Annales
Geophysicae

New insights from a nonlocal generalization of the Farley-Buneman instability problem at high latitudes

J. Drexler, J.-P. St.-Maurice, D. Chen, and D. R. Moorcroft

Department of Physics and Astronomy, University of Western Ontario, London, Canada

Abstract. When their growth rate becomes too small, the E-region Farley-Buneman and gradient-drift instabilities switch from absolute to convective. The neutral density gradient is what gives the instabilities their convective character. At high latitudes, the orientation of the neutral density gradient is close to the geomagnetic field direction. We show that this causes the wave-vector component along the geomagnetic field to increase with time. This in turn leads to wave stabilization, since the increase goes hand-in-hand with an increase in parallel electric fields that ultimately short-circuits the irregularities. We show that from an equivalent point of view, the increase in the parallel wave vector is accompanied by a large upward group velocity that limits the time during which the perturbations are allowed to grow before escaping the unstable region. The goal of the present work is to develop a systematic formalism to account for the propagation and the growth/decay of high-latitude Farley-Buneman and gradient-drift waves through vertical convective effects. We note that our new formalism shies away from a plane wave decomposition along the magnetic field direction. A study of the solution to the resulting nonlinear aspect angle equation shows that, for a host of initial conditions, jump conditions are often triggered in the parallel wave-vector (defined here as the vertical derivative of the phase). When these jump conditions occur, the waves turn into strongly damped ion-acoustic modes, and their evolution is quickly terminated. We have limited this first study to Farley-Buneman modes and to a flow direction parallel to the electron $\boldsymbol{E} \times \boldsymbol{B}$ drift. Our initial findings indicate that, irrespective of whether or not a jump in aspect angle is triggered by initial conditions, the largest amplitude modes are usually near the ion-acoustic speed of the medium (although Doppler shifted by the ion motion), unless the growth rates are small, in which case the waves tend to move at the same drift as the ambient electrons.

Correspondence to: J. Drexler (jdrexler@uwo.ca)

1 Introduction

Understanding the properties of large amplitude Farley-Buneman and gradient-drift waves at high latitudes has remained a challenge in spite of several decades of studies. In large part this is because the large amplitude waves are usually strongly nonlinear by the time they reach the amplitudes that dominate the radar or rocket observations. In the usual Fourier treatment of the instability, this means that one should, in principle, include coupling between all possible modes of the system, even linearly stable ones.

Some success has nevertheless emerged from analytical work dealing with the waves under turbulent conditions. For instance, in work that was originally meant for equatorial applications, Sudan and Keskinen (1979) and Sudan (1983b) have been able to establish that, under weakly-growing conditions, mode-coupling should lead to cascading towards smaller wavelengths and should, at lower frequencies, produce isotropic turbulence in a plane perpendicular to the geomagnetic field.

On a more controversial note, Sudan (1983a) also proposed that the wave fields of faster-growing Farley-Buneman modes might scatter the electrons, thereby producing an "anomalous" amount of diffusion. Robinson (1986) extended this idea to the high latitudes where the growth rates can often be very large, and he used this mechanism to suggest that the large amplitude modes should move at the ion-acoustic speed of the medium. St.-Maurice (1987) countered that there was a problem with the physics used in the anomalous diffusion theories, in that the diffusion produced by unstable waves is perpendicular to their propagation direction when in fact, one needs to introduce diffusion at right angles to that direction if the waves are to saturate in response to increased diffusion. As a result, St.-Maurice (1990) tried to generalize the anomalous diffusion idea by including a bath of background waves pointing in all directions in the plane perpendicular to the magnetic field. He concluded that under anisotropic turbulent conditions, the suggested mechanism could work, but at the expense of having broad-

band density levels of the order of 20%, which seemed contrary to observations. However, for weakly growing conditions with isotropic two-dimensional turbulence, St.-Maurice (1990) found that an equivalent (or "anomalous") collision frequency of the order of four times the electron-neutral collision frequency was possible. This result seemed to be most useful in the equatorial regions near the 100 km altitude region, by providing an explanation for observed current densities that are systematically smaller than inferred from model calculations based on classical conductivity profiles (Gagnepain et al., 1977). However, the inferred anomalous collision frequencies were too small to provide a saturation mechanism for the faster growing high-latitude waves.

