on the synoptic-scale flow yields similar results for both stratification regimes. However the route there differs between the two cases. In the stratospheric regime of moderately strong stratification, various elastic terms appear in the mean-flow equations, demonstrating the unsafe ground one would be on if one neglected everything beyond Boussinesq or anelastic dynamics from the start. An implementation of a WKB ray tracer into weather forecast or climate models would typically supplement a model by the wave-flux terms appearing in the entropy equation (92) and the horizontal-momentum equations (95). However, the latter need in the stratospheric regime an elastic flux term arising from the potential-temperature fluctuations in the pressure-gradient term. This term, appearing as a Coriolis force due to a non-zero mass or buoyancy flux (Grimshaw, 1975b), would supplement an anelastic momentum-flux convergence. In the investigated regime, it is of the same magnitude as the latter, and it gains in importance the stronger the stratification is, preferentially for low-frequency IGWs, and for GMs in a flow with large horizontal scales. Nonetheless, both regimes yield in the end a prognostic equation for quasi-geostrophic potential vorticity (PV), with a wave impact from either IGWs or GMs. For the first time, to the best of our knowledge, we thus show that, in the absence of IGWs and GMs, quasi-geostrophic theory strictly holds at low Rossby numbers also for moderately strong stratification. The derivation coming nearest to this, to the best of our knowledge, has been indicated by Zeitlin *et al.* (2003), but these authors only consider the Boussinesq equations.

At least in the IGW case, the theory also respects the conservation of total PV, consisting of a contribution of synoptic-scale flow and a wave contribution from the vertical curl of wave pseudo-momentum, discussed previously in the IGW context by Bühler and McIntyre (1998, 2005) and Wagner and Young (2015). The sum of wave energy and the energy of the synoptic-scale flow is conserved as well.

Our study is also related to recent work by Xie and Vanneste (2015) and Wagner and Young (2016) on the interaction between near-inertial waves and synoptic-scale flow in the ocean. However the scaling regime investigated there is different. In those studies there is no horizontal length-scale separation. The vertical-scale separation parameter is $\varepsilon^{1/2}$. They assume the synoptic-scale horizontal winds to be weaker than those from the waves. The considered stratification is considerably weaker than considered here, with $f/N = \mathcal{O}(\varepsilon^{1/2})$. Finally, the considered wave field is linear, while our theory considers nonlinear waves. Our study is therefore mostly complementary to those.

The GM dynamics we have investigated might be of relevance for the modelling of subgrid-scale dynamics. Since the GMs are simply advected by the synoptic-scale flow, on top of the nonlinear triad interactions, they do not contribute to vertical coupling in the atmosphere. However, their contribution to horizontal coupling could be important. Although geostrophic, the Rossby number of these modes is large, i.e. $U_{\mathrm{w}}/fL_{\mathrm{w}} = \mathcal{O}(1)$. Nonetheless, their dynamics, as far as we have followed it, can be derived from quasi-geostrophic theory as well, as shown in the Appendix B. Notwithstanding, both their potential-enstrophy equations, and their impact on the synoptic-scale flow, via the vertical curl of an Eliassen–Palm-flux convergence, differ from the quasi-geostrophic dynamics of the interaction between synoptic-scale Rossby waves and a planetary-scale mean flow. The difference in scale between planetary-synoptic versus synoptic-mesoscale interactions seems to be responsible for this discrepancy. It might be worthwhile stressing that our derivations do not lead to an extension of quasi-geostrophic theory to describe the dynamics of mesoscale GMs in interaction with a synoptic-scale flow, as all essentials seem to be embedded in that theory. Notwithstanding the issue of the parametrization of unresolved mesoscale modes, this arises for simulations at sufficiently coarse resolution. This is the merit and purpose of results on GM dynamics and GM–mean-flow interactions in this study.

It might also be remarked that Generalized Lagrangian-Mean (GLM) theory could not be used for studying GM dynamics. That theory assumes a displacement vector ξ so that $\mathrm{D}\xi/\mathrm{D}t = \mathbf{v}$, or $-\mathrm{i}\widehat{\omega}\xi = \mathbf{v}$ in the linear limit. Since the intrinsic frequency of the GM is zero, its displacement vector is not defined. Multi-scale asymptotics, as performed here, do not have this limitation. Thus it is a useful supplementary tool to GLM theory, leaving the latter its undisputed claim for elegance and generality with regard to the dynamics of IGWs and Rossby waves.

