Modulational Stability of Nonlinear Saturated Gravity Waves

MARK SCHLUTOW

Institut für Mathematik, Freie Universität Berlin, Berlin, Germany

ABSTRACT

Stationary gravity waves, such as mountain lee waves, are effectively described by Grimshaw's dissipative modulation equations even in high altitudes where they become nonlinear due to their large amplitudes. In this theoretical study, a wave-Reynolds number is introduced to characterize general solutions to these modulation equations. This nondimensional number relates the vertical linear group velocity with wavenumber, pressure scale height, and kinematic molecular/eddy viscosity. It is demonstrated by analytic and numerical methods that Lindzen-type waves in the saturation region, that is, where the wave-Reynolds number is of order unity, destabilize by transient perturbations. It is proposed that this mechanism may be a generator for secondary waves due to direct wave–mean-flow interaction. By assumption, the primary waves are exactly such that altitudinal amplitude growth and viscous damping are balanced and by that the amplitude is maximized. Implications of these results on the relation between mean-flow acceleration and wave breaking heights are discussed.

1. Introduction

Atmospheric gravity waves generated in the lee of mountains extend over scales across which the background may change significantly. The wave field can persist throughout the layers from the troposphere to the deep atmosphere, the mesosphere and beyond (Fritts et al. 2016, 2018). On this range background temperature and therefore stratification and background density may undergo several orders of magnitude in variation. Also, dynamic viscosity and thermal conductivity cannot be considered constant on such a domain (Pitteway and Hines 1963; Zhou 1995).

Two predominant regimes for the waves can be identified: the homosphere and the heterosphere (Nicolet 1954). These two are separated by the turbopause, usually somewhere in the mesopause region. Below the turbopause molecular viscosity is negligible. Hence, if diffusion occurs, it is caused by turbulence. Due to the missing damping and the thinning background air, mountain waves amplify exponentially when extending upward. This phenomenon is also called anelastic amplification.

At certain heights the amplitudes cannot grow any further due to limiting processes. Those may be static or dynamic instabilities that act on the small scale comparable to the wavelength. For instance, Klostermeyer (1991) showed that all inviscid nonlinear Boussinesq waves are prone to parametric instabilities. The waves do not immediately disappear by the small-scale instabilities, rather the perturbations grow comparably slowly such that the waves persist in their overall structure over several more wavelengths. However, turbulence is produced. Lindzen (1970) modeled the effect of turbulence on the wave by harmonic damping with a constant kinematic eddy viscosity. The eddy viscosity is exactly such that it saturates the wave, meaning that viscous damping and anelastic amplification are balanced (Lindzen 1981; Fritts 1984; Dunkerton 1989; Becker 2012). Pitteway and Hines (1963) referred to this instance as amplitude-balanced wave.

Above the turbopause molecular viscosity takes over. It is usually modeled by a constant dynamic viscosity. In combination with the thinning background density, the kinematic viscosity increases exponentially with height resulting in a rapid decrease in amplitude.

In the process of becoming saturated the amplitude becomes considerably large such that the waves cannot be considered linear. Pioneering work on the mathematical description of nonlinear gravity waves was accomplished by Grimshaw (1972, 1974). This paper aims to extend Lindzen's linear wave saturation theory with the aid of Grimshaw's nonlinear wave description.

Corresponding author: Mark Schlutow, mark.schlutow@fu-berlin.de

2. How the modulation equations solve the compressible Navier–Stokes equations—A brief review

The nonlinear governing equations for our investigations are the two-dimensional dissipative Grimshaw's modulation equations being first introduced by Grimshaw (1974). Before presenting them, we want to give a brief review in this section how they solve the dimensionless compressible, Reynolds-averaged Navier–Stokes equations (NSE) asymptotically. Detailed derivations can be also found in Achatz et al. (2010) and Schlutow et al. (2017). Length and time are nondimensionalized via

$$(\hat{x}, \hat{z}, \hat{t}) = \left(\frac{1}{L_r}x^*, \frac{1}{L_r}z^*, \frac{v_r}{L_r}t^*\right), \tag{1}$$

where $L_r \approx 1...10\,\text{km}$ denotes the reference wavelength (hence the subscript r) and v_r is the reference velocity. Note that variables labeled with an asterisk denote dimensional quantities throughout the paper. To separate the hydrostatic background from the flow field associated with the wave the first ingredient necessary is a small scale separation parameter

$$\varepsilon = L_r/H_\theta \ll 1, \tag{2}$$

where $H_\theta \approx 10...100\,\text{km}$ is the potential temperature scale height. This choice for the scale separation parameter was introduced by Achatz et al. (2010).

The authors considered inviscid flows. To take viscous damping into account, we need to compare inviscid and viscous terms. The (eddy) viscosity is not a constant throughout the atmosphere and among others depends on temperature. Midgley and Liemohn (1966, their Fig. 9) gave a realistic vertical profile of the effective kinematic viscosity combining eddy and molecular effects. Hodges (1969) computed the eddy diffusion by gravity waves near the mesopause to be $K_{r(\text{eddy})} \approx 10^{6...7}\,\text{cm}^2\,\text{s}^{-1}$, which compares well with Midgley and Liemohn (1966) and Lindzen (1981). In the scaling regime of Achatz et al. (2010) as well as Schlutow et al. (2017) these values correspond to Reynolds numbers of $\text{Re} \approx 10...100$ when the Mach number is $\text{Ma} \approx 0.1...0.01$. These numbers are also supported by a review of Fritts (1984).