The limited success of the anomalous diffusion concept led Hamza and St.-Maurice (1993a) to explore the consequences of mode-coupling under strongly growing Farley-Buneman conditions. They found that as long as the nonlinear evolution is dominated by coupling between modes that are perpendicular to B, the waves should indeed saturate at a phase velocity equal to the ion-acoustic speed, while the spectral width would increase to become, in velocity units, comparable to the ion-acoustic speed at right angles to the flow. Furthermore, the Doppler shift would not equal the line-of-sight velocity of the ambient electron drift, even when that speed is less than the ion-acoustic speed to start with. At least as far as predictions on the Doppler width were concerned, a reasonable agreement seems to be found for observations made with radars operating at frequencies exceeding 30 MHz (Eglitis and Robinson, 1998). Hamza and St.-Maurice (1993b) also estimated that the mode-coupling hypothesis was consistent with a broad-band density level of the order of 10% or less. This appears to be consistent with rocket observations. However, just obtaining estimates for the turbulent level proved to be a daunting task, even under the assumption of pure two-dimensional turbulence involving non-dispersive modes.

In order to work around the complexity associated with mode-coupling, St.-Maurice and Hamza (2001) have proposed instead to do away with Fourier analysis under two-dimensional conditions associated with negligible wave electric field components along the geomagnetic direction. They introduced intermittency in lieu of a uniform wave background and replaced the mode-coupling description with an equivalent nonlinear diffusion coefficient that depended on the density itself. Associated with this diffusion coefficient was a total electrostatic field that rotated and decreased inside the blobs and holes, much like in earlier numerical simulations by Otani and Oppenheim (1998). The evolution of the field was such as to bring saturation at the ion-acoustic speed of the medium, but in a direction that was no longer associated with the original electron plasma drift direction. This result had been uncovered earlier in a numerical study by Janhunen (1992) (see below).

As indicated in the previous paragraph, numerical tools have also been useful in improving our understanding of Farley-Buneman (and gradient-drift) waves. Thus far, however, the studies associated with the Farley-Buneman insta-

bilities have been limited to two dimensions. Machida and Goertz (1988), as well as Thiemann and Schlegel (1994), used two-dimensional particle simulations in the plane made by the magnetic field and the $E \times B$ directions. This approach does not allow mode-coupling in the plane perpendicular to the magnetic field, which appears to be a substantial sink of energy for electric fields less than 50 mV/m (Hamza and St.-Maurice, 1993a). Nevertheless, Machida and Goertz (1988) focused on the electron heating question and concluded that the electrons were heated by perpendicular fluctuation fields through what they attributed to be anomalous collisions in spite of the geometrical problems associated with this mechanism. Thiemann and Schlegel (1994) focused on properties observed by radars, such as the phase velocity and the power as a function of electric field. They uncovered phase velocities of the order of the ion-acoustic speed at saturation. Much like Machida and Goertz (1988) before them, they saw the simulated electron temperature go up, but they did not comment much on its physical origin aside from noting that this heating phenomenon was not observed with simulations made in the plane perpendicular to the magnetic field. Their electron temperature increased with electric field strength, in a way reminiscent of observations. This being said, the number of particles being used was relatively small and the fluctuation levels were accordingly large.

In a different kind of numerical study, Janhunen (1992) used a simplified three-dimensional fluid approach to conclude that, for relatively weak electric fields, the high-latitude Farley-Buneman waves reach a limited amplitude because the magnitude of the parallel wave electric field increases monotonically with time. In the process, however, Janhunen (1992) found no evidence for perpendicular turbulence, probably because the electric fields he used were not large (see below). Using a particle simulation, Janhunen (1994) later studied the waves in a plane perpendicular to the magnetic field direction. He found that the largest amplitude waves were at a flow angle for which the phase velocity was equal to the ion-acoustic speed of the medium, while lower amplitude waves could be found in the plasma flow direction moving faster than the ion-acoustic speed. In contrast to the above quoted two-dimensional simulations that included the magnetic field direction, Janhunen (1994) found little evidence for electron heating.

More recently, Oppenheim et al. (1995) also used a particle simulation in which turbulence was limited to the plane perpendicular to the geomagnetic field. Among other things, their simulation produced a rotation of the total electrostatic field inside structures. However, the results from the simulations seem to produce larger rotations than expected from the intermittency theory of St.-Maurice and Hamza (2001). Still, Otani and Oppenheim (1998) were able to associate the rotation obtained in the Oppenheim et al. (1995) simulations, with coupling between dominant modes.

