Shear and Static Instability of Inertia-Gravity Wave Packets: Short-Term Modal and Nonmodal Growth

Ulrich Achatz and Gerhard Schmitz

Leibniz-Institut für Atmosphärenphysik an der Universität Rostock

Kühlungsborn, Germany

submitted to *J. Atmos. Sci.*, April 6th 2004

revised March 1st 2005

Ulrich Achatz
Leibniz-Institut für Atmosphärenphysik an der Universität Rostock
Schlossstr. 6
18225 Kühlungsborn
Germany
Tel.: +49-38293-68340
Fax: +49-38293-6850
e-mail: achatz@iap-kborn.de

Abstract

The problem of nonmodal instabilities of inertia-gravity waves (IGW) in the middle atmosphere is addressed, within the framework of a Boussinesq model with realistic molecular viscosity and thermal diffusion, by singular-vector analyzes of horizontally homogeneous vertical profiles of wind and buoyancy obtained from IGW packets at their statically least stable or most unstable horizontal location. Nonmodal growth is always found to be significantly stronger than that of normal modes, most notably at wave amplitudes below the static instability limit, where normal-mode instability is very weak, whereas the energy gain between optimal perturbation and singular vector after one Brunt-Vaisala period can be as large as two orders of magnitude. Among a multitude of rapidly growing singular vectors for this optimization time, small-scale (wavelengths of a few 100m) perturbations propagating in the horizontal parallel to the IGW are most prominent. These parallel optimal perturbations are amplified by a roll mechanism, while transverse perturbations (with horizontal scales of a few km) are to a large part subject to an Orr mechanism, both controlled by the transverse wind-shear in the IGW at its statically least stable altitude, but further enhanced by reduced static stability. The elliptic polarization of the IGW leaves its traces in an additional impact of the roll mechanism via the parallel-wind shear on the leading transverse optimal perturbation.

1 Introduction

The importance of gravity waves for the dynamics of the middle atmosphere has been recognized for quite a while (Hines, 1960). As they propagate upwards from mostly tropospheric sources their amplitude grows due to the vertical ambient density gradient. Unless they are absorbed at some critical layer (Bretherton, 1966; Booker and Bretherton, 1967) or get deflected at wind or stratification gradients (Chimonas and Hines, 1986; Fritts and Yuan, 1989a; Bühler and McIntyre, 1999) they eventually become subject to wave breaking. The associated momentum and energy deposition is essential for an understanding of the middle atmosphere general circulation (Houghton, 1978; Lindzen, 1981; Holton 1982, 1983; Garcia and Solomon, 1985). Another effect might be the generation of turbulence by gravity wave breaking, thus leading to frictional heating rates of relevance for the mesospheric energy balance. Measurements of turbulent dissipation rates have been published by Lübken (1997) while Becker and Schmitz (2002) showed these to be in agreement with respective values diagnosed from a gravity wave parameterization scheme in a simple general circulation model (GCM). A possible example of a simultaneous observation of an unstable gravity wave and related turbulence was published by Müllemann et al. (2003).

The involved processes still present many open questions. Due to the small scales of the relevant waves they cannot be simulated explicitly in a GCM but must be incorporated via a parameterization. Among the best-known schemes for this task are those by Lindzen (1981), Medvedev and Klaassen (1995), Hines (1997a,b), Alexander and Dunkerton (1999), and Warner and McIntyre (2001). Already the very diversity of approaches argues for an unsatisfactory state of knowledge so that one is confronted with an uncom-

fortably large number of free options in this critical model component. Moreover, recent direct numerical simulations (DNS) of breaking gravity waves have revealed disagreements with several schemes (Afanasyev and Peltier, 2001; Fritts et al., 2003).

Studies of unstable gravity waves by DNS seem to be a key tool for a better understanding of both wave breaking and turbulence excitation. Among these, however, especially three-dimensional studies (Andreassen et al., 1994; Fritts et al., 1994; Isler et al., 1994; Winters and D'Asaro, 1994; Lelong and Dunkerton, 1998a,b; Werne and Fritts, 1999; Afanasyev and Peltier, 2001; Fritts at al., 2003) are still, and will continue to be, a challenge to available computer resources, so that good a priori knowledge about the scales and structures developing in an unstable gravity wave seems highly desirable. Studies of the incipient instability phase via normal mode analysis have proven useful for this purpose. Using the Floquet method, such analyzes have been done for monochromatic gravity waves at higher frequencies by Mied (1976), Klostermeyer (1982, 1983, 1991), Lombard and Riley (1996), and Sonmor and Klaassen (1997), revealing instabilities at all wave amplitudes. To a large part these are three-dimensional, i.e. oblique horizontal propagation with respect to the unstable wave is a characteristic of many growing modes. The local instability of inertia-gravity wave packets, with frequencies near the lower inertial bound, was focus in the studies by Fritts and Yuan (1989b), Yuan and Fritts (1989), Dunkerton (1997) and Kwasniok and Schmitz (2003). It was shown that for sufficiently low-frequent waves dynamic instability due to velocity shear in the wave is more important than static instability. For these waves no strong preference for any horizontal propagation direction of the unstable modes was found. For monochromatic inertia-gravity waves Yau et al. (2004) have recently done a normal-mode analysis. There it is shown that at low intrinsic

frequencies the most rapidly growing modes propagate parallel to the IGW. The authors attribute this to the elliptic polarization in the horizontal wind of the wave.

In several cases normal modes have been used as initial perturbations to a gravity wave in nonlinear simulations. This requires either the normal modes to have much shorter time scales than the gravity wave, or the latter to be a monochromatic wave propagating at a constant phase speed, admitting the use of Floquet analysis. Then a normal-mode approach is especially justified when the initial perturbation, provided by an ambient fluctuation field, can be assumed to be very small so that the leading normal mode has enough time to develop. At moderate perturbation levels another effect might however come into play. It is known from several fields that over finite times the transient growth of so-called singular vectors can be much more rapid than that of normal modes (Farrell, 1988a,b; Butler and Farrell, 1992; Trefethen et al., 1993). This can happen when the normal modes are not orthogonal, which is most typically the case for dynamical systems, so that small initial perturbations are possible where large contributions from different normal modes cancel each other. Provided the time developments of the various normal modes differ enough, those cancellations are removed subsequently so that the strength of the perturbation can rise considerably. This has two consequences. First, it can be that for dynamical states with a growing normal mode the transient growth of a sufficiently strong initial perturbation is so rapid that it leads the dynamical fields directly into the nonlinear phase without the normal mode having had time to develop. Then the normal mode would have been an ill choice for characterizing the incipient instability phase. Secondly, it can even happen that a state diagnosed from the normal-mode analysis to be stable experiences transient growth vigorous enough that it can actually lift weak initial

perturbations to a sufficiently high level for further nonlinear development leading to an irreversible modification of the dynamical fields. In the gravity-wave context this could imply turbulence excitation and wave breaking under conditions where this would not have been expected from a normal-mode analysis. Last, but not least, general gravity wave packets with a time dependence more complex than constant phase propagation in general do no allow a normal-mode analysis at all, so that to these wave packets more general concepts such as singular-vector analyzes must be applied in order to get to grasps with the incipient instability phase.

One can ask oneself what the most rapidly growing optimal perturbations of a gravity wave are and what type of wave decay they can lead to. This is the very question we want to address here. As a starting point we focus on inertia-gravity waves. Both for better comparability of our results to those of previous studies (Fritts and Yuan, 1989b; Yuan and Fritts, 1989; Dunkerton, 1997; Kwasniok and Schmitz, 2003) and for simplicity we additionally first concentrate on the analysis of one-dimensional vertical profiles considered most representative for the IGW packet in question. In this context we also focus on short development times (about one Brunt-Vaisala period). The optimally growing structures for the complete wave packet, their time development, and the dependence on optimization time are studied in a companion paper (Achatz and Schmitz, 2005).

For this purpose we describe the analyzed IGW packets in section 2. Section 3 introduces the linear models used in the stability analysis and gives a short discussion of optimal growth in general. For the most simple approximation of an IGW packet, i.e. by a stratified constant-shear layer, helpful analytic results on optimal growth are derived in section 4, followed by an analysis of a vertical profile of the IGW at the statically

least stable horizontal location in section 5. Our results are summarized and discussed in section 6.

2 Inertia-gravity wave packet

Our study uses the three-dimensional viscous Boussinesq equations with thermal diffusivity on an f-plane

$$\nabla \cdot \widehat{\mathbf{v}} = 0 \tag{1}$$

$$\frac{\partial \widehat{\mathbf{v}}}{\partial t} + (\widehat{\mathbf{v}} \cdot \nabla)\widehat{\mathbf{v}} + f\mathbf{e}_z \times \widehat{\mathbf{v}} + \nabla \widehat{p} - \mathbf{e}_z \widehat{b} = \nu \nabla^2 \widehat{\mathbf{v}} \tag{2}$$

$$\frac{\partial \widehat{b}}{\partial t} + (\widehat{\mathbf{v}} \cdot \nabla)\widehat{b} + N^2 \widehat{w} = \mu \nabla^2 \widehat{b} \quad . \tag{3}$$

Here $\widehat{\mathbf{v}} = (\widehat{u}, \widehat{v}, \widehat{w})$ denotes the three-dimensional velocity field. The buoyancy $\widehat{b} = g(\theta - \overline{\theta}(z))/\theta_0$ is a measure of the deviation of the potential temperature θ from a merely vertically dependent reference profile $\overline{\theta}(z)$, normalized by a characteristic value θ_0. g is the vertical gravitational acceleration. The squared background Brunt-Vaisala frequency $N^2 = (g/\theta_0)d\overline{\theta}/dz$ is assumed to be a constant. $\widehat{p}$ is the pressure field, normalized by a constant reference density, f the Coriolis parameter, and $\mathbf{e}_z$ the vertical unit vector. The Boussinesq equations can be expected to give a reasonably good approximation of the full gravity wave dynamics as long as the focus is on processes with vertical scales of the order or less than the atmospheric scale height. This is the case throughout this study. For viscosity and thermal diffusivity we take the typical mesospheric values $\nu = \mu = 1\mathrm{m}^2/\mathrm{s}$. The f-plane is located at 45° latitude. Our Brunt-Vaisala frequency is $N = 10^{-2}\mathrm{s}^{-1}$. For later reference we note that under usual boundary conditions the Boussinesq equations

conserve total energy

$$E = \frac{1}{2} \oint_V dV(|\widehat{\mathbf{v}}|^2 + \frac{\widehat{b}^2}{N^2}) \qquad . \tag{4}$$

In the inviscid-nondiffusive limit both (1) – (3) and their linearization about the reference atmosphere at rest yield as solutions the plane waves

$$\begin{pmatrix} \widehat{\mathbf{v}} \\ \widehat{b} \end{pmatrix} = a\Re \left[\begin{pmatrix} \mathbf{v}_\pm \\ b_\pm \end{pmatrix} e^{i(\mathbf{K}\cdot\mathbf{x} - \Omega_\pm t)} \right] \tag{5}$$

where

$$\Omega_\pm = \pm\sqrt{\frac{N^2(K^2+L^2)+f^2M^2}{K^2+L^2+M^2}} \tag{6}$$

$$(u_\pm, v_\pm, w_\pm, b_\pm) = \left(-\frac{K\Omega_\pm + iLf}{K^2+L^2}, \frac{iKf - L\Omega_\pm}{K^2+L^2}, \frac{\Omega_\pm}{M}, -i\frac{N^2}{M} \right) \qquad . \tag{7}$$

Here $\mathbf{K} = (K, L, M)$ is the three-dimensional wave vector. The normalization has been chosen so that the wave is statically unstable at $a > 1$.

Having in mind an analysis of the stability of an IGW propagating from the lower atmosphere into the mesosphere, thereby gaining in amplitude due to the vertical ambient density gradient, we are less interested in plane wave trains of infinite vertical extent. Among the gravity wave solutions above the one with the minus sign in front of the square root has an upwardly directed group velocity if $M > 0$. From this we obtain the state of the IGW packet at $t = 0$, i.e.

$$\begin{pmatrix} \mathbf{V_0} \\ B_0 \end{pmatrix} (\mathbf{x}) = a\Re \left[\begin{pmatrix} \mathbf{v}_- \\ b_- \end{pmatrix} e^{i\mathbf{K}\cdot\mathbf{x} - z^2/\sigma_z^2} \right] \qquad . \tag{8}$$

Without loss of generality we assume $L = 0$ so that the wave propagates in the $x - z$ plane. As vertical extent σ_z of the packet we use one vertical wavelength $\Lambda_z = 2\pi/M$. For the wavelengths we have chosen typical observed values: The horizontal wavelength

is $\Lambda_x = 2\pi/K = 500$ km, whereas the vertical wavelength was varied between 3, 6, and 9 km in order to get some information about the effect on our results of varying the ratio $R = f/|\Omega_-|$. For the examined wavelengths it is $R = 0.86, 0.65$, and 0.50, respectively. The maximum of the wave packet is in the center of the vertical domain. There $N^2 + \partial b/\partial z$ minimizes at $x = \Lambda_x/2$, i.e. the statically most unstable region is right in the center of the $x - z$ plane (Fig. 1). This is where also the total Richardson number

$$\mathrm{Ri}_t = \frac{N^2 + \partial\widehat{b}/\partial z}{(\partial\widehat{u}/\partial z)^2 + (\partial\widehat{v}/\partial z)^2} \tag{9}$$

minimizes if $a > 2(1 - R^2)/(2 - R^2)$.

3 Description of an incipient instability

3.1 Linear model

In the following we will be interested in the development of perturbations to the IGW packet over time scales much shorter than its period. At sufficiently small initial strengths these are appropriately described in a model obtained by linearizing the Boussinesq equations about the initial wave packet. The problem is treated on two levels of simplification.

3.1.1 Approximation of the IGW by a one-dimensional profile

As will be seen below and in the companion paper much of the dynamics of an incipient instability of an IGW is controlled by the conditions near its statically least stable location. In this part of our study it is therefore approximated by its initial state at the location $x = x_0 = \Lambda_x/2$ (= 250km) where the strongest static instability is to be expected, subject to the condition following (9). Vertical motion in the IGW is duely neglected (see also Fig.

