Generalized modulation theory for strongly nonlinear gravity waves in a compressible atmosphere

Mark Schlutow* and Erik Wahlén

Abstract: This study investigates strongly nonlinear gravity waves in the compressible atmosphere from the Earth's surface to the deep atmosphere. These waves are effectively described by Grimshaw’s dissipative modulation equations which provide the basis for finding stationary solutions such as mountain lee waves and testing their stability in an analytic fashion. Assuming energetically consistent boundary and far-field conditions, that is no energy flux through the surface, free-slip boundary, and finite total energy, general wave solutions are derived and illustrated in terms of realistic background fields. These assumptions also imply that the wave-Reynolds number must become less than unity above a certain height. The modulational stability of admissible, both non-hydrostatic and hydrostatic, waves is examined. It turns out that, when accounting for the self-induced mean flow, the wave-Froude number has a resonance condition. If it becomes $1/\sqrt{2}$, then the wave destabilizes due to perturbations from the essential spectrum of the linearized modulation equations. However, if the horizontal wavelength is large enough, waves overturn before they can reach the modulational stability condition.

1 Introduction

Gravity waves are an omnipresent oscillation mode in the atmosphere. They redistribute energy vertically but also laterally and thereby affect the dynamics relevant for weather and climate prediction [1, 2]. Usually excited in the troposphere, gravity waves may persist deep into the upper layers of the atmosphere [3, 4, 5] where they interact with the mean flow. They exert drag onto the horizontal mean-flow momentum, produce heat when dissipating [6], and cause increased mixing of tracer constituents such as green-house gases [7].

Atmospheric gravity waves can be excited by various processes such as convection due to latent heat release in clouds or spontaneous emission from jets and fronts [8]. The dominant excitation mechanism is background flow over mountain ranges resulting in mountain lee waves that can be considered quasi-stationary as their horizontal phase speed is essentially opposite to the mean-flow horizontal wind. In the troposphere wave amplitudes are often small such that linear wave theory is applicable [9, 10]. When mountain lee waves extend into the higher layers they get anelastically amplified due to the decreasing background density which is an effect of the compressibility of the atmosphere. Amplitudes get indeed so large that linear theory becomes invalid.

The natural extension to linear wave theory is to incorporate weakly nonlinear effects which remains asymptotically valid as long as the amplitudes are relatively small. This approach works exceedingly well for the description of oceanic internal waves. We will argue, however, that weakly nonlinear theory is in certain cases not sufficient for atmospheric gravity waves. These cases comprise situations when the amplitudes come close to the regime where the horizontal wind perturbation due to the waves is of the same order of

*Corresponding author: Mark Schlutow, Institut für Mathematik, Freie Universität Berlin, Germany, E-mail: mark.schlutow@fu-berlin.de
Erik Wahlén, Centre for Mathematical Sciences, Lund University, Sweden, E-mail: erik.wahlen@math.lu.se
magnitude as the mean-flow horizontal wind. This situation is, in fact, rather common in the stratosphere and higher [11].

Weakly nonlinear theory for gravity waves was studied in [12, 13, 14, 15, 16]. Here, nonlinear effects such as Doppler shift of the frequency and interaction with the mean flow appear as higher-order corrections to the linear model. In the asymptotic limit the model therefore approaches linear theory. An important effect of the weak nonlinearity is the occurrence of modulational instabilities: plane non-hydrostatic Boussinesq waves become modulationally unstable if the second derivative of the dispersion relation with respect to the vertical wavenumber becomes negative. It was shown in [17] that this holds true even for strongly nonlinear waves of the same kind. In the strongly nonlinear theory, Doppler shift and wave-mean-flow interaction appear to leading order such that the perturbation fields are of the same order of magnitude as the background. However, Boussinesq theory does not account for the anelastic amplification. Furthermore, the plane waves extend to the infinities experiencing no lower boundary conditions and no dissipative effects. In [18], a particular flow regime, where anelastic amplification and dissipative forces are exactly balanced, was investigated with respect to modulational stability.

This study aims to generalize the modulation theory incorporating strong nonlinearity, anelastic amplification according to pseudo-incompressible theory, dissipative damping and lower boundary conditions in a comprehensive fashion. Pioneering work on the modulation theory of strongly nonlinear gravity waves was accomplished by [19, 20].

In section 2 of this paper we will introduce Grimshaw’s dissipative modulation equations as our governing equations and link them to asymptotic solutions of the compressible Navier-Stokes equations. Boundary conditions and limit behavior in the far field as derived from physical arguments will be shown in section 3. After the introduction of the antitriptic flow assumption in section 4, which amounts to a leading-order balance of mean-flow pressure gradient and frictional forces, stationary solutions will be found and illustrated in terms of observational data in section 5. In section 6, the modulational stability of the stationary solution will be investigated followed by some concluding remarks in section 7.

2 Model equations

The governing equations for our investigation are Grimshaw’s dissipative modulation equations [20]. Solving them generates an asymptotic solution to the Navier-Stokes equations (NSE). Before stating Grimshaw’s equations let us briefly elaborate on their derivation by means of nonlinear Wentzel-Kramers-Brillouin (WKB) theory. A systematic derivation of the inviscid modulation equations consistent with the Euler equations can be found in [21, 22]. For a review on how viscosity and heat conduction are added to the picture we refer to [18].

The starting point for the derivation are the dimensionless NSE scaled by an appropriate reference length \( L_r \) (a typical wavelength) and time \( t_r \) (a typical oscillation period) favorable for gravity waves. So, the dimensionless coordinates are written as

\[
(x, z, t) = (x^* / L_r, z^* / L_r, t^* / t_r).
\]  

(1)

Throughout this work, dimensional variables are labeled by an asterisk. The appropriate scaling of the NSE is specified in terms of a scale separation parameter \( \varepsilon = L_r / H_\theta \ll 1 \) where \( H_\theta \) denotes the reference scale height of potential temperature, a measure for the reference stratification. The idea is to map \( \varepsilon \) to the Mach, Froude, Reynolds and Prandtl number, individually, which generates a distinguished limit typical for gravity waves. For the sake of brevity, we do not reproduce the full set of the scaled NSE here. The interested reader finds them in [18, Eq. 7].