It should be clear from the above that, aside from a couple of limited numerical studies, most of our progress in the understanding of large amplitude E-region structures has come from two-dimensional studies in a plane perpendicular to the

geomagnetic field. Such studies have, by necessity, had to neglect the effects of electric fields parallel to the geomagnetic field. However, there are two contrasting situations for which the neglect of parallel electric fields cannot so easily be justified.

The first situation is met at high latitudes when the ambient electric field exceeds roughly 50 mV/m. In this case, much of the coupling between perpendicular modes results in the production of modes that are themselves linearly unstable. Perpendicular mode-coupling can, therefore, no longer be invoked for the saturation of the waves. In addition, explosive electron heating is seen to be taking place when the electric field exceeds that same 50 mV/m (e.g. St.-Maurice et al., 1999, and references therein). This suggests that for strong electric fields, mode-coupling must be triggering modes with a large parallel wave vector (i.e. electric field) components along the geomagnetic field (Hamza and St.-Maurice, 1993b). These modes must have a large enough parallel component to be able to decay if they are to eliminate the accumulating wave energy. In the process, the "large" parallel electric fields also heat the electrons. The word "large" has to be put in quotes here because, on average, only wave vectors at 2 or 3° away from perpendicularity are needed (St.-Maurice and Laher, 1985).

The second situation for which parallel electric fields should not be neglected is, somewhat paradoxically, the opposite of the first, namely one for which the growth rates of the waves are so small that the waves become convectively unstable through the inhomogeneity of the medium (St.-Maurice, 1985). This situation, which will be the focus of the present work, has been labeled a "nonlocal" problem in the high-latitude context. The problem becomes nonlocal because the altitude variations of the collision frequencies of ions and electrons are forcing the modes to evolve along the geomagnetic field. This evolution is caused by an altitude change in the eigenfrequency, as well as in the relative ion-electron drift. The resulting parallel electric fields have an important impact on the evolution of the waves because they can quickly short out the waves, i.e. lead to wave saturation (St.-Maurice, 1985; Janhunen, 1992). The fact that the unstable part of the E-region is only 20 km thick and that waves cannot grow if the aspect angle exceeds approximately 1° also means that the structures should not exceed 100 m in the direction perpendicular to the magnetic field (e.g. Janhunen, 1992; St.-Maurice and Laher, 1985).

We should point out that the high-latitude nonlocal problem differs in one important respect from nonlocal equatorial studies (e.g. Riggin and Kadish, 1989; Ronchi et al., 1991; Hu and Bhattacharjee, 1998). At low latitudes, the background inhomogeneities are perpendicular to the horizontal magnetic field lines. This means that the low-latitude structures are not limited as much by an aspect angle constraint: even in the presence of a 0.5° aspect angle, they can reach a few km in size in the east-west direction. However, nonlocal calculations are still needed for these kinds of sizes if the structures are to reach 10 km sizes in the vertical direction (so as to maintain a basic east-west direction for the

wave vector). Nonlocal calculations show that not only is this possible, but in fact nonlinear effects seem to be such as to make the km size structures the preferred size of the system. The bottom line is that the system can sustain a total wave-vector pointing essentially in the east-west direction. These gradient-drift waves are much larger in size than their high-latitudes counterpart because there is little at low latitudes in the way of parallel electric fields to damp the waves.

A strong observational motivation for the study of nonlocal effects at high latitudes is the HF radar detection of modes, that in many respects look rather different from their higher frequency counterpart (Villain et al., 1987, 1990; Hanuise et al., 1991; Eglitis et al., 1995; Jayachandran et al., 2000; Eglitis and Robinson, 1998). Theoretically, we know that 10–15 m modes associated with 10–15 MHz radars (a typical SuperDARN range of radar frequencies) have rather small growth rates, of the order of $1\ s^{-1}$ as opposed to 1000 s^{-1} for their higher frequency counterpart. This means, as we show below, that the former is strongly affected by convective (or nonlocal) effects associated with changing parallel electric fields. This could have important implications for the Doppler shift of the waves. It might be at the origin of some important differences between HF observations and higher frequency observations. For instance, there are many indications that, under moderate to weak electric field conditions, the Doppler shift of the waves is equal to the line-of-sight velocity of the electrons (Villain et al., 1987; Jayachandran et al., 2000). This being said, there is also a statistically clear trend for 10 to 15 m E-region waves to saturate at approximately 450 m/s (Milan et al., 1997; Lacroix and Moorcroft, 2001) under certain conditions. These conditions may well have to do with the larger growth rates associated with stronger electric fields. In that case, convective/nonlocal effects might be less important.