1), i.e. we assume $W_0 = 0$. We set $\widehat{\mathbf{v}} = \mathbf{V}_0 + \mathbf{v}'$, $\widehat{b} = B_0 + b'$ and insert this into (1) – (3). Neglecting all terms quadratic in the perturbations we obtain a set of equations which, due to a corresponding symmetry in the basic wave packet, does not couple different wavenumbers in the horizontally parallel (x) and transverse (y) direction. Thus one can as well set $(\mathbf{v}', b', p') = (\mathbf{v}, b, p)(x, z)\exp[i(kx+ly)]$ for arbitrary wavenumbers k and l. As others have done before (e.g. Dunkerton, 1997) we express the horizontal wavenumbers in terms of parallel wavenumber $k_\parallel$ and azimuth angle α as $(k, l) = k_\parallel(\cos\alpha, \sin\alpha)$. For the further discussion it is also useful to introduce a modified coordinate system, obtained by rotating the horizontal axes so that in the new coordinates $(x_\parallel, y_\perp)$ the $x_\parallel$-axis points into the direction of the horizontal wave vector of the perturbation, i.e. we take

$$x_\parallel = x\cos\alpha + y\sin\alpha \tag{10}$$

$$y_\perp = -x\sin\alpha + y\cos\alpha \quad . \tag{11}$$

With the corresponding horizontal velocity components in perturbation and IGW being denoted by $(U_\parallel, V_\perp)$ and $(u_\parallel, v_\perp)$ the model equations then become

$$ik_\parallel u_\parallel + \frac{\partial w}{\partial z} = 0 \tag{12}$$

$$\frac{Du_\parallel}{Dt} + w\frac{dU_\parallel}{dz} + ik_\parallel p - fv_\perp = \nu\nabla_1^2 u_\parallel \tag{13}$$

$$\frac{Dv_\perp}{Dt} + w\frac{dV_\perp}{dz} + fu_\parallel = \nu\nabla_1^2 v_\perp \tag{14}$$

$$\frac{Dw}{Dt} + \frac{\partial p}{\partial z} - b = \nu\nabla_1^2 w \tag{15}$$

$$\frac{Db}{Dt} + N_{tot}^2 w = \mu\nabla_1^2 b \quad , \tag{16}$$

where $D/Dt = (\partial/\partial t + ik_\parallel U_\parallel)$, $\nabla_1 = (ik_\parallel, 0, \partial/\partial z)$, and $N_{tot}^2 = N^2 + dB_0/dz$. Provided the coriolis effect is weak, which is the case here, one sees from (12)–(16) that $v_\perp$ is coupled to the other model variables only passively.

The equations have been discretized using second-order finite central differences on a staggered grid with $u_\parallel, v_\perp, b$ on full and w on intermediate half levels. Boundary conditions are periodic. Unless specified differently the model domain extension was $L_z = 3\Lambda_z$ and the number of grid points (512 or 1024) was chosen large enough to ensure numerical convergence of our results. The pressure Poisson equation, obtained by taking the divergence of (13)–(15) and using (12), is solved by standard techniques. The time integration is done by two initial fourth-order Runge-Kutta time steps and a third-order Adams-Bashforth scheme thereafter. These time-stepping schemes, as well as the solution of the Poisson equation are discussed in Durran (1999).

As preparation for the constant-shear-layer approximation below we also note that there is, due to (12), a streamfunction ψ so that $(u_\parallel, w) = (-\partial/\partial z, ik_\parallel)\psi$. With this the model equations can be written

$$\frac{D\zeta}{Dt} + ik_\parallel b = \nu\nabla_1^2\zeta - ik_\parallel \frac{d^2U_\parallel}{dz^2}\psi + f\left(\frac{\partial v_\perp}{\partial z} + ik_\parallel u_\parallel\right)\psi \tag{17}$$

$$\frac{Dv_\perp}{Dt} + ik_\parallel \frac{dV_\perp}{dz}\psi = \nu\nabla_1^2 v_\perp \tag{18}$$

$$\frac{Db}{Dt} + ik_\parallel N_{tot}^2\psi = \mu\nabla_1^2 b \quad , \tag{19}$$

where $\zeta = (k_\parallel^2 - \partial^2/\partial z^2)\psi$ is the vorticity component in $y_\perp$-direction.

3.1.2 Approximation by a stratified constant-shear layer

Going one step further one can focus even more on the conditions near the statically least stable location by approximating (see also Fig. 1) the IGW fields by their tangents there, i.e. by assuming

$$U_0 = u_0 = a\Omega_-/K \tag{20}$$

$$V_0 = \beta z \qquad (\beta = afM/K) \tag{21}$$

$$N_{tot}^2 \;=\; N^2(1-a) \qquad , \tag{22}$$

so that $U_\parallel = u_c + \beta_s z$ and $dV_\perp/dz = \beta_c$, where we write $\beta_{s,c} = \beta(\sin\alpha, \cos\alpha)$ and $u_c = u_o \cos\alpha$. Optimal growth in a corresponding nonrotating stratified constant-shear layer with $N_{tot}^2 = N^2$, i.e. without locally reduced static stability, has been studied by Farrell and Ioannou (1993a) and Bakas et al. (2001) whose results we here use and expand on. We neglect in (17)–(19), but not in the basic state, the coriolis effect and set

$$(\psi, \zeta, v_\perp, b)\,(z,t) = \int_{-\infty}^{\infty} dm\,(\psi_m, \zeta_m, v_m, b_m)\,(t) \exp\left[i\left(m_t z - k_\parallel u_c t\right) - D\right] \tag{23}$$

with $m_t = m - k_\parallel \beta_s t$ a time dependent vertical wavenumber, and $D = \nu \int_0^t d\tau K_t^2(\tau)$ the viscous-diffusive damping increment (assuming $\mu = \nu$), while $K_t^2 = k_\parallel^2 + m_t^2$. Thus one obtains the independent three-component systems

$$\frac{d\zeta_m}{dt} \;=\; -ik_\parallel b_m \tag{24}$$

$$\frac{db_m}{dt} \;=\; -ik_\parallel N_{tot}^2 \psi_m \tag{25}$$

$$\frac{dv_m}{dt} \;=\; -ik_\parallel \beta_c \psi_m \tag{26}$$

with $\zeta_m = K_t^2 \psi_m$. Note that these equations conserve the quantity $v_m/\beta_c - b_m/N_{tot}^2$.

3.2 Normal modes and singular vectors

In using the linear model for a stability analysis of the IGW packet we want to identify structures most quickly arising from an arbitrary initial perturbation. Growth of these would indicate an instability possibly leading to turbulent decay. Traditionally normal mode analysis has been used for this purpose, whereas we want to examine singular vectors. Both shall be defined here and compared to each other. This shall be done at

the utmost possible brevity allowing readers to follow the presentation. For details they are referred to Farrell and Ioannou (1996a,b) and Schmid and Henningson (2001).

For compactness of notation we introduce the complex state vector $\mathbf{x}$ of the model which comprises the N values of all variables, i.e. in general $\mathbf{v}'$ and b' at all model grid points, but here all $\mathbf{v}$ and b for the one-dimensional model and all (ζ_m, b_m, v_m) for the stratified constant-shear layer. The model equations can then be written shortly as

$$\frac{d\mathbf{x}}{dt} = \mathcal{A}(t)\mathbf{x} \tag{27}$$

with $\mathcal{A}$ being a possibly time dependent linear operator mapping the state on its tendencies.

Standard normal modes are actually only well-defined for a time-independent background. In this case they are the eigenvectors $\mathbf{n}_\nu$ of the model operator so that

$$\mathcal{A}\mathbf{n}_\nu = -i(\omega_\nu + i\gamma_\nu)\mathbf{n}_\nu \qquad . \tag{28}$$

ω_ν and γ_ν are the frequency and growth rate of the ν-th normal mode. Usually the normal modes form a complete set so that any initial perturbation can be written as

$$\mathbf{x}(0) = \sum_{\nu=1}^{N} a_\nu \mathbf{n}_\nu \tag{29}$$

and, due to (27) and (28),

$$\mathbf{x}(t) = \sum_{\nu=1}^{N} a_\nu e^{-i\omega_\nu t + \gamma_\nu t}\mathbf{n}_\nu \qquad . \tag{30}$$

In the linear approximation, if there is a leading normal mode, i.e. with larger growth rate than all others, it will grow more rapidly (for positive growth rate) or decay slower (for negative growth rate) than the others, and *after a sufficiently long transition time* it will arise from any initial perturbation as the dominant feature. A leading normal mode

with positive growth rate is thus taken as sign of instability of the analyzed background state. It is further often used to characterize the incipient instability phase before the nonlinear development is entered.

The possibility of rapid *transient* growth is examined in a singular-vector analysis. Different from the normal-mode analysis one needs a definition of the strength of a perturbation, i.e. one defines a norm

$$\|\mathbf{x}\|^2 = \overline{\mathbf{x}}^t \mathcal{M} \mathbf{x} \tag{31}$$

where the metric $\mathcal{M}$ is positive definite and hermitian. The upper index t denotes transposition, and the overbar taking the complex conjugate. Different choices are possible for the metric. We here use as norm the discretized version of total energy

$$\|\mathbf{x}\|^2 = \frac{1}{2} \oint_V dV (|\mathbf{v}'|^2 + \frac{|b'|^2}{N^2}) \qquad . \tag{32}$$

We now ask ourselves what initial perturbation $\mathbf{x}(0)$ maximizes for some given finite time τ the ratio $\|\mathbf{x}(\tau)\|^2 / \|\mathbf{x}(0)\|^2$. The answer to this question makes use of the propagator matrix $\mathbf{\Phi}$ which maps the initial perturbation to its state at $t = \tau$ via

$$\mathbf{x}(\tau) = \mathbf{\Phi}(\tau)\mathbf{x}(0) \qquad . \tag{33}$$

By variational analysis one can show that the desired perturbation initializing strongest growth is the leading eigenvector $\mathbf{p}_\nu$ satisfying

$$\mathcal{M}^{-1} \overline{\mathbf{\Phi}}^t(\tau) \mathcal{M} \mathbf{\Phi}(\tau) \mathbf{p}_\nu = \sigma_\nu^2 \mathbf{p}_\nu \tag{34}$$

with the largest possible eigenvalue σ_ν^2, which in turn is the desired (squared) optimal growth factor $\|\mathbf{x}(\tau)\|^2 / \|\mathbf{x}(0)\|^2$. Since $\mathcal{M}$ is symmetric and positive definite there is a Cholesky factorization $\mathcal{M} = \overline{\mathcal{N}}^t \mathcal{N}$ where $\mathcal{N}$ is upper triangular. In our case it is diagonal,

as is easily checked. Inserting the factorization into (34) and defining $\mathbf{q}_\nu = \mathcal{N}\mathbf{p}_\nu$ the eigenvalue problem can be rewritten

$$\overline{\mathcal{L}}^t \mathcal{L}\mathbf{q}_\nu = \sigma_\nu^2 \mathbf{q}_\nu \tag{35}$$

with

$$\mathcal{L} = \mathcal{N}\mathbf{\Phi}(\tau)\mathcal{N}^{-1} \qquad . \tag{36}$$

From (35) one sees that all eigenvalues are positive. Furthermore, the eigenvectors $\mathbf{q}_\nu$ are orthogonal with respect to the euclidian metric, and henceforth also the optimal perturbations $\mathbf{p}_\nu$ with respect to $\mathcal{M}$.

4 Normal modes and singular vectors in the stratified constant-shear layer

Expressing kinetic energy in terms of ψ and $v_\perp$ one finds that for the constant-shear-layer case total energy is given, up to an irrelevant constant factor, by $E = \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dl \int_{-\infty}^{\infty} dm E_{klm}$ with

$$E_{klm} = \frac{e^{-2D}}{2}\left(K_t^2 \left|\psi_m\right|^2 + \left|v_m\right|^2 + \frac{\left|b_m\right|^2}{N^2}\right) \qquad , \tag{37}$$

so that different k and l, i.e. $k_\parallel$ and α, and different initial vertical wavenumbers m are completely decoupled in the energy norm. Thus each of the subsystems (24)–(26) must be considered separately in a singular-vector analysis. This can be done numerically for arbitrary azimuth angles, but the special advantage of the constant-shear-layer approximation is that it admits several closed analytical solutions helpful in getting an oversight of the dependence of optimal growth on the various parameters.

It is instructive to first neglect in (24)–(26) the time dependence of K_t and m_t, yielding a linear system with constant coefficients and normal-mode solutions

$$(\psi_m, b_m, v_m)_\pm = (1, \pm N_{tot} K_t, \pm \beta_c K_t / N_{tot}) / (K_t \sqrt{\epsilon_+}) \tag{38}$$

$$(\psi_m, b_m, v_m)_v = (0, 0, 1) \tag{39}$$

with eigenfrequencies $\omega_\pm = \pm N_{tot} k_\parallel / K_t$ and $\omega_v = 0$, where $\epsilon_+ = 1 + \beta_c^2 / |N_{tot}|^2 + |N_{tot}|^2 / N^2$. The normalization has been chosen so that all three modes have the same initial unit energy $E_{klm} = 1$. For $a > 1$ the convention is $N_{tot} = i |N_{tot}|$ so that in this case the first normal mode is growing exponentially, unless damped at small wavelengths by viscosity and diffusion, while at $a < 1$ all three modes are damped. One sees that two possibilities for near-collinearity between the two first or all normal modes arise. If $\beta_c = 0$ and $|N_{tot}|$ is small the tow first modes are in structure very close to $(1, 0, 0)$, whereas for small $|N_{tot}|$ and large $\beta_c / |N_{tot}|$ they are approximately collinear with the third mode. This collinearity is the very reason for the strong optimal growth discussed below.

Getting back to the real case with time dependent K_t and m_t it is useful to exploit that the subsystem (24)–(25) always yields

$$\frac{d^2 \zeta_m}{dt^2} = -\frac{N_{tot}^2 k_\parallel^2}{K_t^2} \zeta_m \tag{40}$$

admitting for large enough $|\omega_\pm|$ the WKB solution (Mathews and Walker, 1970; Farrell and Ioannou, 1993a) $\zeta_m(t) \propto \sqrt{K_t} \exp(\pm i\phi)$ with $\phi(t) = \int_0^t dt' N_{tot} k_\parallel / K(t')$. Thus we use for the solution of the general initial-value problem the ansatz, validated a posteriori below,

$$\begin{pmatrix} \psi_m \\ b_m \\ v_m \end{pmatrix} (t) = a_+ e^{-i\phi} \begin{pmatrix} \psi \\ b \\ v \end{pmatrix}_+ + a_- e^{i\phi} \begin{pmatrix} \psi \\ b \\ v \end{pmatrix}_- + a_v \begin{pmatrix} \psi \\ b \\ v \end{pmatrix}_v \tag{41}$$

$$(\psi, b, v)_{\pm} = \frac{g^{3/4}}{K_0\sqrt{\epsilon_+}}\left(1, \pm N_{tot}K_t, \pm\beta_c K_t/N_{tot}\right) \quad (42)$$

$$(\psi, b, v)_{v} = (0, 0, 1) \quad (43)$$

with $K_0 = K_t(t=0)$ and $g = K_0^2/K_t^2$. Note that the WKB solution uses as basis time dependent generalizations of the normal modes discussed above, which are exact for the case $\beta_s = 0$, i.e. $\alpha = 0°$. As shown in the Appendix A the approximation yields analytic approximate optimal-growth factors which in the two limit cases of parallel or transverse propagation of the perturbation take an especially simple form.