Taking these arguments into account, a realistic flow regime in terms of Mach, Froude, Reynolds, and Prandtl number most suitable for internal gravity waves in the middle/upper atmosphere region is then found by assuming

$$\frac{v_r}{\sqrt{(1-\kappa)p_r/\rho_r}} \equiv \text{Ma} = O(\varepsilon), \tag{3a}$$

$$\frac{v_r}{\sqrt{gL_r}} \equiv \text{Fr} = O(\varepsilon^{1/2}), \tag{3b}$$

$$\frac{\rho_r v_r L_r}{\mu_r} \equiv \text{Re} = O(\varepsilon^{-1}), \tag{3c}$$

$$\frac{c_p \mu_r}{\kappa_r} \equiv \text{Pr} = O(1), \tag{3d}$$

where ρ_r, p_r, κ_r, and μ_r represent the reference density, pressure, thermal conductivity, and dynamic viscosity, respectively. The constant $\kappa = 2/7$ is the ratio of the ideal gas constant R to the specific heat capacity at constant pressure c_p for diatomic gases. Equations (3) provide a distinguished limit for multiple-scale asymptotic analysis. Introducing compressed coordinates for the horizontal, vertical, and time axis,

$$(x, z, t) = (\varepsilon\hat{x}, \varepsilon\hat{z}, \varepsilon\hat{t}), \tag{4}$$

separates the slowly varying background from the fast oscillating wave fields. Under the assumption of stable stratification to the leading order, the hydrostatic background is determined by the vertical temperature profile $T(z)$. Two dimensionless background variables, which vary on the large scale, will appear in the final modulation equations, the Brunt–Väisälä frequency or stratification measure N and the background density ρ. They have to be calculated from the temperature by solving

$$N^2 = \frac{1}{T}\left(\frac{dT}{dz} + 1\right), \tag{5}$$

$$\frac{1}{\rho}\frac{d\rho}{dz} = -\frac{1}{T}\left(\frac{dT}{dz} + \frac{1}{\kappa}\right), \tag{6}$$

which originate from the hydrostatic assumption in combination with the ideal gas equation of state.

With the given scale separation of background and wave field, we can formulate the scaled two-dimensional compressible NSE,

$$\frac{D\mathbf{v}}{D\hat{t}} + \hat{\nabla}p - Nb\mathbf{e}_z = -\varepsilon(Nb\hat{\nabla}p - N^2 p\mathbf{e}_z) + \varepsilon\Lambda\hat{\nabla}^2\mathbf{v} + \dots, \tag{7a}$$

$$\frac{Db}{D\hat{t}} + Nw = -\varepsilon\left(N^2 + \frac{1}{N}\frac{dN}{dz}\right)wb + \varepsilon\Lambda\hat{\nabla}^2 b + \dots, \tag{7b}$$

$$\hat{\nabla}\cdot\mathbf{v} = -\varepsilon\left(N^2 + \frac{1}{\rho}\frac{d\rho}{dz}\right)w + \dots, \tag{7c}$$

in terms of the wave-field variables, velocity $\mathbf{v}$, buoyancy Nb, and kinematic pressure p, where $\hat{\nabla} = (\partial/\partial\hat{x}, \partial/\partial\hat{z})^\text{T}$

and $\mathbf{e}_z$ the unit vector pointing in the vertical direction; $D/D\hat{t}$ denotes the material derivative, and Λ is the total kinematic viscosity. The dots represent higher-order terms that we do not give explicitly but must not be neglected. The interested reader finds these terms in Schlutow et al. [2017, their (2.19)].

Wentzel–Kramers–Brillouin (WKB) theory provides us with a multiple-scale method to solve the scaled NSE asymptotically by the spectral ansatz

$$\mathbf{U}(\hat{x},\hat{z},\hat{t};\varepsilon) = \mathbf{U}_{0,0}(x,z,t) + [\mathbf{U}_{0,1}(x,z,t)e^{i\phi(x,z,t)/\varepsilon} + \text{c.c.}]$$
$$+ \text{h.h.} + O(\varepsilon), \qquad (8)$$

where c.c. stands for the complex conjugate and h.h. for higher harmonics. The vector $\mathbf{U} = (\mathbf{v}, b, p)^{\mathrm{T}}$ contains the prognostic variables and ϕ is the wave's phase.

We want to give a remark at this point. By construction, the WKB ansatz is a nonlinear approach. In the limit $\varepsilon \to 0$ the amplitudes are finite. In fact, the ansatz converges to the nonlinear plane wave of Boussinesq theory, which is known to be an analytical solution. Mean-flow interaction and Doppler shift are leading-order effects, which is in contrast to weakly nonlinear theory where they appear as higher-order corrections. In the ε limit, the weakly nonlinear approach converges to the linear plane wave.

The nonlinear WKB ansatz is inserted into the compressible NSE and terms are ordered with respect to powers of ε and harmonics.

3. Governing equations: Grimshaw's dissipative modulation equations

To leading order of the nonlinear WKB analysis of the NSE with the turbulence model of Lindzen one obtains Grimshaw's dissipative modulation equations supported by a mean-flow horizontal kinematic pressure gradient [Grimshaw 1974, their (4.1)–(4.6)]. Note that the leading-order WKB analysis of the dissipative pseudoincompressible equations (Durran 1989) lead to the same modulation equations. We present them in slightly different notation as Grimshaw and with the prerequisite that the wave field is horizontally periodic ($0 < k_x = \text{const}$) and only modulated in the z direction,

$$\frac{\partial k_z}{\partial t} + \frac{\partial \omega}{\partial z} = 0, \qquad (9a)$$

$$\rho\frac{\partial a}{\partial t} + \frac{\partial \omega' \rho a}{\partial z} = -\Lambda|\mathbf{k}|^2 \rho a, \qquad (9b)$$

$$\rho\frac{\partial u}{\partial t} + \frac{\partial \omega' k_x \rho a}{\partial z} = -\frac{\partial p}{\partial x}. \qquad (9c)$$