In the present paper, we primarily seek to establish a systematic framework for the study of nonlocal effects in the high-latitude E-region. Our ultimate goal is a proper description of the evolution of slowly growing modes, such as those that are often detected by HF radars. For simplicity, our main focus in the present work will be Farley-Buneman waves, even though we will indicate how our work can be later generalized to gradient-drift waves.

In Sects. 2 and 3, we use a WKB approach to introduce the concept of a "frequency" that is defined in terms of a local variation in the phase of the solution to the standard Farley-Buneman wave equation. Unlike the frequency met in the plane wave decomposition used in standard Farley-Buneman or gradient-drift theory, the generalized frequency is viewed as a non-trivial function of position. This proves very useful for a description of the evolution of slowly growing modes. Once the generalized frequency and wave numbers are introduced, we derive the propagation and amplitude equations through the use of a multi-scale expansion in time and space. We show in Sect. 3 that the amplitude equation thus retrieved agrees with the principle of conservation of wave action as long as allowance for a contribution from spontaneous growth is made, since the waves are,
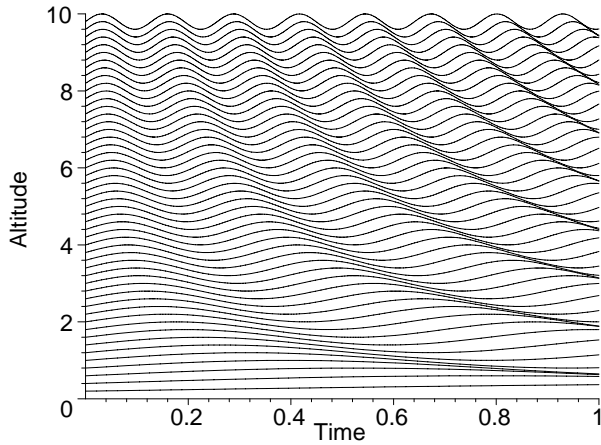
Fig. 1. Phase changes with time and space for a frequency that increases with altitude. All altitudes are assumed to be in phase initially.

Fig. 2. A sketch of the kind of evolution that a single evolving structure (irregularity) would undergo in physical space if it were initially perfectly field-aligned. The vertical is along the magnetic field and the electron $E \times B$ drift direction is to the right.

after all, linearly unstable. Section 4 describes some unusual properties of the resulting equations (intersecting characteristics) and how we used physical and mathematical reasoning to handle the ensuing shocks. Section 5 describes the results of our calculations. We introduce a discussion of the possible implications and our conclusions in Sect. 6.

2 Physical description

Before delving into the mathematical and numerical intricacies of the problem at hand, it should prove useful to describe the physical processes that drive the equations and their solutions. While we go back to the details of the derivation below, we wish to stress here that in the standard analysis of Farley-Buneman (and gradient-drift) waves, the eigenfrequency resulting from a plane wave decomposition is actually a function of position. Specifically, the frequency increases with altitude through the height dependence of the electron and ion collision frequencies (see Eq. 11).

Once it is recognized that the eigenfrequency is a function of altitude, one has to accept that the parallel wave number becomes a function of time. This is illustrated by the sketch shown in Fig. 1. The figure is made of a series of oscillations from various heights. Initially (i.e. $t = 0$ in the figure) the phase of the oscillations is the same at all heights. However, we have introduced a frequency mismatch between various altitudes, with the higher frequencies on top, as in the real ionosphere. After enough oscillations (no matter how many it takes in real situations), a phase change with altitude starts to build up in a way similar to that shown in Fig. 1. Such a phase change is equivalent to the introduction of an oscillation along z at the later times, where there was none initially. A wave vector component along the z direction is, therefore, building up with time. This means that, in principle, the wave vector component parallel to the magnetic field cannot be considered to be a free parameter of the problem.
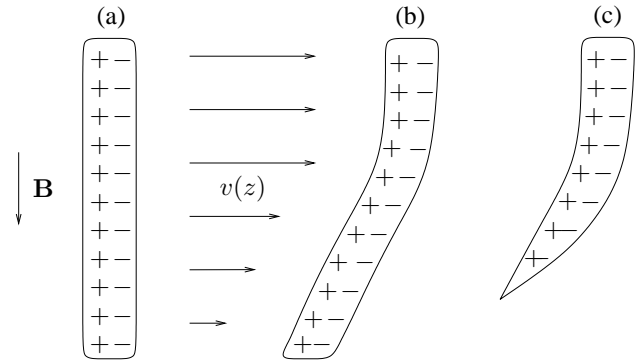
Instead, its value changes with time in a manner that has to be consistent with the variation of the frequency with altitude. To be precise, this relation is given by the Whitham relation shown by Eq. (12) below.