4.1 Parallel singular vectors

In the case $\alpha = 0$ one has $\beta_s = 0$, i.e. m_t and K_t are constant, and thus the WKB solution is exact. The available basic shear is transverse to the wave vector, which corresponds to the situation where the so-called *roll mechanism* (Moffat, 1967; Ellingsen and Palm, 1975; Landahl, 1980) transfers kinetic energy from the basic state to a perturbation in which an initial vertical velocity induces a transverse wind component in the presence of corresponding basic shear. The hence resulting growth factor for the leading optimal perturbation, given in the appendix, can be approximated for $\beta_c^2/\left|N_{tot}\right|^2 = 1/\left|\mathrm{Ri}\right| \gg 1 \gg \left|N_{tot}\right|^2/N^2$ by

$$\sigma_1^2 \approx e^{-2D}\frac{4}{|\mathrm{Ri}|}\frac{N^2}{|N|_{tot}^2}\begin{cases} \sin^4\frac{\phi}{2} & a < 1 \\ \sinh^4\frac{|\phi|}{2} & a > 1 \end{cases} . \quad (44)$$

Via an increase of the collinearity of the involved normal modes both shear and reduced static stability act to enforce the transient growth, which is reduced by viscous-diffusive damping. Most notably, it can be strong even for $a < 1$, when all three normal modes decay.

Given a vertical wavenumber m and an optimization time τ one can ask oneself about the dependence of σ_1^2 on $k_\parallel$, especially at which $k_\parallel$ optimal growth maximizes. For $a > 1$ one finds that due to $|\phi| = |N_{tot}|\, k_\parallel \tau / \sqrt{k_\parallel^2 + m^2}$ the leading-term behavior is $\sigma_1^2 = (4/|\mathrm{Ri}|)(N^2/|N_{tot}|^2) \exp\left\{2\tau\left[|N_{tot}|\, k_\parallel / \sqrt{k_\parallel^2 + m^2} - \nu\left(k_\parallel^2 + m^2\right)\right]\right\}$ so that growth maximizes, just as for the unstable normal mode, at $k_\parallel = k_m$ where

$$0 = \frac{|N_{tot}|\, m^2 k_m}{(m^2 + k_m^2)^{3/2}} - 2\nu k_m^2 \qquad , \tag{45}$$

i.e. $k_m \approx m\,(|N_{tot}|/2\nu m^2)^{1/4}$ for $\nu m^2 \ll |N_{tot}|/2$, and $k_m \approx m\,|N_{tot}|/2\nu m^2$ for $\nu m^2 \gg |N_{tot}|/2$. In this case the wavenumber location of optimal growth is independent of the optimization time.

The situation differs for $a < 1$. For $N_{tot}\tau/2\pi > 1$ there are local maxima of σ_1^2 where $\phi \approx (2n+1)\pi$ for some positive integer $n < (N_{tot}\tau/\pi - 1)/2$, i.e. at $k_\parallel \approx k_n$ where

$$k_n = \frac{m}{\sqrt{\left[\frac{N_{tot}\tau}{(2n+1)\pi}\right]^2 - 1}} \qquad . \tag{46}$$

The number of extrema thus rises with increasing τ where always the one at longest horizontal wavelength, i.e. $k_\parallel = k_0$, is least damped by diffusion and viscosity, so that there $\sigma_1^2 \approx (4/|\mathrm{Ri}|)(N^2/N_{tot}^2)$ for not too strong damping. For $N_{tot}\tau/2\pi < 1$ only one maximum remains which is near

$$k_0 = m\left[\frac{N_{tot}}{2\nu m^2 \tan\left(N_{tot}\tau/2\right)}\right]^{1/4} \qquad , \tag{47}$$

as long as in this approximation $k_0 \gg m$, reducing for $N_{tot}\tau/2 \ll 1$ to $k_0 = m/(\tau\nu m^2)^{1/4}$, i.e. only slowly increasing with decreasing ν and τ. In contrast to the case $a > 1$ we thus here have a dependence of the growth factor on optimization time. For $\lambda_z = 0.1\Lambda_z = 600$ m and four representative combinations of a and τ the leading growth factors are shown in figure 2, giving also a good confirmation of approximation (44).

The structure of the optimal perturbations is quite interesting. In the appendix it is also shown that under the same approximations as used for (44) one has for the strongest-growing structure

$$a_{\pm} \approx \mp \frac{1}{2} \frac{N_{tot}}{|N_{tot}|} \frac{\beta_c}{|\beta_c|} a_v \qquad . \tag{48}$$

Inserting this into (41)–(43) yields for $\beta_c / |N_{tot}| \gg 1$ and $|N_{tot}| / N \ll 1$

$$\psi_m = \frac{a_v}{2K_0} \frac{|N_{tot}|}{|\beta_c|} \frac{N_{tot}}{|N_{tot}|} \frac{\beta_c}{|\beta_c|} \left(e^{i\phi} - e^{-i\phi} \right) \tag{49}$$

$$b_m = -\frac{a_v}{2} N \frac{|N_{tot}|}{|\beta_c|} \frac{N_{tot}}{N} \frac{N_{tot}}{|N_{tot}|} \frac{\beta_c}{|\beta_c|} \left(e^{i\phi} + e^{-i\phi} \right) \tag{50}$$

$$v_m = -\frac{a_v}{2} \left(e^{i\phi} + e^{-i\phi} \right) + a_v \qquad , \tag{51}$$

so that ones sees that it is composed in such a way from the three normal modes that both ψ_m and v_m are initially approximately cancelled, whence due to subsequent removal of this cancellation especially the latter quickly rises in amplitude. This case of transient growth is thus a clear example of the interference effect behind optimal growth in general, here acting to produce rapid growth in the transverse flow field.

4.2 Transverse singular vectors

For $\alpha = 90°$ one has $\beta_c = 0$ and thus the two first WKB normal modes have no parallel-velocity component v_m. The leading optimal perturbation is composed from these so that optimal growth is restricted to ψ_m and b. Now the available shear is parallel to the horizontal wavenumber vector of the perturbation, so that the so-called *Orr mechanism* (Orr, 1907) can extract energy from the transverse flow-field component V_0 of the shear layer. In this case the perturbation leans initially against the shear and intensifies as it is sheared over, thus producing strong vertical velocities. K_t is time dependent so that

the variation of g has an important impact. As also shown in Appendix A one has at $|N_{tot}|^2/N^2 \ll 1$

$$\sigma_1^2 \approx e^{-2D} g^{1/2} \frac{N^2}{|N|_{tot}^2} \begin{cases} \sin^2\phi & a < 1 \\ \sinh^2|\phi| & a > 1 \end{cases} \quad . \tag{52}$$

Also for this case both shear (via g) and reduced static stability (via an enhanced collinearity of the involved normal modes) act to enforce the transient growth, which is reduced by viscous-diffusive damping. Once again $a = 1$ is no real instability threshold.

At fixed initial vertical wavenumber m the squared growth factor σ_1^2 peaks at the maximum of g where K_t^2 minimizes. This is at $k_\parallel = m/\beta\tau$. There

$$g = 1 + \beta^2\tau^2 \tag{53}$$

$$D = \nu m^2 \tau \left(\frac{1}{\beta^2\tau^2} + \frac{1}{3} \right) \tag{54}$$

$$\phi = -\frac{N_{tot}}{\beta} \ln\left(\sqrt{1+\beta^2\tau^2} - \beta\tau \right) \quad , \tag{55}$$

so that the growth factor is of the order $N^2/|N_{tot}|^2$ for small τ, and diverges for large τ as long as viscous-diffusive damping is unimportant. This divergence is $\propto \tau$ for $a < 1$ and $\propto \tau^{\sqrt{4/|\mathrm{Ri}|+1}}$ for $a > 1$, i.e. in contrast to the parallel case algebraic and not exponential. For $\lambda_\parallel = 0.1\Lambda_z = 600$ m and four representative combinations of a and τ the leading growth factors are shown in figure 3, giving also a good confirmation of the WKB approach and approximation (52). For the exact values the propagator matrix has been determined from numerical integration of (24)–(26), followed by a numerical solution of the eigenvalue problem (34).

The structure of the leading optimal perturbation reveals a similar interference effect as in the case above. To lowest approximation in $|N_{tot}|^2/N^2 \ll 1$ one finds (Appendix A)

$a_+ \approx -a_-$ so that

$$\psi_m = -a_+ \frac{g^{3/4}}{K_0} \left(e^{i\phi} - e^{-i\phi} \right) \tag{56}$$

$$b_m = a_+ N_{tot} g^{1/4} \left(e^{i\phi} + e^{-i\phi} \right) \quad , \tag{57}$$

i.e. there is initially near-cancellation in ψ_m whence in the integration it rapidly rises in amplitude due to removal of the destructive interference, so that this structure is dominated by growth in kinetic energy in ψ_m. For both cases, i.e. parallel and transverse propagation, one finds the buoyancy growth to be relatively weak.

4.3 Dependence on azimuth angle

In order to give an impression of the general dependence of optimal growth in the shear-layer approximation we show in figure 4 for the same four cases as in Figs. 2 and 3 the exact growth factors for the azimuth angles $\alpha = 0°, 30°, 60°$, and $90°$. In a comparison between the growth of parallel and transverse perturbations one finds the former to dominate for $a > 1$. This is a result of the exponential dependence on optimization time in the parallel case, while optimal growth of transverse perturbations only rises algebraically with τ. At $a < 1$ parallel optimal growth dominates for small τ, due to the additional factor $4/|\mathrm{Ri}|$ in σ_1^2, while transverse optimal growth takes the lead for larger τ, since the growth factors of parallel perturbations cannot be larger than $(4/|\mathrm{Ri}|)\,(N^2/N_{tot}^2)$, while $\sigma_1^2 \propto \tau$ in the other case. Interestingly for $a < 1$ the growth factor gets largest for intermediate azimuth angles. This is in good agreement with the synergism between Orr mechanism and roll mechanism observed by Farrell and Ioannou (1993b) in unstratified shear layers. Only at larger a and τ the growth of parallel perturbations dominates as a result of the strong exponential dependence of the corresponding growth factor on τ. The dependence

of the optimal wavenumber, at which optimal growth maximizes, on optimization time is for oblique SV, due to the factor g, quite similar to the behavior in the transverse case, i.e. we find optimal growth near $k_\parallel = m/\beta_s\tau$, so that the optimal wavelength rises in proportion with the optimization time. One also recognizes the dependence of the optimal wavelength of parallel optimal growth on τ as predicted for $a < 1$ by (46) ($n = 0$ for $\tau = 1$ h) and (47) (for $\tau = 10$ min). The only case where the optimal wavenumber stays independent of τ is for parallel SV at $a > 1$.

5 Normal modes and singular vectors for the 1D profile

In the approximation of an IGW by a one-dimensional vertical profile one obtains the independent subsystems (12)–(16), one for each combination of k and l, i.e. $k_\parallel$ and α. Moreover total energy can be written, up to an irrelevant constant factor, as $E = \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dl E_{kl}$ with $E_{kl} = |\mathbf{v}|^2/2 + |b|^2/2N^2$, so that the energy norm also does not couple different horizontal wave vectors. Thus both the normal modes and the singular vectors must be determined separately for each combination of $k_\parallel$ and α. This has been done using sparse-matrix techniques, as described in Appendix B.

5.1 Short-term normal-mode growth vs singular-vector growth

First we have analyzed a statically unstable IGW packet ($a = 1.5$) with vertical wavelength $\Lambda_z = 6$ km. The development time considered is $\tau = 10$ min, i.e. approximately one Brunt-Vaisala period. The growth factors ($\exp \gamma_\nu \tau$ for the normal modes and σ_ν for

the singular vectors) of the leading patterns (with largest growth rates, i.e. for $\nu = 1$) obtained in both analyzes are shown for different azimuth angles and parallel wavelengths in figure 5. Because of the Coriolis effect there is a weak asymmetry in the growth factors between (for any β) $\alpha = 90° \pm \beta$. It is however so small that we only show results for $0° \leq \alpha \leq 90°$. In comparing the normal mode growth curves to corresponding results by Dunkerton (1997) and Kwasniok and Schmitz (2003) minor differences are to be expected, since in those studies the impact of the Coriolis effect was taken into account as far as the shape of the basic wave is concerned, but not in the linear model itself. However, we actually find quite good agreement for this value of a. Thus, also two main features in the growth curves obtained there are reproduced. First, at parallel wavelengths of the order of the vertical wavelength of the basic wave we find a local maximum which is especially pronounced for transverse propagation ($\alpha = 90°$). Secondly, towards shorter wavelengths the growth factors for non-transverse propagation rise again to an even higher level. Most pronounced here is the growth factor for parallel propagation ($\alpha = 0°$), maximizing at $\lambda_{\parallel} \approx 400$ m to a value near 30. At even shorter scales viscosity takes over and leads to weaker growth rates, and finally decay. This is in close correspondence to the strong growth of small-scale parallel normal modes in the constant-shear layer discussed above.

The two peaks, as well as the general finding that instability exists at all azimuths, can also be interpreted via the local Richardson number $\mathrm{Ri} = N_{tot}^2/(\partial U_{\parallel}/\partial z)^2$ (Ri_t as defined in (9) is the minimum of Ri over all azimuth angles). Motivation for the use of the local Richardson number is that for a purely z-dependent background without vertical wind the well-known Taylor-Goldstein equation can be derived from the linear equations (12) – (16) without rotation, which has been the basis of the analysis by Fritts and Yuan (1989b),

Yuan and Fritts (1989), Dunkerton (1997), and Kwasniok and Schmitz (2003). This equation had previously been analyzed by Howard (1961) and Miles (1961) who showed that no normal-mode growth can exist if Ri $> 1/4$ everywhere. As a necessary condition for normal-mode instability the local Richardson number should therefore be less than $1/4$ anywhere in the domain. Neglecting viscosity and diffusivity, setting $\mathbf{K}\cdot\mathbf{x}-\omega t = Kx_0 + Mz = \pi + \phi$ with local phase $\phi = Mz$, and applying (5) – (7) one finds

$$\mathrm{Ri} \;=\; \frac{(1-R^2)(1-a\cos\phi)}{a^2(R\sin\alpha\cos\phi+\cos\alpha\sin\phi)^2}\,\frac{1}{1-\Omega_-^2/N^2} \qquad . \tag{58}$$

For $(R,a) = (0.65, 1.5)$ we have plotted $\mathrm{Ri}(\alpha,\phi)$ in figure 6 (assuming $\Omega_-^2 \ll N^2$). The instability at all azimuths is consistent with the observation that for all α there is a wave phase where Ri $< 1/4$. Furthermore we find at $\phi = 0$ negative singularities at $\alpha = 0, \pm\pi$, and saddle points at $\alpha = \pm\pi/2$, indicating the most pronounced instabilities to occur at parallel propagation. Lelong and Dunkerton (1998a) show a similar plot for $a < 1$ where the minimum of Ri lies at $\alpha = \pm\pi/2$. It appears that the growth rate peaks we find for transverse propagation are related to these minima which here at $a > 1$ become masked by the negative singularities at $\alpha = 0, \pm\pi$.