Next, compressed coordinates are introduced,

\[
(X, Z, T) = (\varepsilon^a x, \varepsilon z, \varepsilon^a t) \quad \text{where} \quad a = \begin{cases} 1, & (nh), \\ 2, & (h), \end{cases}
\]  

(2)
The two different versions of compressed coordinates correspond to the hydrostatic (h) and non-hydrostatic (nh) wave regime, respectively. Assuming that the background changes only slowly in comparison with the rapidly oscillating wave (the “WKB assumption”), the thermodynamic background state can be entirely determined by the background temperature profile \( T(z) \). The background density \( \rho \) and the Brunt-Väisälä frequency \( N \) are calculated with regard to a hydrostatic atmosphere being considered an ideal gas,

\[
N^2(z) = \frac{1}{\Theta} \left( \frac{d\Theta}{dZ} + 1 \right), \tag{3a}
\]

\[
\rho(Z) = \exp \left( \int_0^Z \eta(Z') \, dZ' \right) \quad \text{with} \quad \eta(Z) = \frac{1}{\Theta} \left( \frac{d\Theta}{dZ} + \frac{1}{\kappa} \right) \tag{3b}
\]

where \( \kappa = (\gamma - 1)/\gamma \) and \( \gamma \) the heat capacity ratio for ideal gases.

With this preparation, the NSE are ready for the asymptotic ansatz, i.e. the nonlinear WKB expansion. The idea is to expand the state vector of the NSE \( \mathbf{U} = (\mathbf{v}, \beta, \phi)^T \) in a Fourier-like fashion where amplitudes and phase are functions only of the compressed coordinates. Here, \( \mathbf{v}, N\beta = b \) and \( \phi \) denote the two-dimensional velocity vector, buoyancy force and kinematic pressure, respectively. The latter is defined by the product of background potential temperature and Exner pressure of the perturbation. The WKB ansatz reads

\[
\mathbf{U}(x, z; \epsilon) = \mathbf{U}_{0,0}(X, Z, T) + \left( \mathbf{U}_{0,1}(X, Z, T)e^{i\Phi(X,Z,T)/\epsilon} + \text{c.c.} \right) + \text{h.h.} + O(\epsilon) \quad \tag{4}
\]

where c.c. stands for the complex conjugate and h.h. for higher harmonics. This ansatz is substituted into the NSE and terms are ordered in powers of \( \epsilon \) and harmonics, i.e. integer multiples of the phase \( \Phi \). The leading-order solution of the NSE as \( \epsilon \to 0 \) is given in terms of Grimshaw’s dissipative modulation equations [20, Eq. 4.1–4.6]

\[
\begin{align*}
\frac{\partial \mathbf{k}}{\partial T} + \nabla_X (\hat{\omega} + k_X \mathbf{u}) &= 0 \quad \tag{5a} \\
\rho \frac{\partial \mathbf{a}}{\partial T} + \nabla_X \cdot \left[ \left( \frac{\partial \hat{\omega}}{\partial \mathbf{k}} + \mathbf{u} e_x \right) \rho a \right] &= -\Lambda |\mathbf{k}|^2 \rho a \quad \tag{5b} \\
\rho \frac{\partial \mathbf{u}}{\partial T} + \nabla_X \cdot \left( \frac{\partial \hat{\omega}}{\partial \mathbf{k}} k_X \rho a \right) &= -\frac{\partial \rho}{\partial X} \quad \tag{5c} \\
\frac{\partial \mathbf{u}}{\partial X} &= 0. \tag{5d}
\end{align*}
\]

This set of coupled nonlinear partial differential equations governs the evolution of the modulation fields, wavenumber vector \( \mathbf{k} = (k_x, k_z)^T \), wave action density \( \rho a \), and mean-flow horizontal momentum \( \rho \mathbf{u} \). \( \nabla_X = (\partial/\partial X_x, \partial/\partial Z)^T \) denotes the nabla operator and \( e_x \) is the unit vector pointing into the horizontal direction. Intrinsic frequency \( \hat{\omega} \) is a function of \( \mathbf{k} \). It is determined by the dispersion relation for either hydrostatic (h) or non-hydrostatic (nh) gravity waves

\[
\hat{\omega}(k_z) = \begin{cases} 
N k_x / |\mathbf{k}|, & (nh), \\
N k_z / |k_z|, & (h), 
\end{cases}
\quad \tag{6}
\]

with \( |\mathbf{k}|^2 = k_x^2 + k_z^2 \).

The height profile of the kinematic viscosity \( \Lambda \) is determined by the sum of turbulent and molecular viscosity. \( \Lambda \) is negligible at the surface and in the lower atmosphere. At a certain height, however, amplitudes of gravity waves become typically so large due to the anelastic amplification that turbulence is produced by small-scale instabilities and this increases \( \Lambda \). Even higher in the atmosphere, about 100 km, molecular viscosity starts to dominate becoming orders of magnitude larger than the turbulent viscosity. Eventually in the highest layer, the thermosphere, molecular dissipation damps effectively every wave motion.
The modulation fields generate the leading-order variables from the WKB ansatz and hence the asymptotic solution to the NSE via

\[ \nabla_X \Phi = k, \quad (7a) \]

\[ -\frac{\partial \Phi}{\partial T} = \hat{\omega} + k_x u, \quad (7b) \]

\[ U_{0,0} = (u, 0, 0, p)^T, \quad (7c) \]

\[ U_{0,1} = \mathcal{B} U', \quad (7d) \]

where

\[ U' = \left( -\frac{k_x}{k_z N}, i \frac{\hat{\omega}}{N}, 1, -i \frac{k_x}{k_z^2 N} \right)^T \]

represents the polarization vector and

\[ \mathcal{B} = \sqrt{\hat{\omega} a / 2} \]

the amplitude.

In Whitham’s modulation theory [12], equation (5a) represents conservation of waves. The second equation (5b) gives conservation of wave action plus a sink due to dissipation. Finally, (5c) describes the acceleration of mean-flow horizontal momentum due to horizontal pseudo-momentum flux convergence and the horizontal gradient of mean-flow kinematic pressure \( p \). The latter is unspecified at the moment and needs further assumptions to close the system. A thorough discussion of the horizontal gradient of mean-flow kinematic pressure is also found in [20, p. 142].