The modulation equations govern the evolution of vertical wavenumber k_z, wave action density ρa, and mean-flow horizontal wind u. Equations (9a)–(9c) are closed by

$$|\mathbf{k}|^2 = k_x^2 + k_z^2, \qquad (9d)$$

$$\omega = \frac{Nk_x}{|\mathbf{k}|} + k_x u, \qquad (9e)$$

$$\omega' = -\frac{Nk_x k_z}{|\mathbf{k}|^3}, \qquad (9f)$$

where $\mathbf{k}$, ω, and ω' represent the wavenumber vector, extrinsic frequency, and vertical linear group velocity, respectively. Note that primes denote derivative with respect to the vertical wavenumber throughout this paper. Extrinsic frequency is defined by the sum of intrinsic frequency and Doppler shift. It is linked to the wavenumber vector by the dispersion relation for nonhydrostatic gravity waves. It was shown in Schlutow et al. (2017) that the modulation equations equally hold for hydrostatic waves, where the horizontal wavelength is much larger than the vertical, if the dispersion relation is replaced by $\omega = Nk_x/|k_z| + k_x u$. The prognostic variables determine the asymptotic solution as described in the previous section and explained in Schlutow et al. (2017).

The first equation, (9a), essentially describes the evolution of the phase in (8) as by definition $\nabla\phi \equiv \mathbf{k}$ and $-\partial\phi/\partial t \equiv \omega$. Cross differentiating these expressions, they add up to zero. The second equation, (9b), governs the conservation of wave action density being the ratio of wave energy density and the intrinsic frequency. In terms of the polarization relation the leading-order first harmonics are computable via

$$\mathbf{U}_{0,1} = \mathscr{B}\left(-i\frac{k_z}{k_x}\frac{\omega - k_x u}{N}, i\frac{\omega - k_x u}{N}, 1, -i\frac{k_z}{k_x^2}\frac{(\omega - k_x u)^2}{N}\right)^{\mathrm{T}}, \qquad (10)$$

with $\mathscr{B} = \sqrt{(\omega - k_x u)a/2}$ the wave's buoyancy amplitude. The third equation, (9c), accounts for the acceleration of the mean flow. With a slight abuse of notation (we drop the indices) we obtain the zero harmonics,

$$\mathbf{U}_{0,0} = (u, 0, 0, p)^{\mathrm{T}}. \qquad (11)$$

The variable p corresponding to the mean-flow horizontal kinematic pressure is unknown and needs additional investigation to close the system (9).

The governing equations, (9), can be reformulated in vector form:

$$\frac{\partial \mathbf{y}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{y})}{\partial z} = \mathbf{G}(\mathbf{y}), \qquad (12)$$

with a flux $\mathbf{F}$ and an inhomogeneity $\mathbf{G}$ where $\mathbf{y} = (k_z, a, u)^{\mathrm{T}}$ is the prognostic vector.

4. General stationary solutions of the modulation equations

In this section we will explore general stationary solutions before we focus on particular solutions for which we will present stability analysis in the upcoming sections.

In the inviscid limit (i.e., $\Lambda \to 0$), the modulation equations assume stationary solutions where $\partial p/\partial x = 0$, which can be computed analytically by a formula of Schlutow et al. [2017, their (5.20)]. When we multiply (9b) by k_x and subtract (9c), we obtain

$$\rho \frac{\partial}{\partial t}(k_x a - u) = \frac{\partial p}{\partial x} - k_x \Lambda |\mathbf{k}|^2 \rho a. \qquad (13)$$

Thus, to be consistent with the inviscid limit, the dissipative modulation equations assume stationary solutions only if the right-hand side of (13) vanishes, which provides eventually a closure for the mean-flow horizontal kinematic pressure gradient,

$$\frac{\partial p}{\partial x} = k_x \Lambda |\mathbf{k}|^2 \rho a. \qquad (14)$$

This result implicates that the mean-flow horizontal kinematic pressure gradient balances the viscous forces acting on the horizontal mean-flow wind. Such a flow configuration is referred to as antitriptic flow in the literature (Jeffreys 1922).

Then a general stationary solution $\mathbf{Y}(z) = (K_z, A, U)^{\mathrm{T}}$ depicting a mountain lee wave fulfills

$$\omega(K_z, U) = 0, \qquad (15)$$

$$\frac{\partial \Omega' \rho A}{\partial z} = -\Lambda |\mathbf{K}|^2 \rho A, \qquad (16)$$

$$U = U(z), \qquad (17)$$

with $|\mathbf{K}|^2 = K_x^2 + K_z^2$ and $\Omega' = \omega'(K_z)$. Note that we label the stationary solution by capital letters. Given any well-behaved, slowly varying mean-flow horizontal wind $U(z)$, the remaining two variables of the general stationary solution are given explicitly by

$$K_z = -\sqrt{\frac{N^2}{U^2} - K_x^2}, \qquad (18)$$

$$A = \frac{\mathscr{F}}{\rho \Omega'} \exp\left(-\int_0^z \frac{\Lambda |\mathbf{K}|^2}{\Omega'} d\tilde{z} \right). \qquad (19)$$

with $\mathscr{F}$ being the vertical wave action flux at $z = 0$.

5. Lindzen-type mountain lee wave and the wave-Reynolds number

So far, we considered all background variables of our governing PDE as functions of z. In this section we will prescribe these functions in a piecewise fashion in order to construct a typical mountain lee wave that gets saturated by some small-scale instability process comparable to Lindzen (1981). The complete solution is divided into an unsaturated as well as a saturated middle-atmospheric part and a deep-atmosphere part.

a. The unsaturated middle-atmospheric solution

First, we assume that the background atmosphere is piecewise isothermal. From (5), this assumption implies that $N = $ const, some typical value for the middle atmosphere. Then (6) gives

$$\rho(z) = \rho_0 e^{-z/H}, \qquad (20)$$

where $H = \kappa N^{-2} = $ const denotes the dimensionless (local) pressure scale height. Second, we assume that the mean-flow horizontal wind is piecewise constant as well, so $U = $ const. These assumptions can be weakened, which we will discuss in the concluding section 7. It follows immediately by (18) and horizontal periodicity that $K_z = $ const making it a plane wave.