Now turning to the singular vectors (right panel of figure 5) we observe conspicuous similarities to the normal modes. Again we find the most intense growth at short scales and parallel propagation with a secondary peak at longer scales and transverse propagation. The peaks are less pronounced, and the azimuth dependence is also somewhat weaker. Furthermore, the most rapidly growing transverse singular vectors have smaller scales. Whereas the most rapid transverse normal mode growth is at $\lambda_\parallel \approx 8$ km, the respective nonmodal growth maximizes at 5 km. The most important difference, however, is that nonmodal growth within τ =10 min is more rapid, by about a factor 3, than

normal-mode growth. At other wave amplitudes we get the same result. Figure 7 shows the normal-mode and singular-vector growth curves for parallel and transverse propagation at $a = 0.9$ and 0.5. At slightly subcritical amplitudes ($a = 0.9$) optimal growth is still quite strong (nearly a factor 10 for parallel propagation) while the normal modes grow only by an insignificant rate. The largest growth factor we find is $\exp(\gamma_1 \tau) = 1.006$ at $\left(\alpha, \lambda_\parallel\right) = (90°, 8.9\text{km})$. The slightly more intense normal-mode instabilities found by Yau et al. (2004) for subcritical monochromatic IGW seem to be due to their neglect of viscosity and diffusion, and their use of larger values for a and R than ours. Even at rather small amplitudes ($a = 0.5$) optimal growth still exists. This, however, should perhaps not be overrated, since growth by a factor 2 might usually not be sufficient for really destabilizing the basic wave packet. These results verify the prediction from the constant-shear-layer approximation that at short optimization times strong transient growth of nonmodal perturbations can exist, dominated by near-parallel propagation. This persists for $a < 1$, while there modal growth is basically suppressed. We also reproduce the scales of parallel and transverse optimal perturbations, the former being considerably shorter than the latter. A major difference is that no synergism between roll and Orr mechanism is seen in the sense that it is always the parallel perturbations which grow most rapidly, while the shear layer would predict most rapid growth for slightly oblique propagation directions, as also visible in figure 4.

The impact of the vertical wavelength of the basic wave on optimal growth can be seen in figure 8. There we show the leading growth factors (for parallel and transverse singular vectors) for different basic wave amplitudes and vertical wavelengths. As R is increased or Λ_z decreased ($R = 0.86, 0.65$, and 0.50, for $\Lambda_z = 3$, 6, and 9 km, respectively) the

singular-vector growth gets stronger. It appears that the tendency for subcritical growth (at $a < 1$) occurs mainly in inertia-gravity waves with smaller vertical wavelengths (i.e. large R). This is in close analogy to the predictions from normal-mode theory (Fritts and Rastogi, 1985; Dunkerton, 1997). Once again we find this also to be a prediction from the constant-shear-layer approximation where it is found that parallel optimal growth obeys the proportionality $\sigma_1 \propto \sqrt{4/\left|\mathrm{Ri}\right|} = 2\beta/\left|N_{tot}\right|$ while for transverse perturbations one has $\sigma_1 \propto \left(1+\beta^2\tau^2\right)^{1/4}$. Thus smaller vertical wavelengths with larger shear in the transverse wind $\beta = afM/K$ lead to stronger nonmodal instabilities. Furthermore, also here we find for all examined waves stronger short-term growth for parallel singular vectors. We also find a tendency for longer instability wavelengths $\lambda_{\|}$ as Λ_z is increased. At least partially this might be explained as an effect of viscosity and diffusion since the shear-layer approximation would predict, due to the corresponding damping, strongest growth at smallest m which however is limited below by the vertical wavenumber of the wave packet. Smaller M thus allow smaller m not in contradiction to the basic assumption that the perturbative scales are smaller than those of the IGW, while the optimal $k_{\|}$ scales roughly with m.

5.2 Nonmodal structure and growth mechanisms

The similarity between the growth-factor curves obtained in the constant-shear-layer approximation and for the more general 1D profile suggest the singular vectors to be well enough localized in the region of least static stability for the former approximation to hold. Indeed, as can be seen in figure 9, this is the case. There we show the distribution of energy density E_{kl}, within the innermost wavelength ($\Lambda_z = 6$ km) of the IGW packet

with $a = 1.5$, for leading normal mode and singular vector ($\tau = 10$ min) at four representative azimuth angles. For $\alpha \leq 30°$ both types of perturbations are highly concentrated in the innermost region of the wave packet. For larger azimuth angles the normal modes are rather smoothly spread over the wave packet, while the optimal perturbations are still well localized, although for the leading transverse singular vector this localization is smeared out in the course of its further time development. At other values of a and Λ_z the localization turns out to be comparable (not shown).

Next comes the question how well the optimal perturbations can be approximated by a monochromatic wave in the vertical, as is done in the shear-layer approximation. Figure 10 gives a representative impression. There we show for $\Lambda_z = 6$ km and $a = 0.9$ and 1.1 the real part of the structure of the leading parallel and transverse optimal perturbations. One does see a wavy structure with wavelengths very roughly of the order of 0.1 Λ_z, as was assumed in the shear-layer calculations above. Beyond this, however, we also note a dependence of the dominant vertical wavelength on azimuth angle which cannot be predicted by the shear-layer considerations. Furthermore the wave trains are all but constant in local wavelength. Nonetheless, even these rough agreements seem to be satisfactory enough for the shear-layer picture to provide valuable information.

Finally we look at the relative growth of the various model variables. The shear-layer approximation predicts for the parallel singular vectors dominant growth in $v_\perp$ while that of b should be negligible. For the leading transverse perturbation one obtains growth mainly in ψ with at most weak amplification in b, and no growth in $v_\perp$. For a comparison we have calculated how much in an optimal perturbation the contribution of each of these variables to E_{kl} increases within the optimization time. For the same perturbations as

shown in Fig. 10 these amplifications (e.g. $|v_\perp(\tau)|^2 / \left\langle |v_\perp(0)|^2 \right\rangle$ for $v_\perp$ where angle brackets denote a vertical average) are shown in figure 11. For the parallel perturbation one finds the picture from the shear-layer approximation basically verified. More interesting here is the leading transverse perturbation. While indeed the amplification of ψ dominates in the innermost region, we also find a strong contribution from growth in $v_\perp$ away from there. This indicates that the gradient in U_0, neglected in the constant-shear-layer assumption, but non-vanishing for $z \neq 0$, plays an important role so that via the roll mechanism growth of $v_\perp$ becomes possible. The elliptic polarization of the IGW is thus leaving traces.

6 Summary and discussion

The significance of rapid transient growth of nonmodal structures in the incipient instability of gravity waves in the middle atmosphere had not been given a systematic focus so far. Our study represents a first step in this direction. Here we have looked at inertia-gravity waves (IGW) with rather long horizontal scales (500 km) and low intrinsic frequencies ($0.5 \leq R = f/|\omega_{intr}| \leq 0.86$). Their linear stability has been studied in a Boussinesq model with realistic viscosity and diffusion. Motivated by the scenario of an IGW propagating upwards from some source below and gaining in amplitude due to the decreasing ambient density we have not looked at a single monochromatic wave, but rather an upwardly propagating wave packet of limited vertical extent. For different internal frequencies, i.e. vertical wavelengths, and different wave amplitudes we have picked from the wave packet the vertical profile of wind and buoyancy initially exhibiting least static stability. This has then been used as a horizontally uniform and time-independent

basic state for the stability analysis. Both standard normal modes have been calculated and singular vectors. The latter are those among all possible perturbations maximizing the growth of total energy in the disturbance over some finite time, which we have mostly taken to be 10 min, i.e. roughly one Brunt-Vaisala period.

In all cases examined we find the leading singular vector to amplify significantly more rapidly than the leading normal mode. Most importantly, at wave amplitudes slightly below the static instability limit ($a = 0.9$), where normal modes exhibit only insignificant growth over 10 min, the energy amplification in the leading singular vector can be as large as two orders of magnitude. The problem being separable with respect to the horizontal wavenumber of the perturbation, we find among a rather broad spectrum (both with respect to the wavelength and the propagation direction of the perturbation) two main types of nonmodal instabilities. One of these has rather long horizontal wavelength of a few kilometers and propagates in the horizontal perpendicular to the propagation direction of the basic wave (transverse singular vector). The other type is a singular vector with rather short horizontal wavelength (a few 100 m), propagating in the horizontal parallel to the IGW (parallel singular vector). Among the two we always find the latter to amplify more rapidly than the former.

Interestingly, much of the diagnosed transient growth can be understood in a most simple approximation of the examined vertical IGW profile where only the reduced Brunt-Vaisala frequency N^2_{tot} and the transverse-wind shear β at the altitude of least static stability is taken into account. This admits analytic treatment of the two limit cases of parallel and transverse propagation. Basic mechanisms discussed by Farrell and Ioannou (1993a,b) and Bakas et al. (2001) for stratified constant-shear layers turn out to be

responsible for the optimal growth, here however with a significant enhancement by the locally reduced static stability. For parallel perturbations this is the roll mechanism which leads to preferential amplification of the transverse-wind component in the perturbation. Transverse singular vectors are driven by the Orr mechanism where the transverse-vertical circulation is extracting energy from the basic wave. In both cases optimal growth can be explained explicitly via a simple interference effect between contributing normal modes. The scale-difference between the two types of optimal perturbations is reproduced, i.e. at a given vertical wavenumber the growth of parallel optimal perturbations maximizes at shorter wavelengths than that of transverse perturbations. Also the relative amplitudes of the respective growth factors are predicted well by the constant-shear layer approximation. One finds that the energy amplification of parallel perturbations in the statically stable case ($a < 1$) has an upper limit $(4/|\mathrm{Ri}|)\,(N^2/N_{tot}^2)$, with $\mathrm{Ri} = N_{tot}^2/\beta^2$, while that of transverse perturbations increases algebraically with optimization time, as long as viscosity and diffusion are of weaker importance. This indicates that at intermediate optimization times transverse singular vectors might be more important. For very long optimization times we expect the weaker viscous-diffusive damping of parallel perturbations (with constant vertical wavenumber in contrast to the oblique cases) to eventually take the lead, just as shown for the case $N^2 = N_{tot}^2$ by Bakas et al. (2001). One major difference between the results from the shear-layer approximation and the more general calculations is that the shear-layer synergism between roll and Orr mechanism is not found in the more general case, i.e. instead of obliquely propagating perturbations it is always the parallel structures which grow most strongly. Another interesting difference is that the transverse SV seems in reality to be also influenced, via a respective roll

mechanism, by the parallel-wind shear at some distance from the altitude of least static stability (where it vanishes). As a consequence the parallel wind amplifies significantly in this structure. This is in interesting difference to the dynamics of normal modes where the growth rate is only influenced by the gradients in the IGW wind component parallel to the propagation direction of the perturbation, as follows from the Taylor-Goldstein equation (e.g. Dunkerton, 1997; Kwasniok and Schmitz, 2003). It should however be noted that also the normal modes determined here always have a strong amplitude in $v_\perp$ (not shown). What this implies for the feedback of the growing perturbation on the basic wave must remain unclear at this stage. It can, however, be anticipated that it is to a large part the corresponding IGW structure in this component ($v_\perp$) which will be modified. This would influence both the upward momentum transport in the wave, but also the instability dynamics of perturbations propagating at oblique angles.

The parameter dependence of our results seems worth some discussion. The impact of vertical wavelength of the basic wave on the instability is as expected from previous normal mode analyzes and the shear-layer approximation: As it decreases, and thus β increases, the instability gets more vigorous, and both parallel and transverse leading singular vector move to smaller scales. Since the value of the Brunt-Vaisala frequency used here is actually a bit low for the middle atmosphere, although it has e.g. also been employed by Fritts et al. (2003), we have also done calculations with $N = 2 \cdot 10^{-2}\mathrm{s}^{-1}$ at 70° latitude. The results there have been very similar to the ones reported here, with the major difference that the overall instability time scale was shorter by a corresponding factor 2. For better readability we had chosen not to nondimensionalize our equations. By doing so, however, using Brunt-Vaisala period and vertical basic-wave wavelength as

time and length scales, one can see that at zero viscosity and thermal diffusivity the only external controlling parameter for a wave with given scales is f/N. Thus, at fixed f/N and spatial IGW scales, varying N simply implies a proportional variation of the inverse time scale of the problem, thus explaining the above mentioned factor 2. The remaining influence of the different ratio f/N found in our second set of calculations is due to a corresponding variation of R (or β), i.e. an IGW with vertical wavelengths 3, 6, or 9 km has $R = 0.75, 0.5$, or 0.35, respectively. Accordingly the identified growth factors (within 5 min now) are slightly weaker, while the results described here would strictly rather apply to a basic wave with longer horizontal wavelength, i.e. approximately 650 km.

Another important factor might be viscosity and thermal diffusion, the more so as in the atmosphere both rise exponentially with growing altitude due to the decrease in ambient density. We have therefore also done calculations with a value $\nu = \mu = 5\mathrm{m}^2/\mathrm{s}$. Due to its comparatively large spatial scales we find no impact of increased viscosity and diffusion on the leading transverse singular vector. Because of its smaller spatial scales the leading parallel singular vector, however, is affected. Its growth factors decrease and it moves to larger spatial scales. For the latter effect the shear layer-approximation predicts $\lambda_{\|} \propto \nu^{1/4}$, which is roughly reproduced by our calculations. Due to this the comparative importance of parallel singular vector and transverse singular vector seems to be altitude-dependent, as well as the spatial scales of the leading parallel singular vector.

It is certainly too early for a final statement on the relevance of transient nonmodal growth for the problem of IGW breaking in the middle atmosphere, although our results seem to support it. Further investigations are necessary to shed more light on the issue. So it will have to be answered what remains of the identified patterns as the instability

problem with the complete time-dependent IGW packet is treated, and it might also be interesting to examine the development of optimally growing structures on longer time scales than considered here. Due to limited space also the energetics of the identified perturbations has not been given sufficient attention in this study. Their nonlinear development will have to be studied in order to assess their role in wave saturation or damping, and in turbulence generation. Another very important question relates to the ambient perturbations present in the middle atmosphere which must serve to initialize the modelled transient behavior. It will have to be shown that their amplitude can at times be large enough to favor nonmodal over modal growth (which will be comparatively more important the stronger the NM growth rates are), and the role of their spectrum in selecting certain initial perturbations over others will have to be examined. Some of these questions are addressed in a companion paper (Achatz and Schmitz, 2005), but much remains for the future.

Acknowledgements

The authors thank two anonymous reviewers for numerous helpful suggestions for an improvement of the manuscript.