Moreover, the modulation equations exhibit a total energy density being the sum of mean-flow kinetic energy density and wave energy density,

\[ \rho e = \frac{1}{2} \rho u^2 + \rho a \hat{\omega}. \quad (9) \]

It evolves in time as governed by

\[ \rho \frac{\partial e}{\partial T} + \nabla_X \cdot \left[ p u e_x + \left( \frac{\partial \hat{\omega}}{\partial k} + u e_x \right) \rho a (\hat{\omega} + k_x u) \right] = -\hat{\omega} \Lambda |k|^2 \rho a. \quad (10) \]

In conclusion, (9) is a locally conserved quantity in the inviscid limit, i.e. \( \Lambda \to 0 \).

In [22], exact horizontally confined traveling wave packet solutions for the inviscid version of (5) were derived. Note that the class of traveling waves also includes stationary waves (with zero propagation speed).

Grimshaw considered two classes of solutions in [20] for the dissipative case: \((Z, T)\)-dependent modulations, supported and not supported by the horizontal gradient of mean-flow kinematic pressure. In the following analysis, we will also assume modulation only in the vertical direction and we will argue that this implies a non-vanishing horizontal gradient of mean-flow kinematic pressure for stationary waves like mountain lee waves.

If \( k, a \) and \( u \) depend only on \( T \) and \( Z \), (5a) implies that \( k_x = \text{const.} \) And, without loss of generality, we restrict the subsequent considerations with \( k_x > 0 \). Furthermore, (5d) is automatically fulfilled. The remaining equations simplify to

\[ \frac{\partial k_x}{\partial T} + \frac{\partial}{\partial Z} (\hat{\omega} + k_x u) = 0 \quad (11a) \]

\[ \rho \frac{\partial a}{\partial T} + \frac{\partial}{\partial Z} (\hat{\omega} a) = -\Lambda |k|^2 \rho a \quad (11b) \]

\[ \rho \frac{\partial u}{\partial T} + \frac{\partial}{\partial Z} (\hat{\omega} k_x a) = -\frac{\partial p}{\partial X}. \quad (11c) \]

where primes denote derivatives with respect to the vertical wavenumber \( k_z \). The domain is \( Z \in [0, \infty) \), from the Earth’s surface to the deep atmosphere.

3 **Surface boundary and far-field conditions**

The leading-order asymptotic solution to the NSE as defined by (4) governed by the modulation equations already determines the shape of the anticipated wave solution. However, it has some degrees of freedom to
specify physically motivated surface boundary conditions and the far-field limit behavior. Given a mean-flow horizontal wind, three additional conditions are needed to set the horizontal and vertical wavenumber as well as the specific wave action density at $Z = 0$. We will assume that close to the surface the viscosity is negligible. Thus, a free-slip boundary is justified. Let us carve a mountain to the shape of the lowest stream line. Then, this particular, prototypical mountain will define two of the surface boundary conditions. For the third, we will assume that there shall be no energy flux through the boundary and therefore the extrinsic frequency will be zero at the mountain.

### 3.1 No-energy-flux boundary condition

According to (10), if no vertical energy flux through the boundary but finite wave action and wavenumber at the boundary are presumed, then

$$\hat{\omega} + k_x u = 0 \quad \text{at} \quad Z = 0$$

must hold which coincides with the absence of wavenumber flux (cf. (11a)). Let us remark that therefore $u < 0$ as $k_x > 0$ and $\hat{\omega} > 0$ at $Z = 0$.

### 3.2 Free-slip boundary condition: carving a mountain to the wave

To leading order the solution as given by (4) is solenoidal and therefore entirely determined by a stream function. From the polarization vector (8), the stream function can be written in terms of the fast, uncompressed coordinates as

$$\Psi(x, z) = u z - 2B \frac{\hat{\omega}}{Nk_x} \cos(k_x x + k_z z) + O(\varepsilon).$$

(13)

Stream functions of solenoidal flows are constant on stream lines $(x, h(x))$ which can be formulated mathematically by

$$\Psi(x, h(x)) = \text{const.}$$

(14)

Considering the stream line in an $O(\varepsilon)$-neighborhood of the boundary $Z = 0$, (14) defines a parametrization of $h$ implicitly via

$$h(x) = \frac{2B}{u} \frac{\hat{\omega}}{Nk_x} \cos(k_x x + k_z h(x)) + O(\varepsilon).$$

(15)

Here, we have essentially presented the boundary condition of [23, Eq. 7], who define the condition in terms of the vertical displacement of an air parcel $\delta = -\Psi/u$, but in the framework of Grimshaw’s dissipative modulation equations.

Note that in the asymptotic limit $\varepsilon \to 0$, (15) is of the form

$$y = f(y)$$

(16)

with $y(\xi) = k_z h$, $\xi = k_x x$ and $f = q \cos(\xi + y)$ where

$$q = \frac{2B}{u} \frac{\hat{\omega} k_z}{Nk_x}.$$  

(17)

Thus, the stream line can be interpreted as a fixed point of $f$. For every fixed $\xi$ the differentiable function $f$ of $y$ has a Lipschitz constant

$$L = \left| \sup_y f'(y) \right| = |q|.$$  

(18)
Banach’s fixed point theorem states if \( L = |q| < 1 \), then \( f \) is a contraction and hence has a fixed point.

The results of a fixed-point iteration \( y_{n+1} = f(y_n) \) with \( y_0 = 0 \) are plotted in figure 1. It can be observed how the streamlines steepen when \( q \) is increased whereas for small values the streamlines approximate the sinusoidal (linear) profile. Exploiting the polarization (8), we can write
\[
q = \frac{u'}{|u'|}, \quad u' = 2|u_{0,1}| = 2B\hat{\omega}\frac{|k_x|}{Nk_x} \tag{19}
\]
which provides a measure for nonlinearity. If it is small one may assume linear wave excitation. The factor 2 appears simply because of the definition of amplitude due to (4).

Also, using polarization (8) and no-energy-flux (12) in combination with the convergence condition for the fixed-point iteration, we obtain
\[
N^2 > |k_x|b', \quad b' = 2B N \tag{20}
\]
which confirms the classical condition for static stability, lines of constant potential temperature must not overturn [24].