The theory as a whole is nonlinear, with a two-way interaction between finite-amplitude waves and mean flow, and a full consideration of the interaction with and between all nonlinearly induced higher harmonics, found to be negligibly weak in the IGW case, but not in the GM case. Processes resulting from the interaction between waves and a self-induced mean wind (Fritts and Dunkerton, 1984; Sutherland, 2001, 2006; Dosser and Sutherland, 2011) are included in such formulations, as demonstrated, e.g. by Rieper *et al.* (2013) and Muraschko *et al.* (2015). Our results also apply to the interaction between a synoptic-scale flow and small-amplitude wave fields. In the present two-time-scale theory, the wave impact would disappear in this case, as it would be weaker by two orders of ε (e.g. Achatz *et al.*, 2010), but this only implies that the appropriate approach would then be the introduction of a new longer time-scale T_s/ε^2 on which the wave forcing would influence the mean flow. The final results we expect to be the same as presented here. Moreover, the higher harmonics would be suppressed significantly. They are a central result of the present finite-amplitude theory.

An apparent limitation is that our analysis assumes a single basic-wave field, locally monochromatic, superposed by higher harmonics. As soon as various basic-wave fields are allowed, nonlinear interaction terms would supplement even the IGW wave action equation. The wave impact on the large-scale flow also would appear then as the superposition of the wave impacts derived here separately. This approach would eventually imply the use of phase-space wave-action densities (e.g. Bühler and McIntyre, 1999; Hertzog *et al.*, 2002), seemingly a rather powerful tool for the avoidance of numerical instabilities due to crossing rays (Muraschko *et al.*, 2015). It often compares successfully to wave-resolving data, even when nonlinear IGW interactions are neglected. How relevant the latter will be in the end is an open question. Atmospheric waves lead a rather transient life, which might often be too short for nonlinear effects to have a strong impact. However, measurements of atmospheric mesoscale spectra might indicate nonlinear dynamics (Callies *et al.*, 2014; Zhang *et al.*, 2015). For the investigation of corresponding processes, a weakly nonlinear low-amplitude approach might be useful (e.g. Caillol and Zeitlin, 2000; Nazarenko, 2011), as it might be able to yield tractable results. Corresponding investigations seem to be an important line of future research. Another relevant extension could be the consideration of the interaction of small-scale waves, possibly in coexistence with turbulence, with a larger-scale flow containing considerable unbalanced contributions. This might be of interest for subgrid-scale parametrizations in climate and weather forecast models with mesoscale resolution.

Appendices

Appendix A: Reformulation of the nonlinear triad term appearing in the potential-enstrophy dynamics of the geostrophic mode

For the reformulation of the nonlinear triad term (103), we first rewrite the polarization relations (76)–(78), using $\widehat{\omega} = 0$,

$$\mathbf{U}_\beta^{(0)} = \mathbf{e}_z \times \mathrm{i}\beta\mathbf{k}_h\Psi_\beta^{(0)}, \tag{A1}$$

$$W_\beta^{(0)} = 0, \tag{A2}$$

$$B_\beta^{(0)} = \mathrm{i}f_0\beta m\Psi_\beta^{(0)}, \tag{A3}$$

where

$$\Psi_\beta^{(0)} = \frac{c_p}{R}\overline{\Theta}^{(0)}\Pi_\beta^{(0)}/f_0 \tag{A4}$$

is a non-dimensional streamfunction. Quasi-geostrophic theory would imply a corresponding leading-order potential vorticity

$$P_\beta^{(0)} = -\left(!\beta^2\,|\mathbf{k}_h|^2 + \frac{f_0^2}{N_0^2}\beta^2 m^2\right)\Psi_\beta^{(0)} = -\frac{\Psi_\beta^{(0)}}{\widehat{\gamma}_\beta}. \tag{A5}$$

In addition, due to the zero vertical-wind amplitude, one replaces

$$D\left(\lambda, \mathbf{V}_\beta^{(0)}\right) \to D\left(\lambda, \mathbf{U}_\beta^{(0)}\right) = \lambda\left(\nabla_{X,h}\cdot\mathbf{U}_\beta^{(0)}\right) - \mathbf{U}_\beta^{(0)}\cdot\nabla_{X,h}, \tag{A6}$$

where, due to $\nabla_X \times \mathbf{k} = 0$,

$$\nabla_{X,h}\cdot\mathbf{U}_\beta^{(0)} = (\mathbf{e}_z \times \mathrm{i}\beta\mathbf{k}_h)\cdot\nabla_{X,h}\Psi_\beta^{(0)}. \tag{A7}$$