In the middle atmosphere viscosity is negligible. Therefore, below the breaking height z_{break} the integral in (19) vanishes and the amplitude of the wave grows exponentially with the inverse density,

$$A(z) = \frac{A_0}{\rho(z)} \quad \text{in} \quad [0, z_{\text{break}}). \qquad (21)$$

Because of the vanishing dissipation, the mean-flow horizontal kinematic pressure gradient also vanishes according to (14).

b. The saturated middle-atmospheric solution

Above z_{break}, the wave saturates by some small-scale instability process producing turbulence, which balances the anelastic amplification. The exact kinematic eddy viscosity that keeps the amplitude leveled in (19), such that $A = $ const, is then given by

$$\Lambda = \Omega' H^{-1} |\mathbf{K}|^{-2} = \text{const} \quad \text{in} \quad [z_{\text{break}}, z_{\text{turbo}}). \qquad (22)$$

In this region, the mean-flow horizontal kinematic pressure gradient depends on z only by the background density and can be computed inserting (22) into (14),

$$\frac{\partial P}{\partial x} = K_x H^{-1} \Omega' A \rho \quad \text{in} \quad [z_{\text{break}}, z_{\text{turbo}}). \qquad (23)$$

In particular, it is positive while the mean-flow horizontal wind is negative, which can be seen from (9e) and (15).

c. The deep-atmosphere solution

The saturated middle-atmospheric solution is valid below the turbopause at z_{turbo} as in the heterosphere the molecular viscosity dominates, which is modeled by a constant dynamic viscosity μ_{mol} implying

$$\Lambda(z) = \frac{\mu_{\text{mol}}}{\rho(z)} \quad \text{in} \quad [z_{\text{turbo}}, +\infty). \qquad (24)$$

In the deep-atmospheric region the integral in (19) has also an analytic solution for A, which decays quickly for $z \rightarrow +\infty$. Here, typical values N and U for the deep atmosphere are used.

d. The wave-Reynolds number

In Fig. 1, such a prototypical mountain wave, which saturates, is illustrated. By inspection of (22), its behavior in the different regions may be characterized by a "wave-Reynolds" number. It can be written as

$$\frac{\Omega' H^{-1} |\mathbf{K}|^{-2}}{\Lambda} \gtreqqless 1. \qquad (25)$$

When reintroducing the dimensional variables by reversing the nondimensionalization of Schlutow et al. (2017), so

$$H = \frac{\varepsilon}{L_r} H_p^*, \quad \Lambda = \frac{\rho_r}{\mu_r} \nu^*, \quad N = \frac{L_r}{v_r} N^*, \quad \mathbf{K} = L_r \mathbf{K}^*, \quad (26)$$

and using the definitions (3), the wave-Reynolds number reads

$$\text{Re}_{\text{wave}} \equiv \frac{C_{\text{gz}} D}{\nu^*}. \qquad (27)$$

The dimensional linear vertical group velocity is denoted by C_{gz} and represents the velocity scale for this type of Reynolds number. The term $D = H_p^{*-1} |\mathbf{K}|^{*-2}$ defines its length scale. Estimates of the wave-Reynolds number in the different regions are depicted in Fig. 1.

6. Stability of the saturated wave

In this section we will investigate the saturated mountain wave with respect to stability. In particular, we are interested in the region where $\text{Re}_{\text{wave}} = O(1)$ as here the

amplitude is at its maximum and one can expect most likely nonlinear behavior. We assess stability by analyzing the evolution of small perturbations. Due to their smallness, we can linearize the governing equations in vector form (12) around the stationary solution $\mathbf{Y}$. Applying the ansatz

$$\mathbf{y}(z, t) = \hat{\mathbf{y}}(z) e^{\lambda t} \qquad (28)$$

for the perturbation, results in an eigenvalue problem (EVP):

$$\lambda \mathbf{y} + \frac{\partial}{\partial z}\left(\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\bigg|_{\mathbf{Y}} \mathbf{y}\right) = \frac{\partial \mathbf{G}}{\partial \mathbf{y}}\bigg|_{\mathbf{Y}} \mathbf{y}, \qquad (29)$$

where we dropped the hat over $\mathbf{y}$. The Jacobian matrices of the flux and inhomogeneity evaluated at $\mathbf{Y}$ are given by

$$\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\bigg|_{\mathbf{Y}} = \begin{pmatrix} \Omega' & 0 & K_x \\ \Omega'' A & \Omega' & 0 \\ K_x \Omega'' A & K_x \Omega' & 0 \end{pmatrix}, \qquad (30)$$

$$\frac{\partial \mathbf{G}}{\partial \mathbf{y}}\bigg|_{\mathbf{Y}} = \begin{pmatrix} 0 & 0 & 0 \\ H^{-1} \Omega'' A - 2\Lambda K_z A & 0 & 0 \\ K_x H^{-1} \Omega'' A - 2 K_x \Lambda K_z A & 0 & 0 \end{pmatrix}. \qquad (31)$$