A boundary condition close to the surface but sufficiently far away to be considered free-slip assuming \( \Lambda(0) = 0 \) can be determined by a periodic mountain ridge with period \( \mathcal{P} \) and maximum mountain height \( H_m \) which we consider to be given constants hereinafter. In terms of the nonlinearity parameter (17) and the no-energy-flux condition (12), we obtain that \( k_x = 2\pi/\mathcal{P} \) and
\[
H_m = \frac{2B}{N} \quad \text{at} \quad Z = 0. \tag{21}
\]

The combined boundary condition from (21) and (12) may be written in vector form as
\[
B(y) = \begin{pmatrix}
\hat{\omega} + k_x u \\
H_m^2 N^2 - 2a\hat{\omega}
\end{pmatrix} = 0 \quad \text{at} \quad Z = 0. \tag{22}
\]

### 3.3 Far-field condition - limit behavior

Additionally to the lower boundary, the limit of the solution as \( Z \to \infty \) must be specified in order to obtain a physical wave solution. For this argument, we exploit the global energy as derived from (9) being
\[
\int_0^\infty \rho e \, dZ = \int_0^\infty \frac{1}{2} \rho u^2 \, dZ + \int_0^\infty \rho a\hat{\omega} \, dZ. \tag{23}
\]
A physical wave must be such that the global energy is finite.

In the thermosphere, temperature approaches an equilibrium $T_\infty$ as altitude increases [25]. Hence, it is safe to assume that $N \to N_\infty$ and $\eta \to \eta_\infty < 0$ as $Z \to \infty$ (cf. (3a) and (3b)). Due to the high viscosity in the thermosphere and above, we can also assume that $u, a \to 0$ as $Z \to \infty$. These assumptions suffice for the integrals in (23) to converge and also that the energy flux in (10) vanishes. Additionally, we assume that the background variables are exponentially asymptotic, i.e. they converge even when multiplied by an exponentially increasing function [26, p 40].

4 The antitriptic flow assumption and momentum conservation

In the inviscid limit the modulation equations (11) assume stationary solutions where $\partial p/\partial X = 0$ which can be computed analytically by a formula of [22, Eq. 5.20]. We want to point out that mentioned formula is also valid for horizontally confined waves. When we multiply (11b) by $k_x$ and subtract (11c), we obtain the evolution equation for total momentum density,

$$\rho \frac{\partial}{\partial T} (k_x a - u) = \frac{\partial p}{\partial X} - k_x A |k|^2 \rho a.$$  

(24)

Thus, to be consistent with the inviscid limit, the dissipative modulation equations assume stationary solutions if the right hand side of (24) vanishes which provides eventually a closure for the horizontal gradient of mean-flow kinematic pressure,

$$\frac{\partial p}{\partial X} = A |k|^2 \rho k_x a.$$  

(25)

This result implicates that the mean-flow horizontal kinematic pressure gradient balances the dissipation of horizontal pseudo-momentum, $\rho k_x a$. Thereby, total momentum is locally conserved.

A flow configuration where pressure gradient balances viscous forces is referred to as antitriptic flow in the literature [27]. Under this antitriptic flow assumption (25) the modulation equations degenerate as we can integrate (24) with respect to time to obtain the mean-flow horizontal wind,

$$u(Z, T) = k_x a(Z, T) + \bar{U}(Z),$$  

(26)

reducing to a diagnostic variable. We call $\bar{U}$ the background horizontal wind since it is time-independent and we emphasize that in the absence of a wave ($a = 0$), mean-flow and background horizontal wind coincide. The difference between mean-flow and background horizontal wind can be identified as specific horizontal pseudo-momentum. Note that in weakly nonlinear wave theory, the pseudo-momentum is of higher order and would not appear in a leading-order equation as shown here.

Substituting (26) into the governing equations (11) reduces the dimension of the system. We may reformulate the reduced system in vector form

$$\frac{\partial \mathbf{y}}{\partial T} + \frac{\partial \mathbf{F}(\mathbf{y})}{\partial Z} = \mathbf{G}$$  

(27)

where we call $\mathbf{y} = (k_x, a)^T : [0, \infty)^2 \to \mathbb{R}^2$ the prognostic state vector. The nonlinear flux and inhomogeneity are determined by

$$\mathbf{F} = (\dot{\omega} + k_x u, \dot{\omega} \rho a)^T,$$

$$\dot{\mathbf{G}} = (0, -(A |k| + \eta \dot{\omega})a)^T.$$  

(28a)

Before we focus on a specific solution and its stability in the next section, we want to discuss how a general solution to the modulation equations with boundary, far-field and antitriptic-flow conditions may look like in the original variables of the Navier-Stokes equations. A prototypical wave is plotted in Figure 2. Additionally, three different length scales are indicated. The rapid oscillations of the wave occur on the
shortest scale $\sim L_r$ and its modulation fields vary on the intermediate scale $\sim H_\theta$ which separates by a factor $\epsilon$. In the previous sections, we assumed that the wave is horizontally homogeneous posing an unphysical situation as neither waves nor mountains in the real world extend to infinity. Therefore, it becomes clear that there is an even larger scale $\sim H_\theta/\epsilon$ on which the wave and mountain vanish at the infinities. For a systematic approach including also the largest scale, one may employ the method of matched asymptotic expansion which is beyond the scope of this paper. However, considering this largest scale rather qualitatively, we are able to comprehend several properties of the solution. First, one may regard the background horizontal wind $\bar{U}$ as the unperturbed wind upstream away from the mountain. In the region where the wave is excited, the wave drag reduces the wind resulting in the wave-induced mean flow (middle diagram in Fig. 2). Second, the right hand side of (25) does not depend on $X$ being on the intermediate scale and can therefore be integrated to obtain a linear function in the horizontal direction for the mean-flow kinematic pressure. Consequently, it is unbounded on the intermediate scale for $|X| \to \infty$ but on the largest scale it will converge towards the balanced state with constant pressure and constant background horizontal wind (bottom diagram).

5 Stationary solution

This section will examine stationary solutions to Grimshaw’s dissipative modulation equations. They describe typical mountain lee waves which are excited by a background flow over a mountain.