Appendix A Optimal growth in the stratified-shear-layer approximation

Using E_{klm} as norm one finds from application of (37) and (41)–(43) that

$$E_{klm} = \overline{\mathbf{a}}^t \overline{\mathbf{\Phi}}^t \mathcal{M} \mathbf{\Phi} \mathbf{a} \qquad , \tag{A1}$$

where

$$\mathbf{a} = (a_+, a_-, a_v)^t \tag{A2}$$

$$\overline{\mathbf{\Phi}}^t \mathcal{M} \mathbf{\Phi} = \frac{e^{-2D}}{2} \begin{pmatrix} g^{1/2} e^{-i\left(\phi - \overline{\phi}\right)} & g^{1/2} e^{i\left(\overline{\phi} + \phi\right)} \frac{\epsilon_-}{\epsilon_+} & g^{1/4} e^{i\overline{\phi}} \frac{\beta_c}{\sqrt{\epsilon_+} \overline{N}_{tot}} \\ g^{1/2} e^{-i\left(\phi + \overline{\phi}\right)} \frac{\epsilon_-}{\epsilon_+} & g^{1/2} e^{i\left(\phi - \overline{\phi}\right)} & -g^{1/4} e^{-i\overline{\phi}} \frac{\beta_c}{\sqrt{\epsilon_+} \overline{N}_{tot}} \\ g^{1/4} e^{-i\phi} \frac{\beta_c}{\sqrt{\epsilon_+} N_{tot}} & -g^{1/4} e^{i\phi} \frac{\beta_c}{\sqrt{\epsilon_+} N_{tot}} & 1 \end{pmatrix} \tag{A3}$$

with $\epsilon_- = 1 - \beta_c^2 / \left|N_{tot}\right|^2 - \left|N_{tot}\right|^2 / N^2$. $\mathcal{M}\left(t = 0\right)$ results from (A3) by setting $g = 1$, $D = 0$, and $\phi = 0$. The optimal perturbations are given by the three eigenvectors $\mathbf{a}_\nu$ satisfying

$$\left(\overline{\mathbf{\Phi}}^t \mathcal{M} \mathbf{\Phi}\right)(\tau)\, \mathbf{a}_\nu = \sigma_\nu^2 \mathcal{M}\left(0\right) \mathbf{a}_\nu \tag{A4}$$

with a corresponding eigenvalue equation $\det\left[\left(\overline{\mathbf{\Phi}}^t \mathcal{M} \mathbf{\Phi}\right)(\tau) - \sigma_\nu^2 \mathcal{M}\left(0\right)\right] = 0$. This yields a third-order polynomial for σ_ν^2 which can be solved analytically, albeit by rather complicated expressions for the eigenvalues. More interesting, however, is that in the two cases $\beta_s = 0$ (parallel perturbation with $g = 1$ and $\phi = N_{tot} k \tau / \sqrt{k^2 + m^2}$) and $\beta_c = 0$ (transverse perturbation) there is the exact solution $\sigma_2^2 = \exp\left(-2D\right)$. With this knowledge the characteristic polynomial can be reduced to second order with the comparatively simple solutions

$$\sigma_{1,3}^2 = g^{1/2} e^{-2D} \left(\frac{\delta}{2} \pm \sqrt{\frac{\delta^2}{4} - 1}\right) \qquad , \tag{A5}$$

where for $\beta_s = 0$

$$\begin{aligned}\delta &= 2 - \frac{N^2}{4N_{tot}^2}\left(1 - \frac{N_{tot}^2}{N^2}\right)^2 \left(e^{i\phi} - e^{-i\phi}\right)^2 \\ &+ \frac{\beta_c^2 N^2}{4N_{tot}^4}\left[\left(e^{i\phi/2} - e^{-i\phi/2}\right)^4 - \frac{N_{tot}^2}{N^2}\left(e^{i\phi} - e^{-i\phi}\right)^2\right] \quad ,\end{aligned} \tag{A6}$$

and for $\beta_c = 0$

$$\begin{aligned}\delta &= \frac{N^2}{4\left|N_{tot}\right|^2}\left(1 + \frac{\left|N_{tot}\right|^2}{N^2}\right)^2 \\ &\times \left[e^{-i\left(\phi - \overline{\phi}\right)} + e^{i\left(\phi - \overline{\phi}\right)} - \left(\frac{1 - \left|N_{tot}\right|^2/N^2}{1 + \left|N_{tot}\right|^2/N^2}\right)^2 \left(e^{-i\left(\phi + \overline{\phi}\right)} + e^{i\left(\phi + \overline{\phi}\right)}\right)\right] \quad .\end{aligned} \tag{A7}$$

In the limit $\beta_c^2/\left|N_{tot}\right|^2 = 1/\left|\mathrm{Ri}\right| \gg 1 \gg \left|N_{tot}\right|^2/N^2$ one obtains from (A6)

$$\delta \approx \beta_c^2 N^2/\left(4N_{tot}^4\right)\left(e^{i\phi/2} - e^{-i\phi/2}\right)^4 \gg 1 \quad , \tag{A8}$$

whence results (44). Similarly one obtains (52) from (A7) in the limit $\left|N_{tot}\right|^2/N^2 \ll 1$.

The structure of the optimal perturbations is derived by reinserting the growth factor into (A4). One finds at $\beta_s = 0$ for the i-th perturbation

$$\begin{aligned}\frac{a_\pm}{a_v} &= \frac{\beta_c}{\sqrt{\epsilon_+}\overline{N}_{tot}} \frac{1}{\left[e^{-i\left(\phi - \overline{\phi}\right)} - s_i^2\right]\left[e^{i\left(\phi - \overline{\phi}\right)} - s_i^2\right] - \frac{\epsilon_-^2}{\epsilon_+^2}\left|e^{-i\left(\phi + \overline{\phi}\right)} - s_i^2\right|^2} \\ &\times \begin{cases} -\left(e^{i\overline{\phi}} - s_i^2\right)\left[e^{i\left(\phi - \overline{\phi}\right)} - s_i^2\right] - \frac{\epsilon_-}{\epsilon_+}\left(e^{-i\overline{\phi}} - s_i^2\right)\left[e^{i\left(\phi + \overline{\phi}\right)} - s_i^2\right] \\ \left(e^{-i\overline{\phi}} - s_i^2\right)\left[e^{-i\left(\phi - \overline{\phi}\right)} - s_i^2\right] + \frac{\epsilon_-}{\epsilon_+}\left(e^{i\overline{\phi}} - s_i^2\right)\left[e^{-i\left(\phi + \overline{\phi}\right)} - s_i^2\right] \end{cases} \quad ,\end{aligned} \tag{A9}$$

where $s_i^2 = e^{2D}\sigma_i^2$. In the limit of $\beta_c^2/\left|N_{tot}\right|^2 = 1/\left|\mathrm{Ri}\right| \gg 1 \gg \left|N_{tot}\right|^2/N^2$, and hence large s_1^2, this reduces for $i = 1$ to (48). At $\beta_c = 0$ the second optimal perturbation is just $(a_+, a_-, a_v) = (0, 0, 1)$ while for $i = 1, 3$ one obtains $a_v = 0$ and

$$\frac{a_+}{a_-} = \frac{\epsilon_-}{\epsilon_+}\frac{s_i^2 - e^{i\left(\overline{\phi} + \phi\right)}}{e^{-i\left(\phi - \overline{\phi}\right)} - s_i^2} \quad , \tag{A10}$$

where $s_i^2 = e^{2D}g^{-1/2}\sigma_i^2$. For $i = 1$ this yields in the limit $\left|N_{tot}\right|^2/N^2 \ll 1$ the approximate relation $a_+ \approx -a_-$, as long as s_1^2 is large.

Appendix B Technical aspects of the algorithms for normal modes and singular vectors

For not too high-dimensional problems one can calculate normal modes or optimal perturbations directly via eigenvalue analysis of $\mathcal{A}$ or singular value analysis of $\mathcal{L}$, respectively. For the latter one needs the propagator matrix which can either be determined by initializing the tangent linear model with all possible unit vectors or (in the case of a time-independent background) by first doing an eigenvalue analysis of $\mathcal{A}$, transforming to the eigenvector basis, integrating the tangent linear model analytically via (30), and transforming back to the original grid-point representation. In the case of large problems where storage of $\mathcal{A}$ and $\mathbf{\Phi}$ is too demanding for the available computer memory iterative techniques are required. Since high-dimensionality is a problem in the singular-vector calculations for the full two-dimensional IGW packet reported in a companion paper we have used such techniques in our analysis. Provided only a few leading patterns are desired this can also speed up the calculations.

For the calculation of leading normal modes of a time-independent background we use the fact that these are also eigenvectors of $\mathbf{\Phi}(\tau)$ with eigenvalue $\exp(-i\omega_\nu\tau + \gamma_\nu\tau)$. With the help of the software package ARPACK (Lehoucq et al., 1998) a desired number of leading eigenvectors (with largest eigenvalue modulus $\exp(\gamma_\nu\tau)$) have been determined via an implicitly restarted Arnoldi Method. In this approach $\mathbf{\Phi}(\tau)\mathbf{q}$ is required for initial states $\mathbf{q}$ determined in the algorithm at each iteration. Integrating from $t = 0$ to $t = \tau$ this mapping is obtained from the linear model. The iterations stop when the required normal modes, eigenfrequencies, and growth rates have been calculated at a chosen accuracy.

The same software package is also employed in the determination of a required number of leading optimal perturbations. For this we solve the eigenvalue problem (35). As in the normal-mode analysis, ARPACK requires the user at each iteration to provide $\overline{\mathcal{L}}^t\mathcal{L}\mathbf{q}$ for initial states $\mathbf{q}$ determined by the algorithm. After having obtained $\mathbf{r} = \overline{\mathcal{N}}^t\mathcal{N}\mathbf{\Phi}(\tau)\mathcal{N}^{-1}\mathbf{q}$ with the help of the linear model this necessitates also taking the product $\overline{\mathbf{\Phi}}^t(\tau)\mathbf{r}$. For this we used the adjoint Boussinesq model extracted from the linear model code with the help of the adjoint model compiler TAMC (Giering and Kaminski, 1998). Also in this analysis the iterations stop when a required accuracy for the desired number of leading optimal perturbations is reached.

There is one additional aspect to the problem here which needs special treatment. The Boussinesq equations require the flow field to be non-divergent, i.e. purely rotational. The pressure is determined so that the divergence of a flow field is conserved. In the search of the optimal rotational perturbation one must therefore modify the propagator matrix by first decomposing the initial flow field $\mathbf{v}$ into its rotational and its divergent part. For this we have determined a velocity potential χ such that $\nabla_1^2\chi = \nabla_1 \cdot \mathbf{v}$. The projection is done by mapping $\mathbf{v} \mapsto \mathbf{v} - \nabla_1\chi$. Since the kinetic energy of the flow field can be split up into its contributions from the rotational part and the divergent part, this leads to a reduction of the kinetic energy. By maximizing the energy growth in the course of the integration the singular-vector algorithm then automatically identifies among all flow fields with the same rotational part their least energetic purely rotational member.

References

Achatz, U., and G. Schmitz, 2005: Optimal growth in inertia-gravity wave packets: Energetics, long-term development, and three-dimensional structure. *J. Atmos. Sci.*, submitted

Afanasyev, Y. D., and W.R. Peltier, 2001: Numerical simulations of internal gravity wave breaking in the middle atmosphere: The influence of dispersion ad three-dimensionalization. *J. Atmos. Sci.*, **58**, 132-153

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

Andreassen, Ø., Wasberg, C.E., Fritts, D.C., and J.R. Isler, 1994: Gravity wave breaking in two and three dimensions. 1. Model description and comparison of two-dimensional evolutions. *J. Geophys. Res.*, **99**, 8095-8108

Bakas, N.A., Ioannou, P.J., and G.E. Kefaliakos, 2001: The emergence of coherent structures in stratified shear flow. *J. Atmos. Sci.*, **58**, 2790-2806

Becker, E., and G. Schmitz, 2002: Energy deposition and turbulent dissipation owing to gravity waves in the mesosphere. *J. Atmos. Sci.*, **59**, 54-68

Booker, J.R., and F.P. Bretherton, 1967: The critical layer for internal gravity waves in a shear flow. *J. Fluid Mech.*, **27**, 513-539

Bretherton, F.P., 1966: The propagation of groups of internal gravity waves in a shear flow. *Quart. J. Roy. Met. Soc.*, **92**, 466-480

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

Butler, K.M., and B.F. Farrell, 1992: Three dimensional optimal perturbations in viscous shear flow. *Phys. Fluids A*, **4**, 1637-1650

Chimonas, G., and C.O. Hines, 1986: Doppler ducting of atmospheric gravity waves. *J. Geophys. Res.*, **91**, 1219-1230

Dunkerton, T.J., 1997: Shear instability of internal inertia-gravity waves. *J. Atmos. Sci.*, **54**, 1628-1641

Durran, D.R., 1999: *Numerical methods for wave equations in geophysical fluid dynamics.* Springer, 465 pp.

Ellingsen, T., and E. Palm, 1975: Stability of linear flow. *Phys. Fluids*, **18**, 487-488

Farrell, B.F., 1988a: Optimal excitation of neutral Rossby waves. *J. Atmos. Sci.*, **45**, 163-172

Farrell, B.F., 1988b: Optimal excitation of perturbations in viscous shear. *Phys. Fluids*, **31**, 2093-2102

Farrell, B.F., and P.J. Ioannou, 1993a: Transient development of perturbations in stratified shear flow. *J. Atmos. Sci.*, **50**, 2201-2214

Farrell, B.F., and P.J. Ioannou, 1993b: Optimal excitation of three-dimensional perturbations in viscous constant shear flow. *Phys. Fluids A*, **5**, 1390-1400

Farrell, B.F., and P.J. Ioannou, 1996a: Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci.*, **53**, 2025-2040

Farrell, B.F., and P.J. Ioannou, 1996b: Generalized stability theory. Part II: Nonautonomous operators. *J. Atmos. Sci.*, **53**, 2041-2053

Fritts, D.C., and P.K. Rastogi, 1985: Convective and dynamical instabilities due to gravity wave motions in the lower and middle atmosphere: Theory and observations. *Radio Sci.*, **20**, 1247-1277

Fritts, D.C., and L. Yuan, 1989a: An analysis of gravity wave ducting in the atmosphere: Eckarts resonances in thermal and Doppler ducts. *J. Geophys. Res.*, **94**, 18455-18466

Fritts, D.C., and L. Yuan, 1989b: Stability analysis of inertio-gravity wave structure in the middle atmosphere. *J. Atmos. Sci.*, **46**, 1738-1745

Fritts, D.C., J.R. Isler, and Ø. Andreassen, 1994: Gravity wave breaking in two and three dimensions. 2. Three-dimensional evolution and instability structure. *J. Geophys. Res.*, **99**, 8109-8123

Fritts, D.C., C. Bizon, J.A. Werne, and C.K. Meyer, 2003: Layering accompanying turbulence generation due to shear instability and gravity-wave breaking. *J. Geophys. Res.*, **108**, 8452, doi:10.1029/2002JD002406

Garcia, R.R., and S. Solomon, 1985: The effect of breaking waves on the dynamics and chemical composition of the mesosphere and lower thermosphere. *J. Geophys. Res.*, **90**, 3850-3868

Giering, R., and T. Kaminski, 1998: Recipes for Adjoint Code Construction. *ACM Trans. Math. Software*, **24**, 437-474

Hines, C.O., 1960: Internal atmospheric gravity waves at ionospheric heights. *Can. J. Phys.*, **38**, 1441-1481

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

Hines, C.O.,, 1997b: Doppler spread parameterization of gravity-wave momentum deposition in the middle atmosphere. Part 2. Broad spectra and quasi monochromatic waves, and implementation. *J. Atmos. Sol. Terr. Phys.*, **59**, 387-400

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

Holton, J.R., 1983: The influence of gravity wave breaking on the general circulation of the middle atmosphere. *J. Atmos. Sci.*, **40**, 2497-2507

Houghton, J.T., 1978: The stratosphere and mesosphere. *Q. J. R. Meteorol. Soc*, **104**, 1-29

Howard, L., 1961: A note on a paper of John W. Miles. *J. Fluid Mech.*, **10**, 509-512

Isler, J.R., Fritts, D.C., Andreassen, Ø., and C.E. Wasberg, 1994: Gravity wave breaking in two and three dimensions. 3. Vortex breakdown and transition to isotropy. *J. Geophys. Res.*, **99**, 8125-8137

Klostermeyer, J., 1982: On parametric instabilities of finite-amplitude internal gravity waves. *J. Fluid Mech.*, **119**, 367-377

Klostermeyer, J., 1983: Parametric instabilities of internal gravity waves in Boussinesq fluids with large Reynolds numbers. *Geophys. Astrophys. Fluid Dyn.*, **26**, 85-105

Klostermeyer, J., 1991: Two- and three-dimensional parametric instabilities in finite amplitude internal gravity waves. *Geophys. Astrophys. Fluid Dyn.*, **61**, 1-25

Kwasniok, F., and G. Schmitz, 2003: Radiating instabilities of internal inertio-gravity waves. *J. Atmos. Sci.*, **60**, 1257-1269

Landahl, M.T., 1980: A note on an algebraic instability of inviscid parallel shear flows. *J. Fluid Mech.*, **98**, 243-251

Lehoucq, R.B., D.C. Sorensen, and C. Yang, 1998: *ARPACK users' guide: Solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*, SIAM, 160 pp.