All of this is inserted into (103). In the ensuing algebra one makes repeated use of the triad conditions, expressed by the delta functions, and uses replacements of the kind

$$\begin{aligned}\sum_{\beta',\beta''=1}^{\infty} &\beta\beta'\beta''\Psi_{\beta'}^{(0)}\Psi_{\beta''}^{(0)}\delta\left(\beta'+\beta''-\beta\right)\\ &= \sum_{\beta',\beta''=1}^{\infty}\left(\beta'+\beta''\right)\beta'\beta''\Psi_{\beta'}^{(0)}\Psi_{\beta''}^{(0)}\delta\left(\beta'+\beta''-\beta\right)\\ &= \sum_{\beta',\beta''=1}^{\infty} 2\beta'^2\beta''\Psi_{\beta'}^{(0)}\Psi_{\beta''}^{(0)}\delta\left(\beta'+\beta''-\beta\right),\end{aligned} \tag{A8}$$

$$\begin{aligned}\sum_{\beta',\beta''=1}^{\infty} &\beta'\beta''^2\Psi_{\beta'}^{(0)}\Psi_{\beta''}^{(0)*}\delta\left(\beta'-\beta''-\beta\right)\\ &= \sum_{\beta',\beta''=1}^{\infty}\beta'^2\beta''\Psi_{\beta'}^{(0)*}\Psi_{\beta''}^{(0)}\delta\left(-\beta'+\beta''-\beta\right),\end{aligned} \tag{A9}$$

to finally obtain (132).

Appendix B: Mesoscale geostrophic-mode dynamics derived from quasi-geostrophic theory

For the analysis of the interaction between geostrophic synoptic-scale flow and mesoscale geostrophic modes within quasi-geostrophic theory, we use the corresponding PV conservation equation on an f-plane

$$\frac{\partial P}{\partial t} + \nabla\cdot(\mathbf{u}\mathbf{P}) = 0, \tag{B1}$$

with

$$\mathbf{u} = \mathbf{e}_z \times \nabla_h\psi, \tag{B2}$$

$$P = \nabla_h^2\psi + \frac{1}{\overline{\rho}}\frac{\partial}{\partial z}\left(\overline{\rho}\frac{f^2}{N^2}\frac{\partial\psi}{\partial z}\right). \tag{B3}$$

Here

$$\psi = c_p\overline{\theta}_0\left(\pi - \overline{\pi}\right)/f \tag{B4}$$

is the streamfunction, with $\overline{\theta}_0 = T_{00}\overline{\Theta}^{(0)}$ the leading-order part of the reference atmosphere, and $\overline{\pi}$ the reference-atmosphere Exner pressure. $\overline{\rho} = \rho_{00}\overline{R}^{(0)}$ is the leading-order reference-atmosphere density, and $N^2 = \varepsilon^\alpha\left(g/\overline{\Theta}^{(0)}\right)\mathrm{d}\overline{\Theta}^{(\alpha)}/\mathrm{d}z$ the Brunt–Väisälä frequency. Non-dimensionalizing the streamfunction by the wave scales yields the replacement, using (56) and the definitions in section 2,

$$\begin{aligned}\psi \to \frac{c_p\overline{\theta}_0\left(\pi-\overline{\pi}\right)}{fU_\mathrm{w}L_\mathrm{w}} &= \frac{c_p}{R}\frac{\overline{\Theta}^{(0)}}{\varepsilon^{2+\alpha}}\left(\pi - \sum_{j=0}^{\alpha}\varepsilon^j\overline{\Pi}^{(j)}\right)\\ &= \varepsilon^{-1}\sum_{j=0}^{\infty}\varepsilon^j\Psi_0^{(j)}(\mathbf{X},T) + \Re\sum_{\beta=1}^{\infty}\sum_{j=0}^{\infty}\varepsilon^j\Psi_\beta^{(j)}(\mathbf{X},T)\mathrm{e}^{\mathrm{i}\beta\phi(\mathbf{X},T)/\varepsilon},\end{aligned} \tag{B5}$$

where $\Psi_0^{(j)} = (c_p/R)\overline{\Theta}^{(0)}\Pi_0^{(j)}$ and $\Psi_\beta^{(j)} = (c_p/R)\overline{\Theta}^{(0)}\Pi_\beta^{(j)}$ are the various-order synoptic-scale and mesoscale (basic and higher harmonic) streamfunction components. After non-dimensionalization, also by the wave scales L_w, H_w, U_w and $T_\mathrm{w} = L_\mathrm{w}/U_\mathrm{w}$, the PV conservation equation (B1) and the velocity equation (B2) keep their form, while the non-dimensional PV becomes

$$P = \nabla_h^2\psi + \frac{1}{\overline{R}^{(0)}}\frac{\partial}{\partial z}\left(\overline{R}^{(0)}\frac{f_0^2}{N_0^2}\frac{\partial\psi}{\partial z}\right). \tag{B6}$$