5.1 Derivation for stationary waves

A stationary solution $y = Y(Z) = (K_z, A)^T$ must fulfill

$$\frac{\partial F(Y)}{\partial Z} = G(Y).$$
Note that we label the stationary solution with capital letters. The first component of (29) can readily be integrated, so

$$\text{const.} = \Omega = \hat{\Omega} + K_x U.$$  (30)

with $U(A) = K_x A + \hat{U}$ and $\hat{\Omega}^{(n)} = \hat{\omega}^{(n)}(K_z)$. We can solve (30) explicitly exploiting (6) for

$$K_z(A) = \begin{cases} -K_x \sqrt{N^2/(K_x U(A) - \Omega)^2 - 1}, & (nh), \\ K_x N/(K_x U(A) - \Omega), & (h). \end{cases}$$  (31)

We want to point out that in the non-hydrostatic case, the vertical wavenumber may become imaginary which is referred to as evanescence since the ansatz (4) would switch from an oscillatory to an exponential behavior. In fact vertically evanescent waves are observed in the atmosphere and can be modeled by linear theory. However, the nonlinear WKB theory becomes invalid for imaginary wavenumbers. The second component of (29) becomes an explicit, non-autonomous, scalar, ordinary differential equation when we insert (31),

$$\frac{\partial A}{\partial Z} = \Gamma(A, Z)A \text{ in } (0, \infty),$$  (32a)

$$\Gamma = \eta (Re_{\text{wave}}^{-1} - 1) - (1 - a\hat{\Omega})\partial_Z \ln(N) + aK_x \partial_Z \bar{U},$$  (32b)

$$a = \frac{\hat{\Omega}''}{\hat{\Omega}^2},$$  (32c)

$$Re_{\text{wave}} = \frac{\eta \hat{\Omega}'}{\Lambda |K|^2}.$$  (32d)

Definition (32d) can be interpreted as a wave-Reynolds number. It was introduced and discussed by [18] and measures, roughly speaking, the damping of wave amplitude.

From the boundary condition (22), $B(Y) = 0$ at $Z = 0$, we obtain

$$\Omega = 0,$$  (33)

$$A^\pm = \frac{|\hat{U}|}{2K_x} \left( 1 \mp \sqrt{1 - 2Fr_{\text{wave}}^2} \right) \text{ at } Z = 0$$  (34)

where

$$Fr_{\text{wave}} = \frac{Hm N}{|\hat{U}|} \bigg|_{Z=0} \leq \frac{1}{\sqrt{2}}.$$  (35)

We can rule out $A^-$ immediately as non-physical solution as there must be no wave if $H_m = 0$. The constant $Fr_{\text{wave}}$ possesses an upper bound in order to obtain real-valued wave action. It is often referred to as “non-dimensional mountain height” in the literature. Other names are also common. However, we follow a recent discussion in [28] who argue to call it wave-Froude number.

For a brief discussion on the existence of solutions to (32) we restrict ourselves to the hydrostatic and isothermal case, so let $N, \eta$ and $\hat{U}$ be constant. Then, (32a) reduces to

$$\frac{\partial A}{\partial Z} = \eta Re_{\text{wave}}^{-1} - 1 \frac{1}{1 - K_z N^{-1} K_x A} A \text{ in } (0, \infty).$$  (36)

We can immediately observe that the denominator is always greater than one as $K_z < 0$ and due to this fact, a singularity cannot be reached. By definition $Re_{\text{wave}}^{-1} = 0$ at $Z = 0$ and $\eta < 0$ which implies that the solution will grow exponentially from the boundary. This property is known from linear theory and is called anelastic amplification. The viscosity $\Lambda$ increases with height and in conclusion, the wave-Reynolds number drops until it becomes unity. At this point, $A$ assumes its maximum. Physically speaking, it saturates due to turbulent damping.

In order to meet the requirements in the deep-atmosphere limit from section 3.3, $A \rightarrow 0$, we must find

$$Re_{\text{wave}} < 1 \text{ for } Z > Z_{\text{turbo}}.$$  (37)
$Z_{\text{turbo}}$ marks the turbopause, i.e. the transition zone from turbulent-dominated damping to energy diffusion by molecular viscosity [29]. Due to the strong molecular dissipation above the turbopause increasing $\Lambda$, there can be no doubt that the bound (37) will be reached. But its position $Z_{\text{turbo}}$ is hard to predict as it depends on small-scale processes which are not governed by the modulation equations anymore. However, we will see in section 6 that the actual value of $Z_{\text{turbo}}$ has no influence on the stability of the waves.

The general case, non-isothermal and non-hydrostatic, will be investigated numerically in the following section.

5.2 Illustrative example

An illustrative example for a typical mountain wave is plotted in figure 3. We took observation data from the zonal mean COSPAR International Reference Atmosphere (CIRA-86, [30]) for the dimensional temperature profile $T^*$ (Panel a) and the dimensional background horizontal wind $\bar{U}^*$ (Panel d). The values are taken for March in the northern hemisphere at $50^\circ$ as here conditions of background wind for waves that extend deep into the atmosphere are optimal.

We defined the dimensional kinematic viscosity profile $\Lambda^*(z^*) \propto \tanh$ in extrapolation of [31, Fig. 1] since viscosity is unavailable in the CIRA-86 data set. In contrast to the common definition, we span the zonal $x$-axis from East to West due to our sign convention ($k_x > 0$). Background and wave fields are computed accordingly (see figure caption for equation references).

The data set contains important features of the atmosphere such as clearly pronounced troposphere, stratosphere, mesosphere and thermosphere due to the temperature inflection points. Also, the polar and mesospheric jets are distinctly visible. A typical wave envelope profile is generated in terms of standard numerical solvers for (32) (panel d). It grows exponentially with height as the background density decreases until viscosity starts to dominate damping the wave to disappearance. As a final remark, we must recognize that the data set does not reach far enough into the higher atmosphere to observe the far-field behavior as discussed in Section 3.3 and shown in [25].