Lelong, M.-P., and D.J. Dunkerton, 1998a: Inertia-gravity wave breaking in three dimensions. Part I: Convectively stable waves. *J. Atmos. Sci.*, **55**, 2473-2488

Lelong, M.-P., and D.J. Dunkerton, 1998b: Inertia-gravity wave breaking in three dimensions. Part II: Convectively unstable waves. *J. Atmos. Sci.*, **55**, 2489-2501

Lindzen, R.S., 1981: Turbulence and stress owing to gravity wave and tidal breakdown. *J. Geophys. Res.*, **86**, 9707-9714

Lombard, P.N., and J.R. Riley, 1996: Instability and breakdown of internal gravity waves. I. Linear stability analysis. *Phys. Fluids*, **8**, 3271-3287

Lübken, F.-J., 1997: Seasonal variation of turbulent energy dissipation rates at high latitudes as determined by in situ measurements of neutral density fluctuations. *J.*

Geophys. Res., **102**, 13441-13456

Mathews, J., and R.L. Walker, 1970: *Mathematical methods of physics*. Addison-Wesley, 501 pp.

Medvedev, A.S., and G.P. Klaassen, 1995: Vertical evolution of gravity wave spectra and the parameterization of associated gravity wave drag. *J. Geophys. Res.*, **100**, 25841-25853

Mied, R.P., 1976: The occurrence of parametric instabilities in finite amplitude internal gravity waves. *J. Fluid Mech.*, **78**, 763-784

Miles, J.W., 1961: On the stability of heterogeneous shear flows. *J. Fluid Mech.*, **10**, 496-508

Moffat, H.K., 1967: The interaction of turbulence with strong shear. *Atmospheric Turbulence and Radio Wave Propagation*, A.M. Yaglom and V.I. Tatarsky, Eds., Nauka, 139-161

Müllemann, A., M. Rapp, and F.-J. Lübken, 2003: Morphology of turbulence in the polar summer mesopause region during the MIDAS/SOLSTICE campaign 2001. *Adv. Space Res.*, **31**, 2069-2074

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

Orr,W. M'F., 1907: The stability or instability of the steady motions of a perfect liquid and of a viscous liquid, *Proc. Roy. Irish Acad.*, **27A**, 9-217

Schmid, P.J., and D.S. Henningson, 2001: *Stability and transition in shear flows.* Springer, 556 pp.

Sonmor, L.J., and G.P. Klaassen, 1997: Toward a unified theory of gravity wave stability, *J. Atmos. Sci.*, **54**, 2655-2680

Trefethen, L.N., A.E. Trefethen, S.C. Reddy, and T.A. Driscoll, 1993: Hydrodynamic stability without eigenvalues. *Science*, **261**, 578-584

Warner, C.D., and M.E. McIntyre, 2001: An ultra-simple spectral parameterization for non-orographic gravity waves. *J. Atmos. Sci.*, **58**, 1837-1857

Werne, J., and D.C. Fritts, 1999: Stratified shear turbulence: Evolution and statistics. *Geopys. Res. Lett.*, **26**, 439-442

Winters, K.B., and E.A. D'Asaro, 1994: Three-dimensional wave instability near a critical level. *J. Fluid Mech.*, **272**, 255-284

Yau, K.-H., G.P. Klaassen, and L.J. Sonmor, 2004: Principal instabilities of large amplitude inertio-gravity waves. *Phys. Fluids*, **16**, 936-951

Yuan, L., and D.C. Fritts, 1989: Influence of a mean shear on the dynamical instability of an inertia-gravity wave. *J. Atmos. Sci.*, **46**, 2562-2568

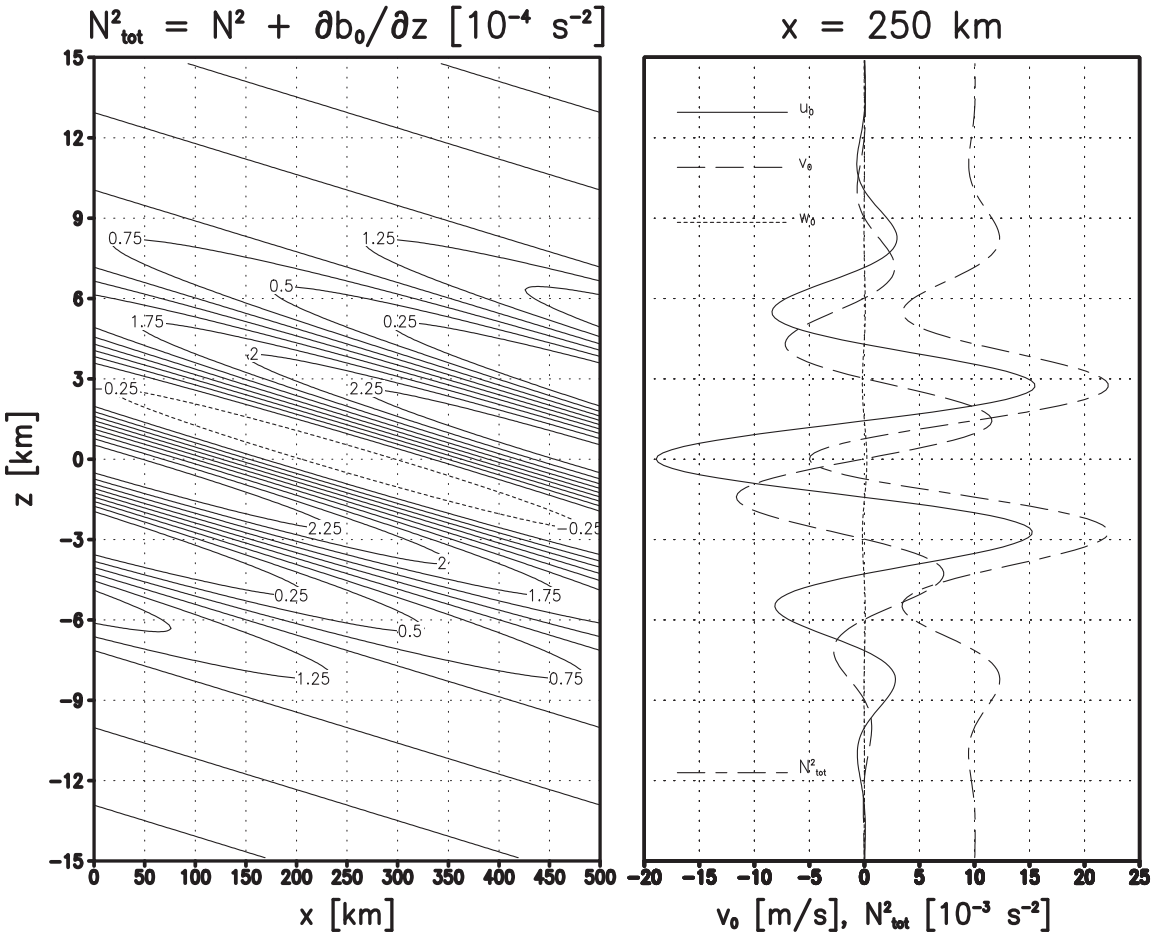

Figure 1: Squared total Brunt-Vaisala frequency in a gravity wave packet with $a = 1.5$ (left panel, negative values are indicated by dashed contours), and at $x = \Lambda_x/2$ its vertical dependence and that of the three-dimensional velocity field (right).

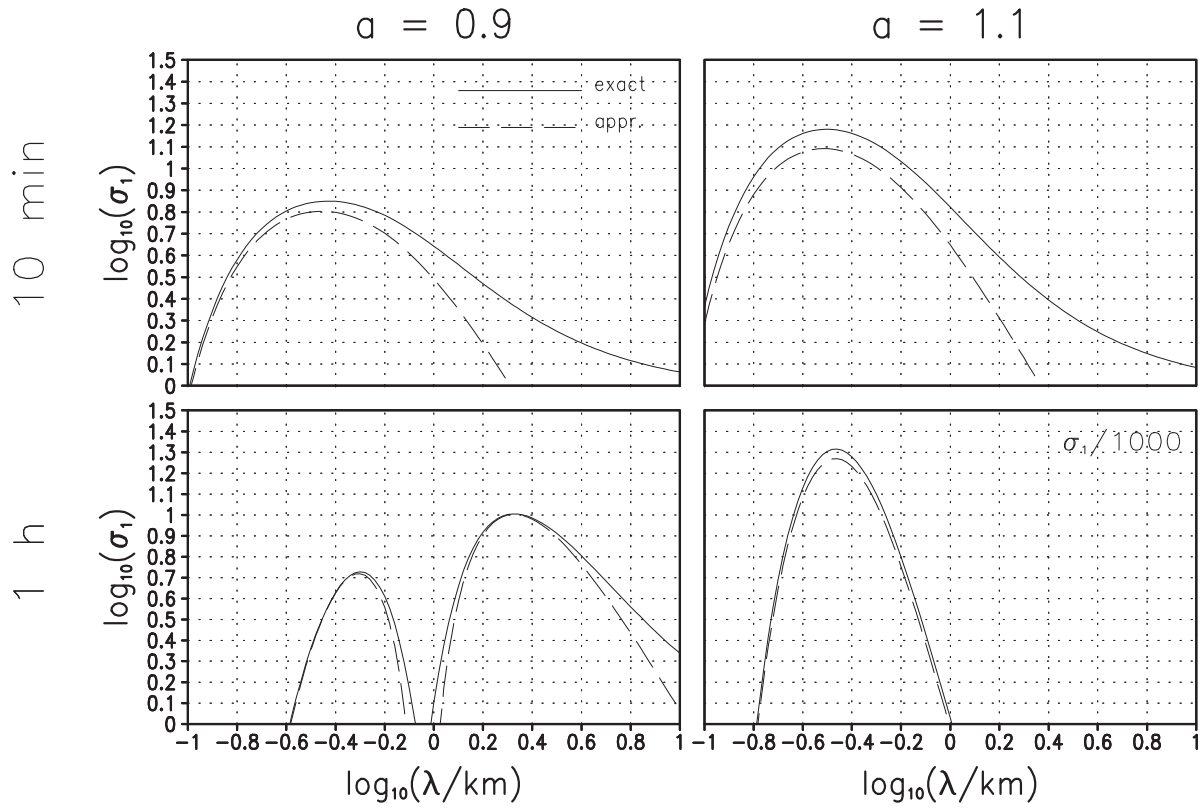

Figure 2: For the shear-layer approximation of optimal growth of parallel perturbations ($\alpha = 0$, vertical wavelength $\lambda_z = 600$ m) in an IGW packet ($\Lambda_z = 6$ km) with two different amplitudes a the leading growth factors for optimization times $\tau = 10$ min and 1 h, in dependence of the horizontal wavelength of the perturbation. Shown are both the exact values and the approximation given by equation (44). In the case $(a, \tau) = (1.1, 10\text{min})$ the growth factors have been divided by a factor 1000.

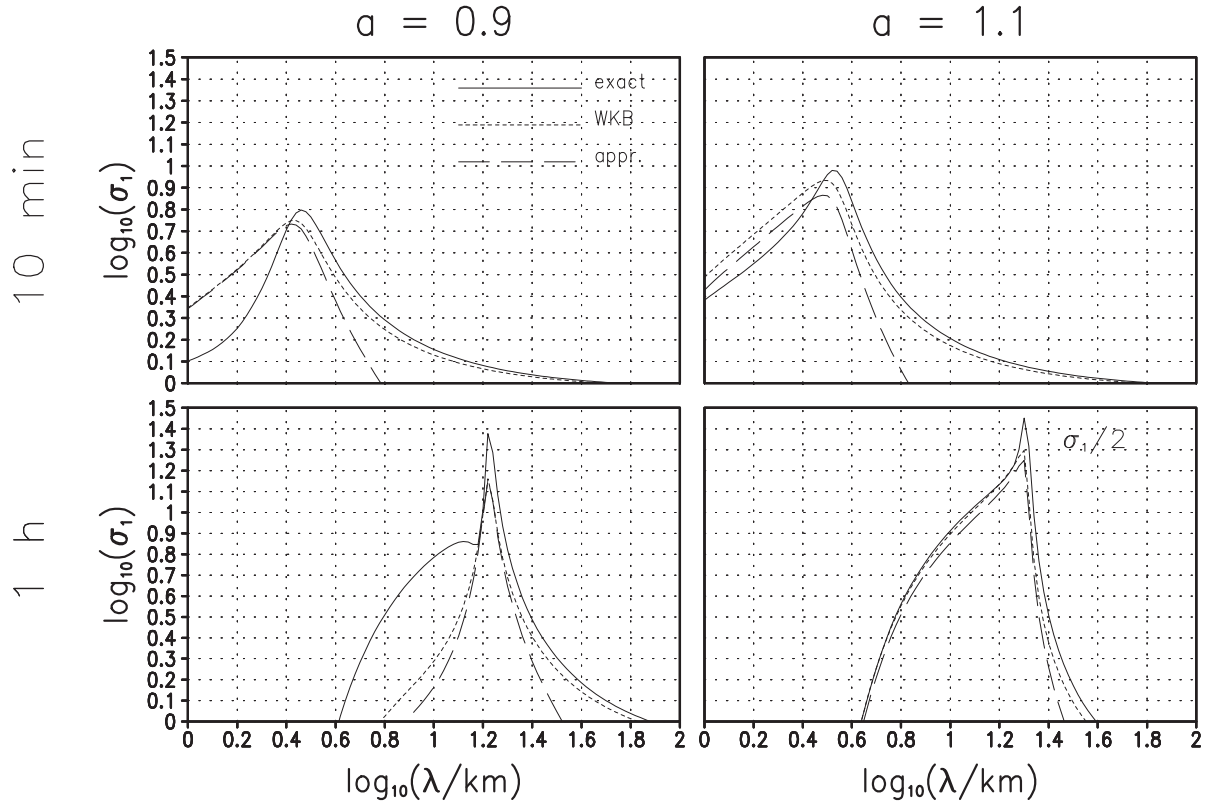

Figure 3: For the shear-layer approximation of optimal growth of transverse perturbations ($\alpha = 90°$, vertical wavelength $\lambda_z = 600$ m) in an IGW packet ($\Lambda_z = 6$ km) with two different amplitudes a the leading growth factors for optimization times $\tau = 10$ min and 1 h, in dependence of the horizontal wavelength of the perturbation. Shown are both the exact values, their WKB approximation by equation (A5), and the approximation given by equation (52). In the case $(a, \tau) = (1.1, 10\text{min})$ the growth factors have been divided by a factor 2.