We should add that the temporal evolution of the parallel wave vector is not necessarily important for the evolution of the waves. For it to be important, in practice, the wave has to grow so slowly with time that nonlinear amplitude saturation processes (e.g. Sudan, 1983a; Hamza and St.-Maurice, 1993a) will not take control of the wave evolution before the parallel wave vector does. To assess what kind of growth rates are required for the wave vector to matter, we can compare a typical growth time with the time it takes for a wave packet to travel through the unstable high-latitude E-region. If the travel time is such that the growth is too small for nonlinear processes to take place, then convective effects associated with the parallel wave-vector evolution will have to matter.

This leads us to the fact that the wave vector enters the wave evolution picture mathematically partly through its impact on the parallel component of the group velocity. Unless the parallel wave vector is strictly zero (in which case the parallel component of the group velocity is also zero), the parallel group velocity is of the order of 10 to 30 km/s (Moorcroft, 1984). Since the unstable part of the E-region is only approximately 25 km thick, this means that waves with a growth rate of the order of 1 to 10 s^{-1} are very likely to be affected by convective effects before nonlinear amplitude saturation mechanisms have a chance to come into the picture. Alternatively, we can link the parallel wave vector to a parallel electric field: as the magnitude of the parallel wave vector increases, a point is reached at which the waves can no longer grow because, physically, the parallel electric field associated with the parallel wave vector (the waves are electrostatic) is short circuiting the instability through the response of the highly mobile electrons along the magnetic field lines.

A different way to visualize the role played by the parallel wave vector is to imagine the instability process as a positive

feedback mechanism that enhances an initial density perturbation in time and space. This approach has been followed by St.-Maurice and Hamza (2001) for the two-dimensional case for which parallel electric fields and wave vectors were neglected. When adding the latter, the picture becomes three-dimensional. This situation is illustrated in Fig. 2 through a two-dimensional cut for the case of an initial elongated shoebox-like density enhancement along the magnetic field (zero initial parallel wave-vector situation in Fourier space). For simplicity, the magnetic field in our illustration is aligned with the vertical direction. Also notice that the lack of an initial structure along the magnetic field is arbitrary but it is consistent with the picture drawn in Fig. 1. Finally, as described in St.-Maurice and Hamza (2001), it is important to realize that the instability starts by having the side of the blob that is in the direction of the electron $\boldsymbol{E} \times \boldsymbol{B}$ drift become negatively charged.

As far as the time evolution of the initial shoebox is concerned, the fact that the frequency increases with height means that the box becomes distorted after some time, due to a faster motion (i.e. horizontal group velocity; see Appendix A) at the top than at the bottom of the shoebox. This image seems to be closely analogous to what Janhunen (1992) calls a "tilting wave tower". In any event, as the distortion, or "tilting" becomes measurable, the positively charged edge of the box ends up resting on the same magnetic field line as the other, negatively charged edge. This means that the size of the box is shrinking, particularly lower down: as parallel electric fields form, the more mobile electrons from the negative side of the irregularities are moving up, in response to these fields. The shape of the box is, therefore, changing as the negative bottom moves upwards toward the positively charged top. The motion is faster in regions with larger tilts, i.e. in regions with larger shears in the phase velocity ω_r/k. This motion is reflected in the 15 km/s or greater upward parallel group velocities that can be found particularly near the bottom, where the aspect angles tend to be largest initially, at least for the case evolving from a zero aspect angle initial condition.

We conclude that parallel electric fields have to be expected everywhere around the distorted shoebox, albeit with smaller values at the top than at the bottom of the structures, since the shears are larger lower down. For the geometry and situation depicted in Fig. 2, this means the creation after some initial time of upward group velocities throughout the unstable layer, with the faster group velocities at the bottom initially. This means not only that wave energy tends to be advected towards the top of the unstable layer, but also that it becomes enhanced at the top through a compression associated with a convergence in the parallel group velocity. This picture should help us to understand the features that emerge from the formal calculations undertaken in the sections that follow.