6 Modulational stability of the stationary solution

This section is dedicated to the stability of the stationary solution of the modulation equations. In order to assess stability we linearize the governing equations (27) and the boundary condition (22) around the stationary solution (29) and apply the ansatz for the perturbation

$$y(Z, T) = \hat{y}(Z)e^{\lambda T}. \quad (38)$$

This transforms the problem of stability into a boundary eigenvalue problem (BEVP),

$$\lambda \hat{y} + \frac{\partial}{\partial Z} \left[ D_F(Y) \hat{y} \right] = D_G(Y)\hat{y} \quad \text{for} \quad Z \in (0, \infty),$$

$$D_B(Y)\hat{y} = 0 \quad \text{at} \quad Z = 0 \quad (39a)$$

with the Jacobian matrices

$$D_F(Y) = \begin{pmatrix} \hat{\Omega}' & K_x^2 \\ \hat{\Omega}'' \Lambda & \hat{\Omega}' \end{pmatrix}, \quad (40a)$$

$$D_G(Y) = \begin{pmatrix} 0 & 0 \\ -\eta \hat{\Omega}'' \Lambda - 2\Lambda K_x \Lambda & -\eta \hat{\Omega}' - \Lambda |K|^2 \end{pmatrix}, \quad (40b)$$

$$D_B(Y) = \begin{pmatrix} \hat{\Omega}' & K_x^2 \\ -2\Lambda \hat{\Omega}' & -2\hat{\Omega} \end{pmatrix}. \quad (40c)$$
Fig. 3: Illustrative example. CIRA-86 zonally averaged temperature (panel a) and background horizontal wind (panel d, thick blue line) for March at 50° N. Kinematic viscosity (panel b), Brunt-Väisälä frequency (panel c) computed by (3a). Vertical wave number (panel e) and specific wave action density (panel f) computed by (31) and (32a), respectively, assuming horizontal wavenumber $K^*_x = 0.2 \text{ km}^{-1}$. Wave-induced mean flow (panel d, dashed blue line).
Solving the BEVP really means to find the spectrum of the linear differential operator $\mathcal{L}$ defined due to (39). The wave is stable if there is no spectrum on the right-hand side of the complex plane. We can decompose the spectrum into the essential (continuous) and the point (matrix-like) spectrum. A comprehensive introduction in this method can be found in [32]. In the following sections, we will study each part of the spectrum individually.

### 6.1 Essential spectrum

The linear operator $\mathcal{L}$ of the BEVP can be approximated by an asymptotic differential operator having constant coefficients, $\mathcal{L}_{\infty}$. Utilizing Fredholm operator theory one can prove that it has the same essential spectrum as the original operator [26]. The BEVP of the asymptotic operator can be reformulated as an initial value problem

$$\frac{\partial \hat{Y}}{\partial Z} = C_{\infty}(\lambda) \hat{Y} \quad \text{for} \quad Z \in (0, \infty),$$

$$D_B(Y) \hat{Y} = 0 \quad \text{at} \quad Z = 0$$

(41a)

(41b)

where the constant coefficient matrix is given by

$$C_{\infty}(\lambda) = \lim_{Z \to +\infty} C(Z, \lambda),$$

$$C(Z, \lambda) = D^{-1}_Y(\theta) \left( D_Y(\theta) - \frac{\partial D_Y(\theta)}{\partial Z} - \lambda \right).$$

(42)

(43)

Existence of the limit is granted due to the assumptions of section 3.3.

The asymptotic operator $\mathcal{L}_{\infty} - \lambda$, and hence the original operator $\mathcal{L} - \lambda$, are Fredholm if $C_{\infty}$ is hyperbolic, i.e. all its eigenvalues have non-zero real part. We find two distinct spatial eigenvalues

$$\nu_1(\lambda) = -\frac{\lambda}{\hat{\Omega}_{\infty}},$$

$$\nu_2(\lambda) = -\frac{\lambda + \eta_{\infty} \hat{\Omega}'_{\infty} + A_{\infty} |K_{\infty}|^2}{\hat{\Omega}_{\infty}}.$$  

(44a)

(44b)

Thus, the Morse index, which is defined as the dimension of the unstable subspace of a hyperbolic matrix, is

$$i_{\infty}(\lambda) = \begin{cases} 
0 & \text{if} \quad 0 < \Re(\lambda), \\
1 & \text{if} \quad -\eta_{\infty} \hat{\Omega}'_{\infty} - A_{\infty} |K_{\infty}|^2 < \Re(\lambda) < 0, \\
2 & \text{if} \quad \Re(\lambda) < -\eta_{\infty} \hat{\Omega}'_{\infty} - A_{\infty} |K_{\infty}|^2.
\end{cases}$$

(45)

Note that $-\eta_{\infty} \hat{\Omega}'_{\infty} - A_{\infty} |K_{\infty}|^2 < 0$ due to (37). On lines in the complex plane where $\Re(\lambda) = 0$ and $\Re(\lambda) = -\eta_{\infty} \hat{\Omega}'_{\infty} - A_{\infty} |K_{\infty}|^2$ the matrix $C_{\infty}$ is not hyperbolic and hence the operator is not Fredholm.

The Fredholm index tells us where the essential spectrum lies. According to [33, their formula 1.10] and also [34, p 391], it can be written as

$$\text{ind} = \dim(\ker D_B(Y_0)) - i_{\infty}(\lambda)$$

(46)

where $Y(0) = Y_0$. The essential spectrum is the set of $\lambda$’s for which the operator $\mathcal{L} - \lambda$ is Fredholm but ind $\neq 0$ or it is not Fredholm. The point spectrum, on the other hand, lies where the operator is Fredholm and ind $= 0$ but the operator is not invertible. We will investigate the point spectrum in section 6.2.

Having a closer look on (45), it turns out that

$$\dim(\ker D_B(Y_0)) = 0$$

(47)

must be true in order to obtain a stable essential spectrum, i.e. no essential spectrum on the right hand side of the complex plane. In particular, if the kernel is non-empty, the waves are unstable and the operator is even ill-posed as the complete right plane is in the essential spectrum.
The criterion can be rephrased: stable waves necessitate Dirichlet boundary conditions. This is violated and hence the wave destabilizes due to perturbations from the essential spectrum if

$$\frac{H_m N}{|U|} = \sqrt{2} \quad \text{at} \ Z = 0$$

which can be called the gross wave-Froude number. This instability criterion is not determined by an inequality like most fluid dynamical stability criteria. In fact the instability condition has to be fulfilled by equality. It is therefore more similar to a catastrophic resonance condition.