Inserting (B5), one obtains

$$P = \varepsilon\sum_{j=0}^{\infty}\varepsilon^jP_0^{(j)}(\mathbf{X},T) + \Re\sum_{\beta=1}^{\infty}\sum_{j=0}^{\infty}\varepsilon^jP_\beta^{(j)}(\mathbf{X},T)\mathrm{e}^{\mathrm{i}\beta\phi(\mathbf{X},T)/\varepsilon} \tag{B7}$$

with

$$P_0^{(j)} = \nabla_{X,h}^2\Psi_0^{(j)} + \frac{1}{\overline{R}^{(0)}}\frac{\partial}{\partial Z}\left(\overline{R}^{(0)}\frac{f_0^2}{N_0^2}\frac{\partial\Psi_0^{(j)}}{\partial Z}\right) \tag{B8}$$

and where

$$P_\beta^{(0)} = -\left(\beta^2\left|\mathbf{k}_h\right|^2 + \frac{f_0^2}{N_0^2}\beta^2m^2\right)\Psi_\beta^{(0)}, \tag{B9}$$

$$\begin{aligned}P_\beta^{(1)} = &-\left(\beta^2\left|\mathbf{k}_h\right|^2 + \frac{f_0^2}{N_0^2}\beta^2m^2\right)\Psi_\beta^{(1)}\\ &+\mathrm{i}\left[2\beta\mathbf{k}_h\cdot\nabla_{X,h} + \left(\nabla_{X,h}\cdot\beta\mathbf{k}_h\right) + 2\beta m\frac{f_0^2}{N_0^2}\frac{\partial}{\partial Z}\right.\\ &\left.+\frac{1}{\overline{R}^{(0)}}\frac{\partial}{\partial Z}\left(\overline{R}^{(0)}\frac{f_0^2}{N_0^2}\beta m\right)\right]\Psi_\beta^{(0)}\end{aligned} \tag{B10}$$

are the PV wave amplitudes explicitly needed below. Likewise the non-dimensional wind is

$$\mathbf{u} = \sum_{j=0}^{\infty}\varepsilon^j\mathbf{U}_\mathbf{0}^{(\mathbf{j})}(\mathbf{X},\mathbf{T}) + \Re\sum_{\beta=1}^{\infty}\sum_{j=0}^{\infty}\varepsilon^\mathbf{j}\mathbf{U}_\beta^{(\mathbf{j})}(\mathbf{X},\mathbf{T})\mathrm{e}^{\mathrm{i}\beta\phi(\mathbf{X},\mathbf{T})/\varepsilon} \tag{B11}$$

with

$$\mathbf{U}_0^{(j)} = \mathbf{e}_z \times \nabla_{X,h}\Psi_0^{(j)} \tag{B12}$$

and for $\beta \geq 1$

$$\mathbf{U}_\beta^{(0)} = \mathbf{e}_z \times \mathrm{i}\beta\mathbf{k}_h\Psi_\beta^{(0)}, \tag{B13}$$

$$\mathbf{U}_\beta^{(j\geq 1)} = \mathbf{e}_z \times \left(\mathrm{i}\beta\mathbf{k}_h\Psi_\beta^{(j)} + \nabla_{X,h}\Psi_\beta^{(j-1)}\right). \tag{B14}$$

Obviously

$$\mathbf{k}_h\cdot\mathbf{U}_\beta^{(0)} = 0, \tag{B15}$$

which is used frequently below. With the expansions above it is useful to also expand the PV flux

$$\mathbf{u}P = \sum_{j=0}^{\infty}\varepsilon^j\mathbf{F}_0^{(j)}(\mathbf{X},T) + \Re\sum_{\beta=1}^{\infty}\sum_{j=0}^{\infty}\varepsilon^j\mathbf{F}_\beta^{(j)}(\mathbf{X},T)\mathrm{e}^{\mathrm{i}\beta\phi(\mathbf{X},T)/\varepsilon}, \tag{B16}$$

where only the contributions

$$\mathbf{F}_0^{(0)} = \frac{1}{2}\Re\sum_{\beta=1}^{\infty}\mathbf{U}_\beta^{(0)}P_\beta^{(0)*}, \tag{B17}$$

$$\mathbf{F}_0^{(1)} = \mathbf{U}_0^{(0)}P_0^{(0)} + \frac{1}{2}\Re\sum_{\beta=1}^{\infty}\left(\mathbf{U}_\beta^{(0)}P_\beta^{(1)*} + \mathbf{U}_\beta^{(1)}P_\beta^{(0)*}\right), \tag{B18}$$

$$\mathbf{F}_\beta^{(0)} = \mathbf{U}_0^{(0)} P_\beta^{(0)} + \frac{1}{2} \sum_{\beta',\beta''=1}^{\infty} \Big[\mathbf{U}_{\beta'}^{(0)} P_{\beta''}^{(0)} \delta\left(\beta'+\beta''-\beta\right) + \mathbf{U}_{\beta'}^{(0)} P_{\beta''}^{(0)*} \delta\left(\beta'-\beta''-\beta\right) + \mathbf{U}_{\beta'}^{(0)*} P_{\beta''}^{(0)} \delta\left(-\beta'+\beta''-\beta\right) \Big] \quad \text{(B19)}$$