3 Basic derivation

To derive the differential equation that describes the perturbed potential, we follow the standard linearization procedure that can be found in many papers (e.g. Sudan, 1983b) but with allowance for a time-space description that includes three dimensions, not just two (St.-Maurice, 1985). Briefly stated: we use the electron and ion continuity equations in the absence of chemical production terms, and the electron and ion momentum equations with the assumption of isothermal electrons and ions. The electrons are furthermore assumed to be fully magnetized ($\nu_e \ll \Omega_e$, where ν_e is the electron-neutral collision frequency and Ω_e is the electron cyclotron frequency), while the ions are taken to be weakly magnetized ($\nu_i > \Omega_i$). Using charge neutrality ($n_i = n_e$), linearizing the equations and assuming an $\exp(i\boldsymbol{k}_\perp \cdot \boldsymbol{x})$ dependence in the plane perpendicular to the magnetic field, we then obtain the following linearized partial differential equation in terms of the perturbed density $\delta n/n_0$:

$$
\frac{\partial^2}{\partial z^2} \left[\left(\frac{\partial^2}{\partial t'^2} + \nu_i \frac{\partial}{\partial t'} + k_\perp^2 c_s^2 \right) \frac{\delta n}{n_0} - \frac{k_\perp^2 \nu_e^2}{\Omega_e^2} \right.
$$
$$
\cdot \left[\frac{\partial^2}{\partial t'^2} + \nu_i \left(1 + \frac{1}{\Psi_0} \right) \frac{\partial}{\partial t'} \right.
$$
$$
\left. + k_\perp^2 c_s^2 + \frac{i \nu_i}{\Psi_0} \boldsymbol{k}_\perp \cdot \boldsymbol{V}_d \right] \frac{\delta n}{n_0} = 0, \tag{1}
$$

where z is along the magnetic field direction and where

$$
\Psi_0 = \frac{\nu_e \nu_i}{\Omega_e \Omega_i}, \tag{2}
$$

$$
\frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + i\boldsymbol{k}_\perp \cdot \boldsymbol{V}_{i0} \tag{3}
$$

$$
\boldsymbol{V}_d = \boldsymbol{V}_{e0} - \boldsymbol{V}_{i0} \tag{4}
$$

As a check on the algebra, we note that if we Fourier analyze Eq. (1) in time and along the magnetic field in space, we recover the standard local dispersion relation (e.g. Kelley, 1989), namely

$$
\omega'(\omega' + i\nu_i) + i \frac{\nu_i}{\Psi} (\omega' - \boldsymbol{k}_\perp \cdot \boldsymbol{V}_d) - k^2 c_s^2 = 0, \tag{5}
$$

where

$$
\Psi = \Psi_0 \left(\frac{k_\perp^2}{k^2} + \frac{k_\parallel^2 \Omega_e^2}{k^2 \nu_e^2} \right) \tag{6}
$$

and where $\omega' = \omega - \boldsymbol{k}_\perp \cdot \boldsymbol{V}_{i0}$. This being stated, it should be noted that while the perturbed potential and perturbed densities are usually assumed to be proportional to one another, the equations for the perturbed potential and the perturbed densities are not quite the same because the high order derivatives do not commute. This means that in the first term in Eq. (1), the order of the time and space derivatives for the perturbed potential is the reverse of what we obtain for the perturbed density field. This might imply, in turn, that the evolution

of the potential and perturbed densities is not quite the same along the magnetic field direction.

As observed in the previous section, the problem with a plane wave decomposition in ω and $k_\parallel$ is that when the collision frequencies and $\boldsymbol{V}_d$ are changing in space, the frequency should be taking a value that changes with altitude. This is contrary to the very assumption used in obtaining an eigenvalue plane wave decomposition and constitutes an essential part of the so-called "nonlocal" problem. A decomposition into a superposition of plane waves was used by St.-Maurice (1985) to tackle this nonlocal problem, using a Laplace transform to solve the problem for impulses given at various altitudes. It was found that the nonlocal/convective character of the instability was important for waves greater than 10 m in size, and that the waves could reach a peak amplitude as they convected through the unstable region, without requiring a nonlinear saturation process. The approach used by St.-Maurice (1985) was unfortunately subject to numerical difficulties in that Laplace transforms are difficult to invert numerically and the results are difficult to interpret at times. Therefore, we have used a different approach here, which has the added advantage of shedding some new light on the physics at work.