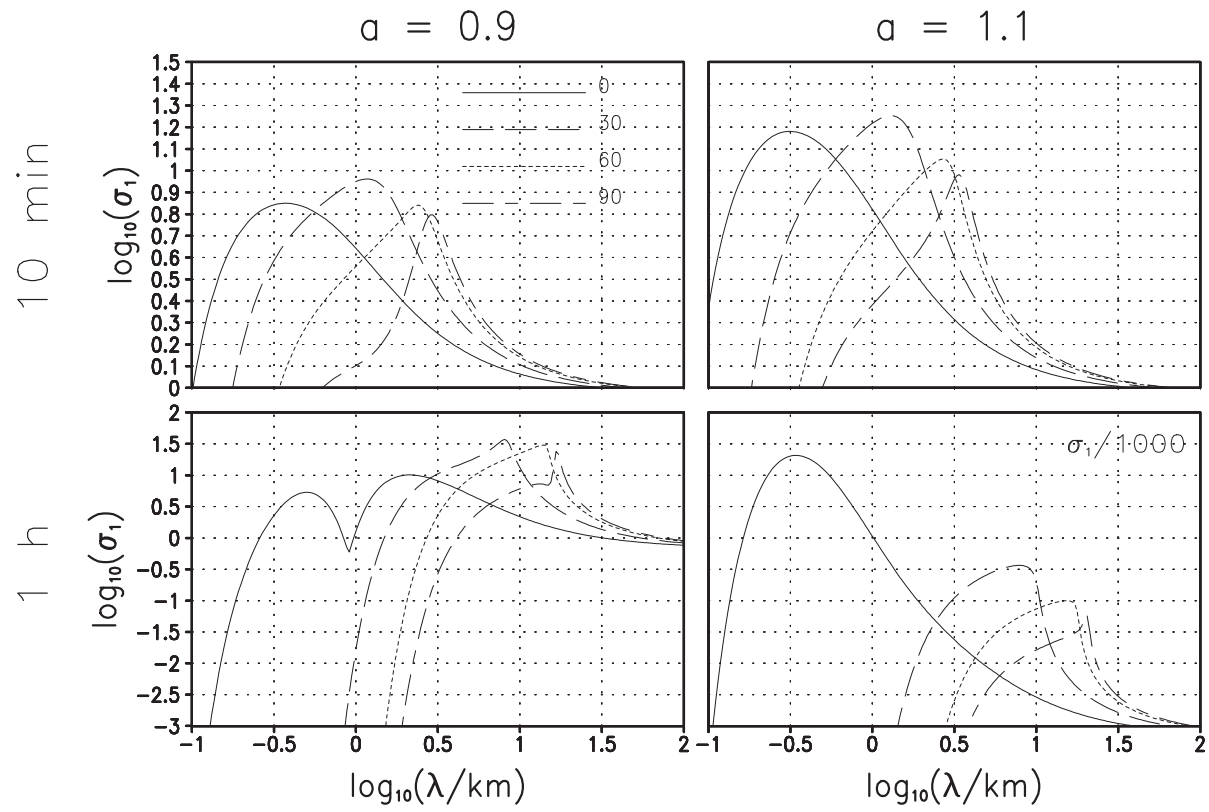

Figure 4: For the same four cases shown in Figs. 2 and 3, the optimal growth factors for the azimuth angles $\alpha = 0°, 30°, 60°$, and $90°$.

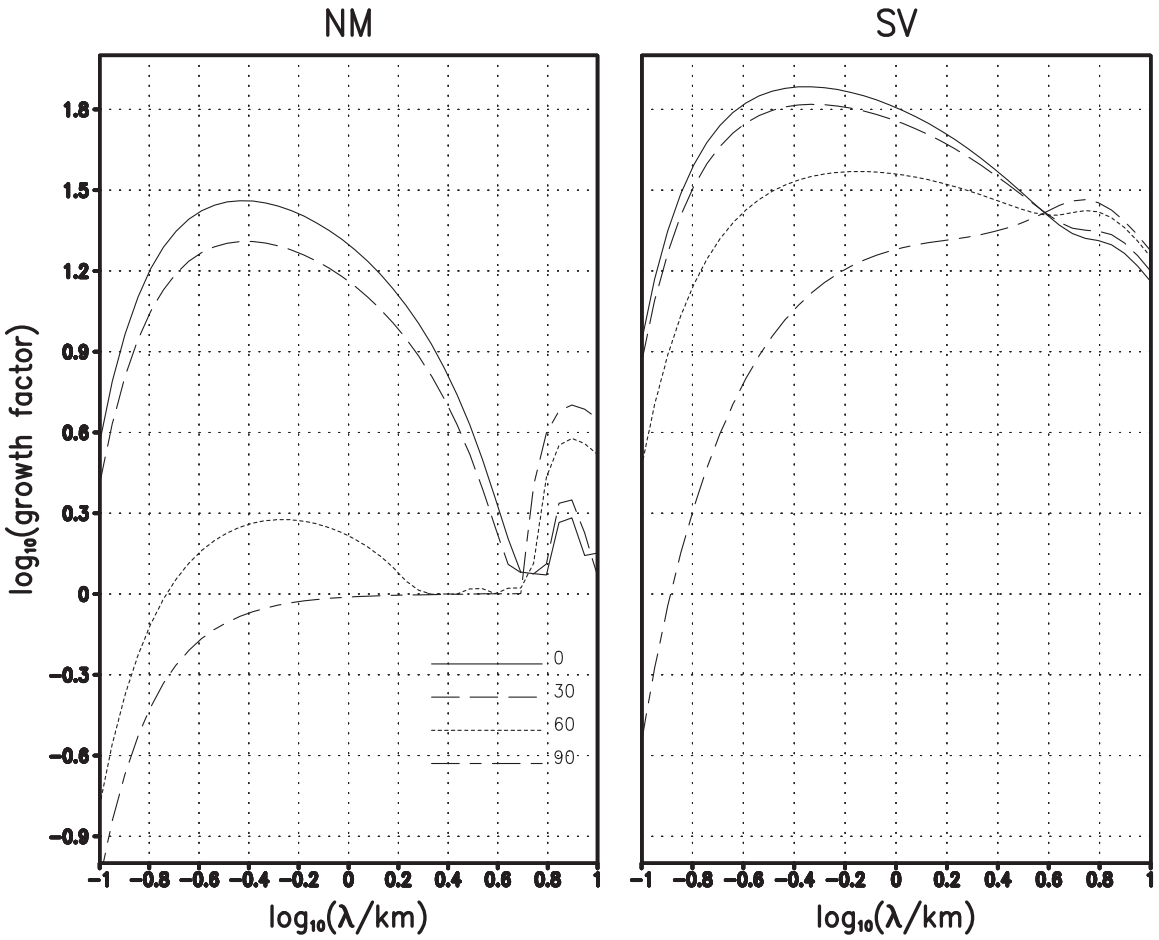

Figure 5: For the vertical profile at the horizontal location of maximal static instability, of a statically unstable ($a = 1.5$) IGW packet with vertical wavelength $\Lambda_z = 6$ km, the growth factors of the most unstable normal modes (left panel) and leading singular vectors (right panel), propagating in the horizontal at various azimuth angles α with respect to the basic wave, as a function of the parallel wave length. Integration time is $\tau = 10$ min.

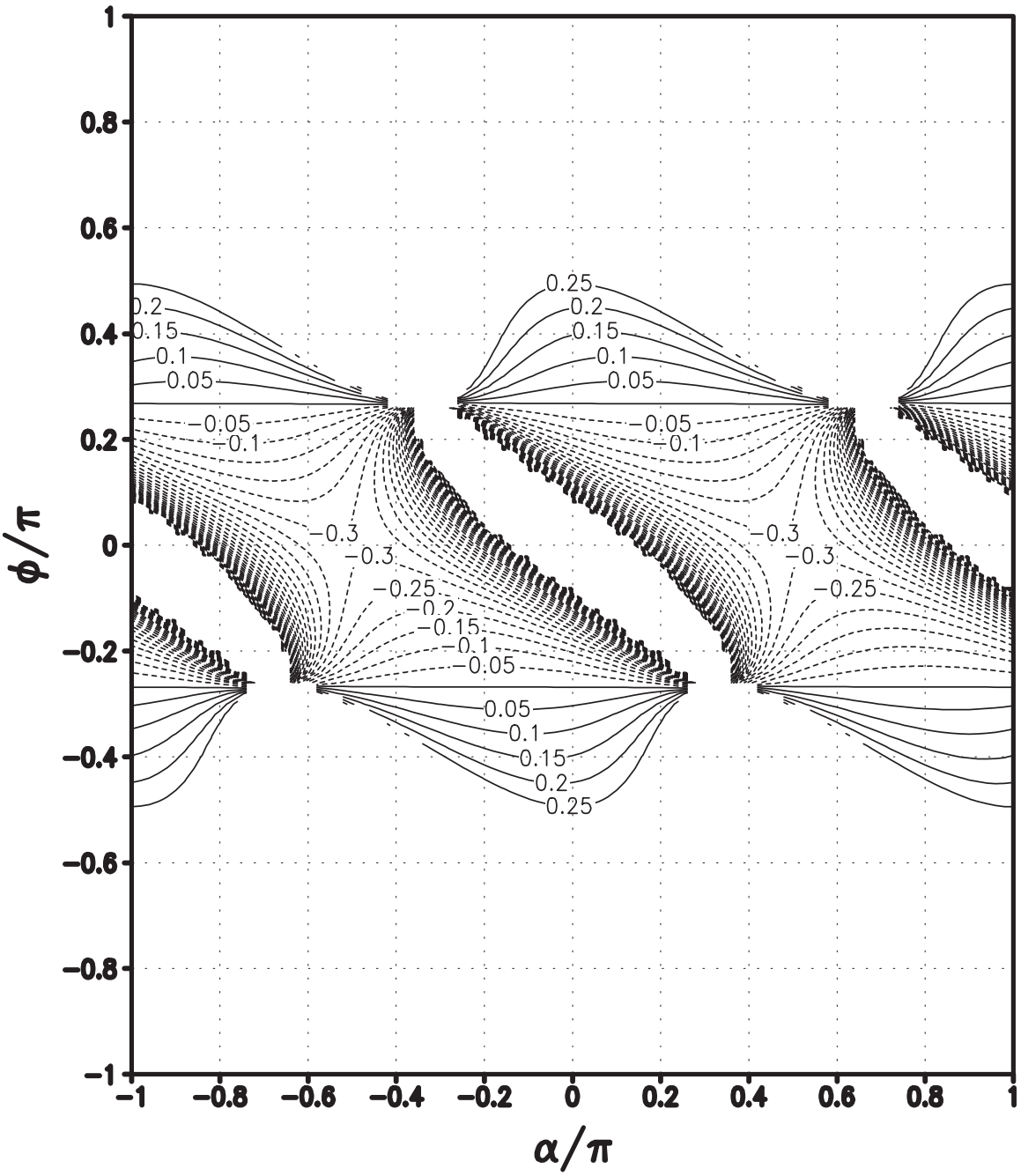

Figure 6: As a function of instability azimuth α and wave phase ϕ, the local Richardson number in the vertical IGW profile used in the analysis shown in figure 5. Only values $-2 \leq \mathrm{Ri} \leq 1/4$ are shown.

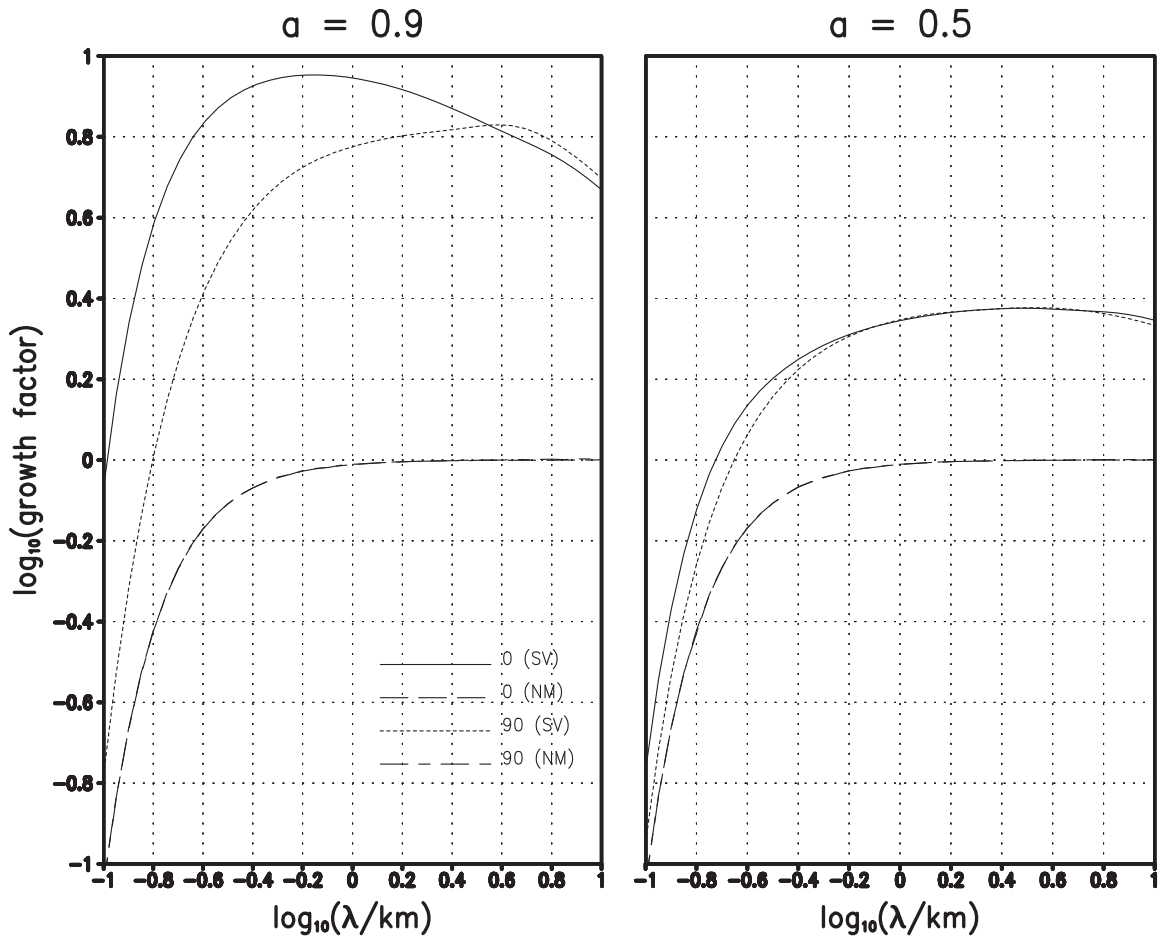

Figure 7: As Fig. 5, but just for transverse and parallel perturbations, and the basic wave amplitudes $a = 0.9$ and 0.5.