$$\mathbf{F}_\beta^{(1)} = \mathbf{U}_0^{(0)} P_\beta^{(1)} + \mathbf{U}_0^{(1)} P_\beta^{(0)} + \mathbf{U}_\beta^{(0)} P_0^{(0)} + \frac{1}{2} \sum_{\beta',\beta''=1}^{\infty} \Big[\left(\mathbf{U}_{\beta'}^{(0)} P_{\beta''}^{(1)} + \mathbf{U}_{\beta'}^{(1)} P_{\beta''}^{(0)}\right) \delta\left(\beta'+\beta''-\beta\right) + \left(\mathbf{U}_{\beta'}^{(0)} P_{\beta''}^{(1)*} + \mathbf{U}_{\beta'}^{(1)} P_{\beta''}^{(0)*}\right) \delta\left(\beta'-\beta''-\beta\right) + \left(\mathbf{U}_{\beta'}^{(0)*} P_{\beta''}^{(1)} + \mathbf{U}_{\beta'}^{(1)*} P_{\beta''}^{(0)}\right) \delta\left(-\beta'+\beta''-\beta\right) \Big] \quad \text{(B20)}$$

are used explicitly below.

The leading $\mathcal{O}(1)$ of the PV-conservation equation (B1) is found with this only to have wave parts, yielding

$$-\mathrm{i}\beta\omega P_\beta^{(0)} + \mathrm{i}\beta \mathbf{k}_h \cdot \mathbf{F}_\beta^{(0)} = 0. \quad \text{(B21)}$$

Due to the solenoidality (B15), the nonlinear triad part in the PV flux (B19) has a vanishing scalar product with $\mathbf{k}_h$ so that one obtains the polarization relation

$$0 = \beta\widehat{\omega} = \beta\left(\omega - \mathbf{k}_h \cdot \mathbf{U}_0^{(0)}\right). \quad \text{(B22)}$$

The next $\mathcal{O}(\varepsilon)$ has a synoptic-scale part

$$0 = \nabla_{X,h} \cdot \mathbf{F}_0^{(0)}. \quad \text{(B23)}$$

However, using (B17), (B13), and (B9), one finds that $\mathbf{F}_0^{(0)} = 0$, so that this equation is satisfied trivially. The corresponding wave parts are

$$0 = -\mathrm{i}\beta\omega P_\beta^{(1)} + \frac{\partial P_\beta^{(0)}}{\partial T} + \mathrm{i}\beta\mathbf{k}_h \cdot \mathbf{F}_\beta^{(1)} + \nabla_X \cdot \mathbf{F}_\beta^{(0)}. \quad \text{(B24)}$$

Using the explicit flux contributions (B19) and (B20), the solenoidality (B15) and the dispersion relation (B22), this becomes

$$0 = \left(\frac{\partial}{\partial T} + \mathbf{U}_0^{(0)} \cdot \nabla_{X,h}\right) P_\beta^{(0)} + \mathrm{i}\beta\mathbf{k}_h \cdot \mathbf{U}_0^{(1)} P_\beta^{(0)} + \frac{1}{2} \sum_{\beta',\beta''=1}^{\infty} \times \Big\{ \Big[\mathrm{i}\beta\mathbf{k}_h \cdot \mathbf{U}_{\beta'}^{(1)} P_{\beta''}^{(0)} + \nabla_{X,h} \cdot \left(\mathbf{U}_{\beta'}^{(0)} P_{\beta''}^{(0)}\right)\Big] \delta\left(\beta'+\beta''-\beta\right) + \Big[\mathrm{i}\beta\mathbf{k}_h \cdot \mathbf{U}_{\beta'}^{(1)} P_{\beta''}^{(0)*} + \nabla_{X,h} \cdot \left(\mathbf{U}_{\beta'}^{(0)} P_{\beta''}^{(0)*}\right)\Big] \delta\left(\beta'-\beta''-\beta\right) + \Big[\mathrm{i}\beta\mathbf{k}_h \cdot \mathbf{U}_{\beta'}^{(1)*} P_{\beta''}^{(0)} + \nabla_{X,h} \cdot \left(\mathbf{U}_{\beta'}^{(0)*} P_{\beta''}^{(0)}\right)\Big] \delta\left(-\beta'+\beta''-\beta\right) \Big\} \quad \text{(B25)}$$

However, due to (B13), (B14), and $\nabla_X \times \mathbf{k} = 0$, one has

$$\mathrm{i}\beta\mathbf{k}_h \cdot \mathbf{U}_{\beta'}^{(1)} = \mathrm{i}\frac{\beta}{\beta'}\beta'\mathbf{k}_h \cdot \mathbf{U}_{\beta'}^{(1)} = -\frac{\beta}{\beta'}\nabla_X \cdot \mathbf{U}_{\beta'}^{(0)}, \quad \text{(B26)}$$

so that multiplication of (B25) by $\overline{R}^{(0)} P_\beta^{(0)*}/2$ and taking the real part of the product yields the potential-enstrophy equation (136).