Note that the second and third row of each Jacobian are linearly dependent. Hence, we can solve the third equation of (29) for

$$u = K_x a, \qquad (32)$$

which reduces the dimension of the system, so with a slight abuse of notation, $\mathbf{y} = (k_z, a)^{\text{T}}$ and

$$\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\bigg|_{\mathbf{Y}} = \begin{pmatrix} \Omega' & K_x^2 \\ \Omega'' A & \Omega' \end{pmatrix}, \qquad (33)$$

$$\frac{\partial \mathbf{G}}{\partial \mathbf{y}}\bigg|_{\mathbf{Y}} = \begin{pmatrix} 0 & 0 \\ H^{-1} \Omega'' A - 2\Lambda K_z A & 0 \end{pmatrix}. \qquad (34)$$

If one assumes that the perturbation decays sufficiently fast toward the edges of the region where $\text{Re}_{\text{wave}} = O(1)$ (in Fig. 1 at 70 and 110 km), in other words, that the edges have no influence on the perturbation, then one can extend the domain to the infinities. Consequently, we consider the differential operator associated with the EVP as closed and densely defined on L^2, the space of vector valued square integrable functions on the real line equipped with the norm

$$\|\mathbf{y}\|_{L^2} = \sqrt{\int_{-\infty}^{+\infty} k_z^2 + a^2 + u^2 \, dz}. \qquad (35)$$

In conclusion, we translated the problem of stability to finding the spectrum of a linear operator of an EVP.
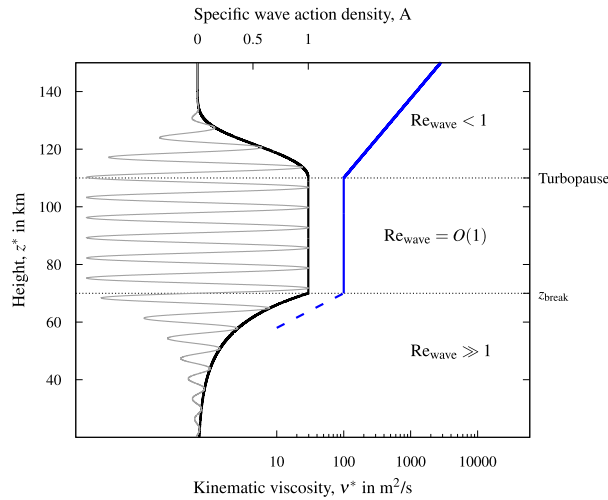
FIG. 1. Schematic of a saturated plane mountain lee wave (thin gray line) with amplitude profile (thick black line) and effective kinematic viscosity (thick blue line) being characterized by the wave-Reynolds number.

Broadly speaking, the spectrum is the set of all $\lambda \in \mathbb{C}$ for which the EVP has admissible solutions.

The EVP is solved by the Fourier transform

$$\mathbf{y} = \int_{-\infty}^{+\infty} \tilde{\mathbf{y}} e^{i\mu z} d\mu, \tag{36}$$

which yields an algebraic equation

$$\left(\lambda + \frac{\partial \mathbf{F}}{\partial \mathbf{y}} \bigg|_{\mathbf{Y}} i\mu - \frac{\partial \mathbf{G}}{\partial \mathbf{y}} \bigg|_{\mathbf{Y}} \right) \tilde{\mathbf{y}} = 0. \tag{37}$$

It has nontrivial solutions only if the coefficient matrix is singular $[\det(M) = 0]$, which happens if

$$(\lambda + \Omega' i\mu)^2 - K_x^2 A[(2\Lambda K_z - H^{-1}\Omega'')i\mu - \Omega''\mu^2] = 0. \tag{38}$$

This characteristic polynomial has two zeros,

$$\lambda_{1,2}(\mu) = -\Omega' i\mu \pm K_x \sqrt{A} \sqrt{(2\Lambda K_z - H^{-1}\Omega'')i\mu - \Omega''\mu^2}, \tag{39}$$

which determine the spectrum of the linear operator of the EVP as curves in the complex plane parameterized by the spatial eigenvalue μ, which can also be interpreted as the vertical wavenumber of the perturbation. The real part of λ represents the instability growth rate, if it is positive, and the imaginary part is a frequency. Thus, (39) provides also a dispersion relation linking the perturbation's wavenumber with its frequency.

One can readily show that either λ_1 or λ_2 of (39) has positive real part for reasonable wave solutions implying that they are unconditionally unstable. We want to point out that when $H \to \infty$ and $\Lambda \to 0$, (39) reduces to the spectrum of the inviscid nonlinear Boussinesq plane waves (Schlutow et al. 2019) having the classical modulational stability criterion $\Omega'' > 0$ (Grimshaw 1977).

a. Transient (in)stability

In this subsection we want to investigate the characteristics of the instability that is presented in the previous section. We put the "in" of the section title into parentheses because there is the possibility that the primary wave "survives" the instability. One particular type of those harmless instabilities is called transient instability. Despite the fact that the instability's norm grows exponentially in time, the instability vanishes at each given point in space if one waits long enough.

To show such characteristics a prerequisite for mathematical rigor must be fulfilled: the linear differential operator of the EVP must be well-posed, which means its spectrum has a maximum real part or, in other words, the instability growth rate is bounded. The reader finds this tedious verification in the appendix. We want to give an interesting remark at this point. The "well-posedness" depends on a criterion, $\Omega'' > 0$, which turns out to be equivalent to the modulational stability criterion of plane waves in Boussinesq theory.