6.2 Point spectrum

In this section we will prove the non-existence of unstable point spectrum. Let us assume a stable essential spectrum and the existence of an unstable eigenvalue. For this eigenvalue the Fredholm index is zero and hence it may belong to the point spectrum. Then, the eigenfunction solves the associated ODE

$$\frac{\partial \hat{y}}{\partial Z} = C(Z, \lambda) \hat{y}.$$  \hspace{1cm} (49)

By assumption, the essential spectrum is stable and hence the kernel of the Jacobian of the boundary condition is empty or, in other words, we get a Dirichlet boundary condition, \( \hat{y} = 0 \) at \( Z = 0 \). The Dirichlet boundary is the initial condition for the ODE which then assumes the trivial solution, so there is no eigenfunction. This contradiction completes the proof that there is no unstable point spectrum. In conclusion, modulational instabilities, given that the resonance condition is true, originate from the essential spectrum.

7 Summary and conclusion

Let us summarize the results of the previous section. The gross wave-Froude number depending on the wave itself due to the induced mean flow is specified in terms of (34) and (35) by

$$\frac{H_m N}{|U|} = \frac{2Fr_{\text{wave}}}{\sqrt{1 - 2Fr_{\text{wave}}^2} + 1} \quad \text{at} \ Z = 0.$$  \hspace{1cm} (50)

In conclusion, stationary waves extending from the surface to the deep atmosphere, that experience

$$Re_{\text{wave}} < 1 \quad \text{above} \ Z_{\text{turbo}},$$

are

- non-evanescent and statically stable if

$$\begin{cases} 0, & (h), \\ H_m K_x, & (nh) \end{cases} \frac{H_m N}{|U|} < \begin{cases} 1, & (h), \\ \sqrt{1 + H_m^2 K_x^2}, & (nh) \end{cases} \quad \text{at} \ Z = 0.$$  \hspace{1cm} (52)

Here, the lower bound originates from (31). It provides a real-valued and negative vertical wavenumber. The upper bound is due to (20) and guarantees static stability.

- modulationally stable if

$$\frac{H_m N}{|U|} \neq \sqrt{2} \quad \text{at} \ Z = 0.$$  \hspace{1cm} (53)

This non-resonance criterion stems from the investigation of the essential spectrum in section 6.1. An illustration combining these criteria for the wave-Froude number computed by inversion of (50) is plotted in figure 4. It turns out that the resonance condition for the wave-Froude number is the same as the threshold
that guarantees real-valued wave action in (34). Combining this result with the upper bound for real-valued wave action (35), we obtain a strict inequality,

$$Fr_{\text{wave}} < \frac{1}{\sqrt{2}},$$

for the existence and modulational stability of the wave. That the critical values are the same does not come as a surprise because $A^+$ and $A^-$ coincide at this value which is linked to the singularity of the Jacobian of the boundary condition by the inverse function theorem.

Also, beyond $H_m K_x = \sqrt{2}$ waves cease to exist. Only solutions where $H_m K_x > 1$ may become modulationally unstable since waves in the region $H_m K_x < 1$ overturn before they ever reach the resonance condition.

The main result of this paper is that the stability of stationary strongly nonlinear gravity waves with respect to Grimshaw’s dissipative modulation equations depends on three characteristic parameters: The wave-Reynolds number, the wave-Froude number and the mountain’s height-period ratio. Considering vertically modulated waves from the surface to the deep atmosphere, we find that the stability is completely determined by the boundary and far-field conditions. These results are valid for fairly general wave solutions that posses only minor restrictions on the background fields: the background is hydrostatic and exhibits a physical far-field behavior. Other than this, the background temperature and horizontal wind are unconditioned.

We want to give some remarks on the characteristic parameters that we hope are particularly useful for interpreting measurements and numerical simulations of atmospheric gravity waves. When reformulated in dimensional variables, the wave-Reynolds number reads

$$Re_{\text{wave}} = \frac{C_{g\ell} D^*}{A^*},$$

where $C_{g\ell}$ represents the vertical group velocity and $D^* = H_p^{-1} |K^*|^{-2}$ defines a length scale with $H_p$ the local pressure scale height.

The net wave-Froude number as defined in this work is readily redimensionalized, so

$$Fr_{\text{wave}} = \frac{H_m N^*}{|U^*|}.$$
It has to be distinguished from the gross wave-Froude number as the latter contains the induced mean-flow and is therefore not independent of the wave itself. In the literature, only the gross wave-Froude number is considered which is reasonable since from observations one gets the mean-flow horizontal wind. Usually, it is not feasible to ask for the background flow of a mountain wave which really is the flow without the mountain or alternatively the flow far upstream. However, knowing the wave parameters of the excited wave, it is possible to compute the background wind by the total momentum equation (26). Also, in weakly nonlinear theory they are indeed the same as the induced mean flow is a higher-order correction. Still, the strongly nonlinear description in this paper reduces the mean-flow wind and therefore gains wave energy from the mean flow when excited. Hence, from a theoretical point of view the net wave-Froude number should not depend \textit{a priori} on the wave that is excited by the background flow over the mountain.

In an envisaged companion paper we want to extend our investigations to the stability of gravity waves governed by three-dimensional modulation equations including the Coriolis force. The basis for such a study was already founded in [35].

Acknowledgments: This research was supported by the German Research Foundation (DFG) through Grants KL 611/25-2 of the Research Unit FOR1898 and Research Fellowship SCHL 2195/1-1. Comments of Prof. Rupert Klein, that helped to improve the manuscript, are gratefully acknowledged.