Finally, from the $\mathcal{O}(\varepsilon^2)$, we only use the synoptic-scale part

$$0 = \frac{\partial P_0^{(0)}}{\partial T} + \nabla_{X,h} \cdot \mathbf{F}_0^{(1)} \quad \text{(B27)}$$

or, using $\nabla_{X,h} \cdot \mathbf{U}_0^{(0)} = 0$ and (B18),

$$0 = \left(\frac{\partial}{\partial T} + \mathbf{U}_0^{(0)} \cdot \nabla_{X,h}\right) P_0^{(0)} + \frac{1}{2} \nabla_{X,h} \cdot \Re \sum_{\beta=1}^{\infty} \left(\mathbf{U}_\beta^{(0)} P_\beta^{(1)*} + \mathbf{U}_\beta^{(1)} P_\beta^{(0)*}\right). \quad \text{(B28)}$$

Inserting (B13), (B14), (B9), and (B10), again using $\nabla_X \times \mathbf{k} = 0$, and finally resorting to the definitions (128), (129), and (137), one obtains after some algebra the prognostic equation (154) for the leading-order synoptic-scale PV. We point out that the equation system derived here is not closed. The solution of the potential-enstrophy equations requires knowledge of the phase of the PV amplitudes $P_\beta^{(0)}$. As can be seen from (B25), one needs for this $\mathbf{U}_0^{(1)}$. For this one would have to solve the next-order equation for the synoptic-scale PV, involving $P_\beta^{(1)}$ and so forth.

Acknowledgements

We thank O. Bühler and two anonymous reviewers for helpful comments on earlier versions of the manuscript. UA and BR thank the German Federal Ministry of Education and Research (BMBF) for partial support through the program *Role of the Middle Atmosphere in Climate (ROMIC)* and through grant 01LG1220A. UA and RK thank the German Research Foundation (DFG) for partial support through the research unit *Multiscale Dynamics of Gravity Waves (MS-GWaves)* and through grants AC 71/8-1, AC 71/9-1, AC 71/10-1, and KL 611/25-1.

References

Achatz U. 2007. Gravity-wave breaking: Linear and primary nonlinear dynamics. *Adv. Space Res.* **40**: 719–733, doi: 10.1016/j.asr.2007.03.078.

Achatz U, Klein R, Senf F. 2010. Gravity waves, scale asymptotics, and the pseudo-incompressible equations. *J. Fluid Mech.* **663**: 120–147.

Alexander M, Dunkerton T. 1999. A spectral parameterization of mean-flow forcing due to breaking gravity waves. *J. Atmos. Sci.* **56**: 4167–4182.

Alexander MJ, Geller M, McLandress C, Polavarapu S, Preusse P, Sassi F, Sato K, Eckermann S, Ern M, Hertzog A, Kawatani Y, Pulido M, Shaw TA, Sigmond M, Vincent R, Watanabe S. 2010. Recent developments in gravity-wave effects in climate models and the global distribution of gravity-wave momentum flux from observations and models. *Q. J. R. Meteorol. Soc.* **136**: 1103–1124.

Andrews D, McIntyre M. 1976. Planetary waves in horizontal and vertical shear: The generalized Eliassen-Palm relation and the mean zonal acceleration. *J. Atmos. Sci.* **33**: 2031–2048, doi: 10.1175/1520-0469(1976)033<2031:pwihav>2.0.co;2.

Andrews D, McIntyre M. 1978a. An exact theory of nonlinear waves on a Lagrangian-mean flow. *J. Fluid Mech.* **89**: 609–646.

Andrews D, McIntyre M. 1978b. On wave-action and its relatives. *J. Fluid Mech.* **89**: 647–664.

Bierdel L, Snyder C, Park SH, Skamarock WC. 2016. Accuracy of rotational and divergent kinetic energy spectra diagnosed from flight-track winds. *J. Atmos Sci.* **73**: 3273–3286, doi: 10.1175/JAS-D-16-0040.1.

Bölöni G, Ribstein B, Muraschko J, Sgoff C, Wei J, Achatz U. 2016. The interaction between atmospheric gravity waves and large-scale flows: An efficient description beyond the non-acceleration paradigm. *J. Atmos. Sci.* **73**: 4833–4852, doi: 10.1175/JAS-D-16-0069.1.

Borchert S, Achatz U, Fruman MD. 2014. Gravity wave emission in an atmosphere-like connfguration of the differentially heated rotating annulus experiment. *J. Fluid Mech.* **758**: 287–311.

Bretherton F. 1966. The propagation of groups of internal gravity waves in a shear flow. *Q. J. R. Meteorol. Soc.* **92**: 466–480.

Bretherton F. 1969. On the mean motion induced by internal gravity waves. *J. Fluid Mech.* **36**: 785–803.

Bühler O. 2009. *Waves and Mean Flows*. Cambridge University Press: Cambridge, UK.