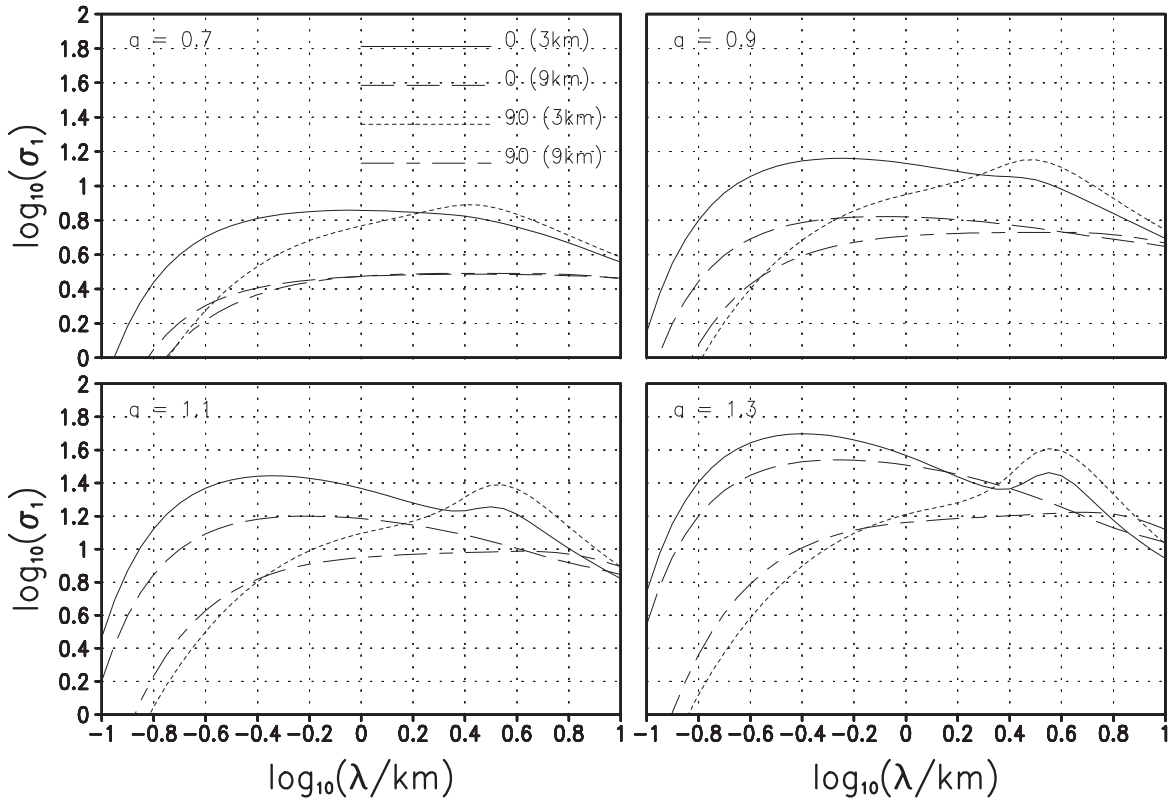

Figure 8: For the basic-wave vertical wavelengths $\Lambda_z = 3$ km and 9 km, the wavelength-dependence of the leading growth factors for parallel and transverse perturbations, at basic-wave amplitudes $a = 0.7$, 0.9, 1.1, and 1.3.

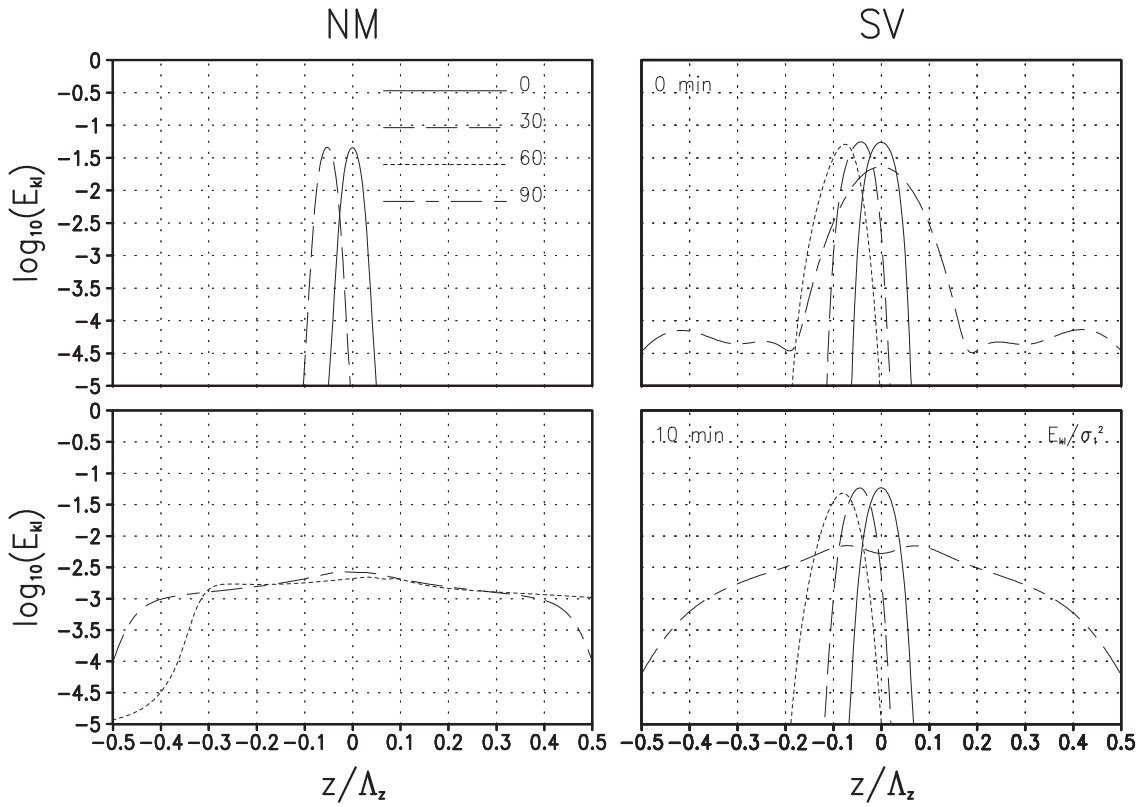

Figure 9: For an IGW packet with $\Lambda_z = 6$ km and $a = 1.5$ and azimuth angles $\alpha = 0°, 30°, 60°$, and $90°$ the energy density E_{kl} of the leading normal modes (left panels) and the leading singular vectors at initialization (upper right) and optimization time (lower right, divided by σ_1^2), in arbitrary units.

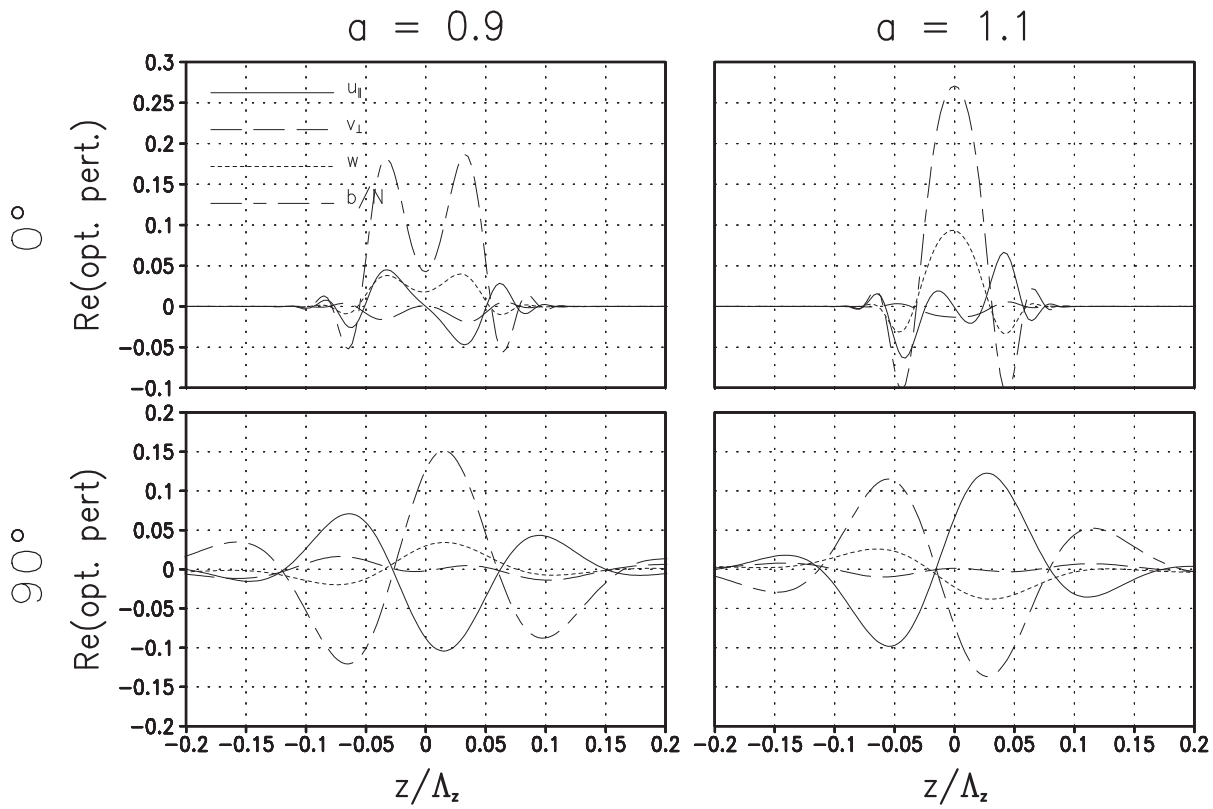

Figure 10: For an IGW packet with $\Lambda_z = 6$ km and $a = 0.9$ (left column) or 1.1 (right) and azimuth angles $\alpha = 0°$ (top row) and 90° (bottom) the altitude-dependence of the real part of the four model variables in the leading optimal perturbation, in arbitrary units.

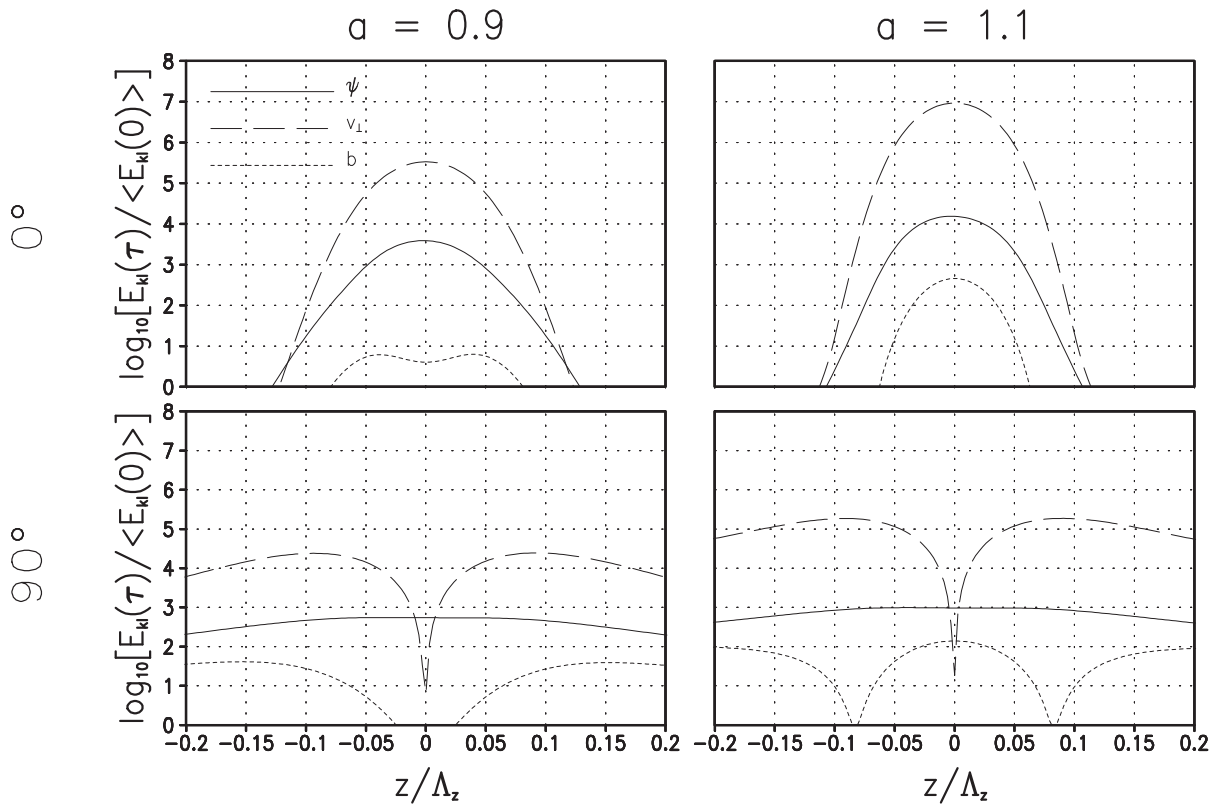

Figure 11: For an IGW packet with $\Lambda_z = 6$ km and $a = 0.9$ (left column) or 1.1 (right) and azimuth angles $\alpha = 0°$ (top row) and 90° (bottom) the altitude-dependence of the relative amplification during the optimization time $\tau = 10$ min in the four model variables in the leading optimal perturbation.