The key idea for transient instabilities (Kapitula and Promislow 2013) is to ask for solutions of the EVP, (29), in a weighted space L_α^2 with $\alpha \in \mathbb{R}$ an exponential weight, such that

$$\|\mathbf{y}\|_{L_\alpha^2} = \|\mathbf{y} e^{\alpha z}\|_{L^2}. \tag{40}$$

Doing so, we get the same curves as in (39) but with $i\mu \to i\mu - \alpha$, so

$$\lambda_{1,2}^\alpha(\mu) = -\Omega'(i\mu - \alpha) \pm K_x \sqrt{A} \sqrt{(2\Lambda K_z - H^{-1}\Omega'')(i\mu - \alpha) + \Omega''(i\mu - \alpha)^2}. \tag{41}$$

When we choose

$$\alpha < \frac{1}{2} \frac{2\Lambda K_z - H^{-1}\Omega''}{\sqrt{\Omega''}\Omega' + \Omega''} < 0, \tag{42}$$

provided $\Omega'' > 0$ the spectrum is stabilized, so $\mathrm{Re}(\lambda_{1,2}^\alpha) \leq 0$ for all $\mu \in \mathbb{R}$. The negative exponential weight penalizes solutions at $-\infty$. To converge in the weighted norm, the perturbation in L^2 must therefore
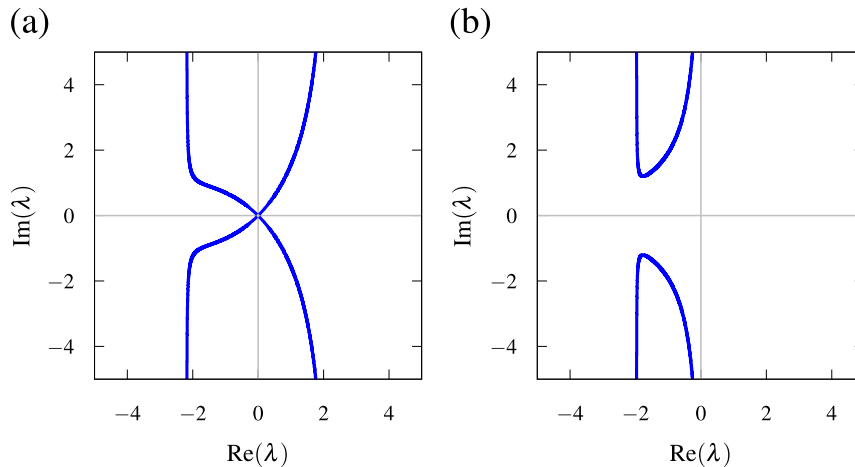
(a)

(b)

FIG. 2. (a) Unstable spectrum (blue lines) in L^2 and (b) stable spectrum in the exponentially weighted L_α^2 for the saturated wave of Fig. 3. The exponential weight is the largest α as defined by (42).

decay exponentially at $-\infty$ for all times. But simultaneously, its L^2 norm grows exponentially in time. This seeming paradox resolves when the perturbation propagates sufficiently fast toward $+\infty$ (i.e., upward). Perturbations of this behavior are called transient instabilities (Sandstede and Scheel 2000). For illustrative purpose, the unweighted and weighted spectrum of an example wave, which we will discuss in more detail in the following section, are plotted in Fig. 2.

b. Numerical investigation of the transient instabilities

We use the finite-volume numerical solver presented in Schlutow et al. (2019) for the governing equations, (9), to compute the evolution of a tiny Gaussian initial perturbation of the saturated wave in the region where $\mathrm{Re}_{\mathrm{wave}} = O(1)$. The results are shown in Fig. 3. The simulation is set up by $N = 1$ and $K_x = 1$. We discretize the equations on 2000 grid points in z and integrate in t over 6000 time steps. As can be seen in the figure, the perturbation amplifies exponentially and propagates to the right (i.e., upward), as theory suggests. We can also observe that the perturbation's wavelength compares with the scale height. Or in other words the instability varies on the large scale. The corresponding spectra to this case as computed by (39) and (41) are plotted in Fig. 2. The amplitude and the vertical wavenumber undergo strong modulations due to the exponentially growing perturbation. Furthermore, the initially constant mean-flow horizontal wind experiences acceleration. The analytic maximum instability growth rate according to (A6) is 2.2 for this particular case. In terms of an approximated L_2 norm that we compute numerically, the actual growth rate of the Gaussian

perturbation is found to be 1.6. The difference between theoretical and observed rate occurs because the perturbation is not optimal.

7. Summary and discussion

In this paper we investigated nonlinear mountain waves, which are governed by Grimshaw's dissipative modulation equations being asymptotically consistent with the compressible Navier–Stokes equations and the dissipative pseudoincompressible equations likewise.

We introduced a wave-Reynolds number characterizing the stationary solution. When this dimensionless quantity is of order unity, the wave amplitude saturates by small-scale instabilities and assumes a maximum as anelastic amplification by the thinning background air is exactly balanced by the turbulent damping. We analyzed this regime with respect to modulational stability as nonlinearities dominate for large-amplitude waves. It turned out that transient instabilities emerge that propagate upward. We tested this analysis solving the modulation equations numerically and found excellent agreement with the theory.

In the framework of saturated nonlinear wave theory, which we presented in this paper, the saturated mountain wave does initially not accelerate the mean-flow horizontal wind. Instead, a mean-flow horizontal kinematic pressure gradient emerges that keeps the horizontal wind constant by balancing the viscous forces. The wave persists structurally and loses energy directly to turbulence, which in turn damps altitudinal amplification. Eventually, the mean flow is accelerated by a
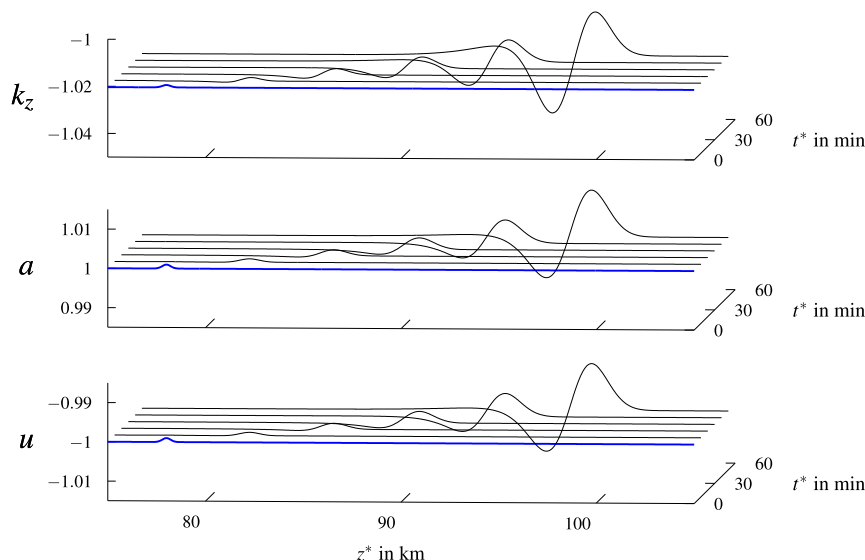
FIG. 3. Numerical corroboration of the governing equations, (9). Evolution of an initial Gaussian perturbation of the saturated wave centered at 78 km (blue lines).