Bühler O. 2010. Wave–vortex interactions in fluids and superfluids. *Annu. Rev. Fluid Mech.* **42**: 205–228.

Bühler O, McIntyre M. 1998. On non-dissipative wave–mean interactions in the atmosphere or oceans. *J. Fluid Mech.* **354**: 301–343.

Bühler O, McIntyre M. 1999. On shear-generated gravity waves that reach the mesosphere. Part II: Wave propagation. *J. Atmos. Sci.* **56**: 3764–3773.

Bühler O, McIntyre M. 2003. Remote recoil: a new wave–mean interaction effect. *J. Fluid Mech.* **492**: 207–230.

Bühler O, McIntyre M. 2005. Wave capture and wave-vortex duality. *J. Fluid Mech.* **534**: 67–95, doi: 10.1017/S0022112005004374.

Caillol P, Zeitlin V. 2000. Kinetic equations and stationary energy spectra of weakly nonlinear internal gravity waves. *Dyn. Atmos. Oceans* **32**: 81–112, doi: 10.1016/S0377-0265(99)00043-3.

Callies J, Ferrari R, Bühler O. 2014. Transition from geostrophic turbulence to inertia-gravity waves in the atmospheric energy spectrum. *Proc. Natl. Acad. Sci. U.S.A.* **111**: 17033–17038.

Charney J. 1947. The dynamics of long waves in a baroclinic westerly current. *J. Meteorol.* **4**: 135–163.

Charney JG. 1948. On the scale of atmospheric motion. *Geofys. Publ. Oslo* **17**: 1–17.

Dosser HV, Sutherland BR. 2011. Anelastic internal wave packet evolution and stability. *J. Atmos. Sci.* **68**: 2844–2859, doi: 10.1175/JAS-D-11-097.1.

Durran D. 1989. Improving the anelastic approximation. *J. Atmos. Sci.* **46**: 1453–1461.

Eady ET. 1949. Long waves and cyclone waves. *Tellus* **1**: 33–52.

Eliassen A, Palm E. 1961. On the tansfer of energy in stationary mountain waves. *Geofys. Publ. Oslo* **22**: 1–23.

Fritts D, Alexander M. 2003. Gravity wave dynamics and effects in the middle atmosphere. *Rev. Geophys.* **41**: 1003, doi: 10.1029/2001RG000106.

Fritts DC, Dunkerton TJ. 1984. A quasi-linear study of gravity-wave saturation and self-acceleration. *J. Atmos. Sci.* **41**: 3272–3289.

Gage KS. 1979. Evidence for a law inertial range in mesoscale two-dimensional turbulence. *J. Atmos Sci.* **36**: 1950–1954.

Grimshaw R. 1975a. Internal gravity waves: Critical layer absorption in a rotating fluid. *J. Fluid Mech.* **70**: 287–304.

Grimshaw R. 1975b. Nonlinear internal gravity waves in a rotating fluid. *J. Fluid Mech.* **71**: 497–512, doi: 10.1017/S0022112075002704.

Hertzog A, Souprayen C, Hauchecorne A. 2002. Eikonal simulations for the formation and the maintenance of atmospheric gravity wave spectra. *J. Geophys. Res.* **107**: ACL 4-1–ACL 4-14, doi: 10.1029/2001JD000815.

Hines C. 1997. Doppler spread parameterization of gravity-wave momentum deposition in the middle atmosphere. Part 1. Basic formulation. *J. Atmos. Sol. Terr. Phys.* **59**: 371–386.

Holton JR. 1982. The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere. *J. Atmos. Sci.* **39**: 791–799.

Kim YJ, Eckermann SD, Chun HY. 2003. An overview of the past, present and future of gravity-wav drag parametrization for numerical climate and weather prediction models. *Amos. Ocean* **41**: 65–98.

Klein R. 2011. 'On the regime of validity of sound-proof model equations for atmospheric flows'. In Workshop on Non-hydrostatic Modelling, 8–10 November 2010. ECMWF: Reading, UK.

Klein R, Achatz U, Bresch D, Knio OM, Smolarkiewicz PK. 2010. Regime of validity of sound-proof atmospheric flow models. *J. Atmos. Sci.* **67**: 3226–3237.

Lilly DK. 1983. Stratified turbulence and the mesoscale variability of the atmosphere. *J. Atmos. Sci.* **40**: 749–761.

Lindborg E. 2015. A Helmholtz decomposition of structure functions and spectra calculated from aircraft data. *J. Fluid Mech.* **762**: doi: 10.1017/jfm.2014.685.

Lindzen RS. 1981. Turbulence and stress owing to gravity wave and tidal breakdown. *J. Geophys. Res.* **86**: 9707–9714, doi: 10.1029/JC086iC10p09707.