References

[1] Fritts DC, Alexander MJ. Gravity wave dynamics and effects in the middle atmosphere. Reviews of Geophysics. 2003;41(1):1003.
[2] Becker E. Dynamical control of the middle atmosphere. Space Science Reviews. 2012;168(1-4):283–314.
[3] Fritts DC, Smith RB, Taylor MJ, Doyle JD, Eckermann SD, Dörnbrack A, et al. The Deep Propagating Gravity Wave Experiment (DEEPWAVE): An Airborne and Ground-Based Exploration of Gravity Wave Propagation and Effects from Their Sources throughout the Lower and Middle Atmosphere. Bulletin of the American Meteorological Society. 2016 mar;97(3):425–453.
[4] Fritts DC, Vosper SB, Williams BP, Bossert K, Plane JMC, Taylor MJ, et al. Large-Amplitude Mountain Waves in the Mesosphere Accompanying Weak Cross-Mountain Flow During DEEPWAVE Research Flight RF22. Journal of Geophysical Research: Atmospheres. 2018 sep;123(18):9992–10022.
[5] Fritts DC, Wang L, Taylor MJ, Pautet PD, Criddle NR, Kaifler B, et al. Large-Amplitude Mountain Waves in the Mesosphere Observed on 21 June 2014 During DEEPWAVE: 2. Nonlinear Dynamics, Wave Breaking, and Instabilities. Journal of Geophysical Research: Atmospheres. 2019 sep;124(17-18):10006–10032.
[6] Becker E. Direct heating rates associated with gravity wave saturation. Journal of Atmospheric and Solar-Terrestrial Physics. 2004;66(6-9):683–696.
[7] Schlutow M, Becker E, Körnich H. Positive definite and mass conserving tracer transport in spectral GCMs. Journal of Geophysical Research: Atmospheres. 2014;119(20):56,577.
[8] Elsassen A, Palm E. On the Transfer of Energy in Stationary Mountain Waves. Geofysiske Publikasjoner. 1961;22:1–23.
[9] Pütz C, Schlutow M, Klein R. Initiation of ray tracing models: evolution of small-amplitude gravity wave packets in non-uniform background. Theoretical and Computational Fluid Dynamics. 2019 oct;33(5):509–535.
[10] Err M, Trinh QT, Preusse P, Gille JC, Mlynczak MG, Russell III JM, et al. GRACILE: a comprehensive climatology of atmospheric gravity wave parameters based on satellite limb soundings. Earth System Science Data. 2018 apr;10(2):857–892.
[11] Whitham GB. Linear and Nonlinear Waves. John Wiley & Sons, Inc.; 1974.
[12] Grimshaw R. The Modulation of an Internal Gravity-Wave Packet, and the Resonance with the Mean Motion. Studies in Applied Mathematics. 1977 jun;56(3):241–266.
[13] Sutherland BR. Finite-amplitude internal wavepacket dispersion and breaking. Journal of Fluid Mechanics. 2001 feb;429:343–380.
[14] Sutherland BR. Weakly nonlinear internal gravity wavepackets. Journal of Fluid Mechanics. 2006;569:249–258.
[15] Tabaei A, Akylas TR. Resonant long-short wave interactions in an unbounded rotating stratified fluid. Studies in Applied Mathematics. 2007;119(3):271–296.
[16] Schlutow M, Wahlén E, Birken P. Spectral stability of nonlinear gravity waves in the atmosphere. Mathematics of Climate and Weather Forecasting. 2019;5(1):12–33.
[17] Schlutow M. Modulational Stability of Nonlinear Saturated Gravity Waves. Journal of the Atmospheric Sciences. 2019 nov;76(11):3327–3336.
[19] Grimshaw R. Nonlinear internal gravity waves in a slowly varying medium. Journal of Fluid Mechanics. 1972;54(2):193–207.
[20] Grimshaw R. Internal gravity waves in a slowly varying, dissipative medium. Geophysical Fluid Dynamics. 1974;6:131–148.
[21] Achatz U, Klein R, Senf F. Gravity waves, scale asymptotics and the pseudo-incompressible equations. Journal of Fluid Mechanics. 2010;663:120–147.
[22] Schlutow M, Klein R, Achatz U. Finite-amplitude gravity waves in the atmosphere: travelling wave solutions. Journal of Fluid Mechanics. 2017 sep;826:1034–1065.
[23] Lilly DK, Klemp JB. The effects of terrain shape on nonlinear hydrostatic mountain waves. Journal of Fluid Mechanics. 1979;95(2):241–261.
[24] Bölöni G, Ribstein B, Muraschko J, Sgoff C, Wei J, Achatz U. The interaction between atmospheric gravity waves and large-scale flows: An efficient description beyond the nonacceleration paradigm. Journal of the Atmospheric Sciences. 2016 aug;73(12):4833–4852.
[25] Schmidtke G, Suchy K, Rawer K. Geophysik III / Geophysics III. vol. 10 / 49 / of Handbuch der Physik / Encyclopedia of Physics. Rawer K, editor. Berlin, Heidelberg: Springer Berlin Heidelberg; 1984.
[26] Kapitula T, Promislow K. Spectral and Dynamical Stability of Nonlinear Waves. vol. 185 of Applied Mathematical Sciences. New York, NY: Springer New York; 2013.
[27] Jeffreys H. On the dynamics of wind. Quarterly Journal of the Royal Meteorological Society. 1922 aug;48(201):29–48.
[28] Mayer FT, Fringer OB. An unambiguous definition of the Froude number for lee waves in the deep ocean. Journal of Fluid Mechanics. 2017;831:1–9.
[29] Nicolet M. Dynamic Effects of the High Atmosphere. In: Kuiper GP, editor. The Earth as a Planet. The University of Chicago Press; 1954. p. 644–712.
[30] Fleming EL, Chandra S, Barnett JJ, Corney M. Zonal mean temperature, pressure, zonal wind and geopotential height as functions of latitude. Advances in Space Research. 1990 jan;10(12):11–59.
[31] Walterscheid RL, Hickey MP. Group velocity and energy flux in the thermosphere: Limits on the validity of group velocity in a viscous atmosphere. Journal of Geophysical Research Atmospheres. 2011 jun;116(D12):D12101.
[32] Sandstede B. Stability of travelling waves. In: Fiedler B, editor. Handbook of dynamical systems. vol. 2. Gulf Professional Publishing; 2002. p. 983–1055.
[33] Ben-Artzi A, Gohberg I, Kaashoek MA. Invertibility and dichotomy of differential operators on a half-line. Journal of Dynamics and Differential Equations. 1993 jan;5(1):1–36.
[34] Gohberg I, Goldberg S, Kaashoek MA. Classes of Linear Operators Vol. I. Basel: Birkhäuser Basel; 1990.
[35] Achatz U, Ribstein B, Senf F, Klein R. The interaction between synoptic-scale balanced flow and a finite-amplitude mesoscale wave field throughout all atmospheric layers: weak and moderately strong stratification. Quarterly Journal of the Royal Meteorological Society. 2017;143(702):342–361.