transient instability in the saturation zone propagating upward while growing and transferring kinetic energy to the mean flow.

Our investigations have two implications being of interest for gravity wave parameterizations in numerical weather prediction and climate modeling. First, the induced mean flow behaves wavelike. Its evolution is governed by a dispersion relation for the linear perturbation. In conclusion, it may be interpreted as an upward-traveling secondary wave of larger scale than the primary wave. Secondary waves with wavelengths comparable to the primary modulation scale were investigated by Vadas and Fritts (2002), Vadas et al. (2003), and Becker and Vadas (2018). The authors propose a generating mechanism based on body forces produced by the dissipating primary wave. In contrast to this model, our secondary wave is generated by direct wave–mean-flow interaction that compares to Wilhelm et al. (2018).

Second, this novel picture of saturated mountain waves may explain the bias between the onset of small-scale instability and the actually observed mean-flow acceleration (Achatz 2007). We give an extension to the established picture where waves become unstable at the breaking height and deposit their momentum and energy onto the mean flow at this level. As it turns out, only in combination with modulational instability does an initially saturated nonlinear wave induce a mean flow. This modulational instability has its own transient velocity and growth rate, which we can compute explicitly. Then the mean-flow acceleration depends on these two quantities.

In our derivations we assumed piecewise analytic solutions being matched. This assumption can be weakened to almost arbitrary background, fulfilling hydrostatics and the ideal gas law, as well as almost any mean-flow horizontal wind. However, these functions of height must be restricted to converge to constant values at the infinities. The resulting spectrum describing the temporal evolution of the perturbation would be much more complicated but still analytically assessable by Fredholm operator theory (Schlutow et al. 2019). In conclusion, our results are valid in a much more realistic atmosphere.

APPENDIX

Well-Posedness: Are the Instability Growth Rates Finite?

In this appendix we demonstrate that the linear differential operator of the EVP, (29), is well-posed. This property is essential for existence, uniqueness and continuous dependency of the solution for the perturbation on the initial data. In section 6 we showed that the spectrum of the operator extents into the right half of the complex plane rendering the

saturated wave unconditionally unstable. But is the spectrum bounded from above on the real axis? Let us denote the discriminant of the square root appearing in $\lambda_{1,2}$ of (39) by $D_Y(\mu)$. The real part being the instability growth rate may be expressed by

$$\text{Re}(\lambda_{1,2}(\mu)) = \pm K_x \sqrt{A} |D_Y(\mu)|^{1/2} \cos\left\{\frac{1}{2}\arg(D_Y(\mu))\right\}, \tag{A1}$$

where

$$|D_Y(\mu)|^{1/2} = |\mu| \sqrt[4]{(2\Lambda K_z - H^{-1}\Omega'')^2 \mu^{-2} + \Omega''^2}, \tag{A2}$$

$$\arg(D_Y(\mu)) = \begin{cases} \arctan\left(\dfrac{H^{-1}\Omega'' - 2\Lambda K_z}{\Omega'' \mu}\right), & \Omega'' < 0 \\[4mm] \arctan\left(\dfrac{H^{-1}\Omega'' - 2\Lambda K_z}{\Omega'' \mu}\right) \pm \pi, & \Omega'' > 0. \end{cases} \tag{A3}$$

The function $\text{Re}(\lambda_{1,2}(\mu))$ has no poles. Thus, we are only concerned with its behavior at the infinities. We have

$$|D_Y(\mu)|^{1/2} = O(|\mu|) \quad \text{as} \quad \mu \to \pm\infty, \tag{A4}$$

$$\cos\left\{\frac{1}{2}\arg(D_Y(\mu))\right\} = \begin{cases} O(1), & \Omega'' < 0 \\[4mm] \pm\dfrac{1}{2}\dfrac{H^{-1}\Omega'' + 2\Lambda K_z}{\Omega''}\mu^{-1} + O(|\mu|^{-3}), & \Omega'' > 0 \end{cases} \quad \text{as} \quad \mu \to \pm\infty. \tag{A5}$$

Therefore,

$$\text{Re}(\lambda_{1,2}(\mu)) \to \begin{cases} \pm\infty, & \Omega'' < 0 \\[4mm] \pm\dfrac{1}{2}\dfrac{2\Lambda K_z - H^{-1}\Omega''}{\sqrt{\Omega''}}, & \Omega'' > 0 \end{cases} \quad \text{as} \quad \mu \to \pm\infty. \tag{A6}$$

In conclusion the operator is well-posed (Kapitula and Promislow 2013) if $\Omega'' > 0$ and ill-posed otherwise.

REFERENCES

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

——, R. Klein, and F. Senf, 2010: Gravity waves, scale asymptotics and the pseudo-incompressible equations. *J. Fluid Mech.*, **663**, 120–147, https://doi.org/10.1017/S0022112010003411.