Medvedev A, Klaassen G. 1995. Vertical evolution of gravity wave spectra and the parameterization of associated gravity wave drag. *J. Geophys. Res.* **100**: 25841–25853, doi: 10.1029/95JD02533.

Müller P. 1976. On the diffusion of momentum and mass by internal gravity waves. *J. Fluid Mech.* **77**: 789–823.

Muraschko J, Fruman M, Achatz U, Hickel S, Toledo Y. 2015. On the application of WKB theory for the simulation of the weakly nonlinear dynamics of gravity waves. *Q. J. R. Meteorol. Soc.* **141**: 676–697.

Nazarenko S. 2011. *Wave Turbulence*. Springer: Berlin.

Pedlosky J. 1987. *Geophysical Fluid Dynamics*. Springer: Berlin.

Plougonven R, Zhang F. 2007. On the forcing of inertia-gravity waves by synoptic-scale flows. *J. Atmos. Sci.* **64**: 1737–1742.

Plougonven R, Zhang F. 2014. Internal gravity waves from atmospheric jets and fronts. *Rev. Geophys.* **52**: 33–76, doi: 10.1002/2012RG000419.

Plumb R. 1986. Three-dimensional propagation of transient quasi-geostrophic eddies and its relationship with the eddy forcing of the time-mean flow. *J. Atmos. Sci.* **43**: 1657–1678.

Ribstein B, Achatz U, Senf F. 2015. The interaction between gravity waves and solar tides: results from 4-d ray tracing coupled to a linear tidal model. *J. Geophys. Res.* **120**: 6795–6817, doi: 10.1002/2015JA021349.

Rieper F, Achatz U, Klein R. 2013. Range of validity of an extended WKB theory for atmospheric gravity waves: One-dimensional and two-dimensional case. *J. Fluid Mech.* **729**: 330–363.

Rossby CG. 1938. On the mutual adjustment of pressure and velocity distribution in certain simple current systems, II. *J. Mar. Res.* **1**: 239–263.

Senf F, Achatz U. 2011. On the impact of middle-atmosphere thermal tides on the propagation and dissipation of gravity waves. *J. Geophys. Res.* **116**: D24110, doi: 10.1029/2011JD015794.

Sutherland BR. 2001. Finite-amplitude internal wavepacket dispersion and breaking. *J. Fluid Mech.* **429**: 343–380.

Sutherland BR. 2006. Weakly nonlinear internal gravity wavepackets. *J. Fluid Mech.* **569**: 249–258, doi: 10.1017/S0022112006003016.

Sutherland BR. 2010. *Internal Gravity Waves*. Cambridge University Press: Cambridge, UK.

Tulloch R, Smith KS. 2006. A theory for the atmospheric energy spectrum: depth-limited temperature anomalies at the tropopause. *Proc. Natl. Acad. Sci. U.S.A.* **103**: 14690–14694.

Vallis GK. 2006. *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulation*. Cambridge University Press: Cambridge, UK.

Vallis GK, Shutts GJ, Gray MEB. 1997. Balanced mesoscale motion and stratified turbulence forced by convection. *Q. J. R. Meteorol. Soc.* **123**: 1621–1652.

Wagner GL, Young WR. 2015. Available potential vorticity and wave-averaged quasi-geostrophic flow. *J. Fluid Mech.* **785**: 401–424.

Wagner GL, Young WR. 2016. A three-component model for the coupled evolution of near-inertial waves, quasi-geostrophic flow and the near-inertial second harmonic. *J. Fluid Mech.* **802**: 806–837.

Xie JH, Vanneste J. 2015. A generalised-lagrangian-mean model of the interactions between near-inertial waves and mean flow. *J. Fluid Mech.* **774**: 143–169.

Yasuda Y, Sato K, Sugimoto N. 2015a. A theoretical study on the spontaneous radiation of inertia-gravity waves using the renormalization group method. Part I: Derivation of the renormalization group equations. *J. Atmos. Sci.* **72**: 957–983, doi: 10.1175/JAS-D-13-0370.1.

Yasuda Y, Sato K, Sugimoto N. 2015b. A theoretical study on the spontaneous radiation of inertia-gravity waves using the renormalization group method. Part II: Verification of the theoretical equations by numerical simulation. *J. Atmos Sci.* **72**: 984–1009, doi: 10.1175/JAS-D-13-0371.1.

Zeitlin V, Reznik GM, Jelloul MB. 2003. Nonlinear theory of geostrophic adjustment. Part 2. Two-layer and continuously stratified primitive equations. *J. Fluid Mech.* **491**: 207–228.

Zhang F, Wei J, Zhang M, Bowman KP, Pan LL, Atlas E, Wofsy SC. 2015. Aircraft measurements of gravity waves in the upper troposphere and lower stratosphere during the START08 field experiment. *Atmos. Chem. Phys.* **15**: 7667–7684, doi: 10.5194/acp-15-7667-2015.