Becker, E., 2012: Dynamical control of the middle atmosphere. *Space Sci. Rev.*, **168**, 283–314, https://doi.org/10.1007/s11214-011-9841-5.

——, and S. L. Vadas, 2018: Secondary gravity waves in the winter mesosphere: Results from a high-resolution global circulation model. *J. Geophys. Res. Atmos.*, **123**, 2605–2627, https://doi.org/10.1002/2017JD027460.

Dunkerton, T. J., 1989: Theory of internal gravity wave saturation. *Pure Appl. Geophys.*, **130**, 373–397, https://doi.org/10.1007/BF00874465.

Durran, D. R., 1989: Improving the anelastic approximation. *J. Atmos. Sci.*, **46**, 1453–1461, https://doi.org/10.1175/1520-0469(1989)046<1453:ITAA>2.0.CO;2.

Fritts, D. C., 1984: Gravity wave saturation in the middle atmosphere: A review of theory and observations. *Rev. Geophys. Space Phys.*, **22**, 275–308, https://doi.org/10.1029/RG022i003p00275.

——, and Coauthors, 2016: 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. *Bull. Amer. Meteor. Soc.*, **97**, 425–453, https://doi.org/10.1175/BAMS-D-14-00269.1.

——, and Coauthors, 2018: Large-amplitude mountain waves in the mesosphere accompanying weak cross-mountain flow during DEEPWAVE research flight RF22. *J. Geophys. Res. Atmos.*, **123**, 9992–10 022, https://doi.org/10.1029/2017JD028250.

Grimshaw, R., 1972: Nonlinear internal gravity waves in a slowly varying medium. *J. Fluid Mech.*, **54**, 193–207, https://doi.org/10.1017/S0022112072000631.

——, 1974: Internal gravity waves in a slowly varying, dissipative medium. *Geophys. Fluid Dyn.*, **6**, 131–148, https://doi.org/10.1080/03091927409365792.

——, 1977: The modulation of an internal gravity-wave packet, and the resonance with the mean motion. *Stud. Appl. Math.*, **56**, 241–266, https://doi.org/10.1002/sapm1977563241.

Hodges, R. R., 1969: Eddy diffusion coefficients due to instabilities in internal gravity waves. *J. Geophys. Res.*, **74**, 4087–4090, https://doi.org/10.1029/JA074i016p04087.

Jeffreys, H., 1922: On the dynamics of wind. *Quart. J. Roy. Meteor. Soc.*, **48**, 29–48, https://doi.org/10.1002/qj.49704820105.

Kapitula, T., and K. Promislow, 2013: *Spectral and Dynamical Stability of Nonlinear Waves*, Applied Mathematical Sciences, Vol. 185, Springer, 361 pp., https://doi.org/10.1007/978-1-4614-6995-7.

Klostermeyer, J., 1991: Two- and three-dimensional parametric instabilities in finite-amplitude internal gravity waves.

Geophys. Astrophys. Fluid Dyn., **61**, 1–25, https://doi.org/10.1080/03091929108229035.

Lindzen, R. S., 1970: Internal gravity waves in atmospheres with realistic dissipation and temperature part I. Mathematical development and propagation of waves into the thermosphere. *Geophys. Fluid Dyn.*, **1**, 303–355, https://doi.org/10.1080/03091927009365777.

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

Midgley, J. E., and H. B. Liemohn, 1966: Gravity waves in a realistic atmosphere. *J. Geophys. Res.*, **71**, 3729–3748, https://doi.org/10.1029/JZ071i015p03729.

Nicolet, M., 1954: Dynamic effects of the high atmosphere. *The Earth as a Planet*, G. P. Kuiper, Ed., University of Chicago Press, 644–712.

Pitteway, M. L. V., and C. O. Hines, 1963: The viscous damping of atmospheric gravity waves. *Can. J. Phys.*, **41**, 1935–1948, https://doi.org/10.1139/p63-194.

Sandstede, B., and A. Scheel, 2000: Absolute and convective instabilities of waves on unbounded and large bounded domains. *Physica D*, **145**, 233–277, https://doi.org/10.1016/S0167-2789(00)00114-7.

Schlutow, M., R. Klein, and U. Achatz, 2017: Finite-amplitude gravity waves in the atmosphere: Travelling wave solutions. *J. Fluid Mech.*, **826**, 1034–1065, https://doi.org/10.1017/jfm.2017.459.

——, E. Wahlén, and P. Birken, 2019: Spectral stability of nonlinear gravity waves in the atmosphere. *Math. Climate Wea. Forecasting*, **5**, 12–33, https://doi.org/10.1515/mcwf-2019-0002.

Vadas, S. L., and D. C. Fritts, 2002: The importance of spatial variability in the generation of secondary gravity waves from local body forces. *Geophys. Res. Lett.*, **29**, 1984, https://doi.org/10.1029/2002GL015574.

——, ——, and M. J. Alexander, 2003: Mechanism for the generation of secondary waves in wave breaking regions. *J. Atmos. Sci.*, **60**, 194–214, https://doi.org/10.1175/1520-0469(2003)060<0194:MFTGOS>2.0.CO;2.

Wilhelm, J., T. R. Akylas, G. Bölöni, J. Wei, B. Ribstein, R. Klein, and U. Achatz, 2018: Interactions between mesoscale and submesoscale gravity waves and their efficient representation in mesoscale-resolving models. *J. Atmos. Sci.*, **75**, 2257–2280, https://doi.org/10.1175/JAS-D-17-0289.1.

Zhou, Q., 1995: A simplified approach to internal gravity wave damping by viscous dissipation. *J. Atmos. Terr. Phys.*, **57**, 1009–1014, https://doi.org/10.1016/0021-9169(94)00107-Y.