Our method of solution is based on a WKB-style solution to the differential equation given by Eq. (1). We simply write that the perturbed density (and the perturbed potential for that matter) has a solution of the form

$$\frac{\delta n}{n_0} = A(z, t) \exp[i S(z, t)]. \tag{7}$$

However, in a departure from the standard WKB procedure, we prescribe that both A and S be real with the key, but reasonable assumption, that temporal and spatial derivatives of the amplitude A are much smaller than the temporal and spatial derivatives of the phase S. Therefore, we model growth not through an imaginary part of the phase, but rather through a time derivative of the amplitude itself. We also define the "frequency" ω_r and the "parallel wave vector" $k_\parallel$ in terms of local derivatives of the phase, viz.

$$\omega_r(z, t) = -\partial S(z, t)/\partial t$$
$$k_\parallel(z, t) = \partial S(z, t)/\partial z. \tag{8}$$

With these definitions and assumptions and if we keep all leading order real and imaginary terms (even though real and imaginary terms are not of the same order), we obtain

$$\frac{\partial}{\partial t'}\left(\frac{\delta n}{n_0}\right) = \left(\frac{\partial A}{\partial t} - i\omega_r' A\right) e^{iS} \tag{9a}$$

$$\frac{\partial^2}{\partial t'^2}\left(\frac{\delta n}{n_0}\right) \approx \left(-\omega_r'^2 A - \left\{2i\omega_r'\frac{\partial A}{\partial t} + i\frac{\partial \omega_r'}{\partial t'}A\right\}\right) e^{iS} \tag{9b}$$

$$\frac{\partial}{\partial z}\left(\frac{\delta n}{n_0}\right) = \left(\frac{\partial A}{\partial z} + ik_\parallel A\right) e^{iS} \tag{9c}$$

$$\frac{\partial^2}{\partial z^2}\left(\frac{\delta n}{n_0}\right) \approx \left(-k_\parallel^2 A + \left\{2ik_\parallel\frac{\partial A}{\partial z} + i\frac{\partial k_\parallel}{\partial z}A\right\}\right) e^{iS} \tag{9d}$$

$$\frac{\partial^2}{\partial z^2}\left[\frac{\partial}{\partial t'}\left(\frac{\delta n}{n_0}\right)\right] \approx \left(ik_\parallel^2\omega_r' A + 2k_\parallel\frac{\partial(\omega_r' A)}{\partial z}\right.$$
$$\left. + \omega_r'\frac{\partial k_\parallel}{\partial z}A - k_\parallel^2\frac{\partial A}{\partial t}\right) e^{iS} \tag{9e}$$

$$\frac{\partial^2}{\partial z^2}\left[\frac{\partial^2}{\partial t'^2}\left(\frac{\delta n}{n_0}\right)\right] \approx \left(k_\parallel^2\omega_r'^2 A + i\left\{2k_\parallel^2\omega_r'\frac{\partial A}{\partial t}\right.\right.$$
$$\left.\left. + k_\parallel^2\frac{\partial \omega_r'}{\partial t}A - 2k_\parallel\frac{\partial(A\omega_r'^2)}{\partial z} - \omega_r'^2\frac{\partial k_\parallel}{\partial z}A\right\}\right) e^{iS}, \tag{9f}$$

where we recall that $\partial/\partial t'$ is defined by Eq. (3).

3.1 Phase relationships

Equations for the local phase change, as well as for the amplitude, can be obtained from Eq. (1), using Eq. (9a) to (9f) if we assume that $\Psi_0 \ll 1$. After dividing by the term $\exp(iS)$, the leading order balance between the remaining imaginary terms then gives the equation for the frequency, whereas the balance between the real terms produces the amplitude equation. We also neglect temporal and/or spatial derivatives of the kind $|\partial \ln A/\partial t|$ when compared to $|\omega_r'|$: as a reference point, note that this means that none of the terms in the curly brackets in Eq. (9a) to (9f) end up being used in the process.

With the procedure just described, the leading order equation for the frequency becomes

$$\omega_r' = \frac{\boldsymbol{k} \cdot \boldsymbol{V}_d}{1 + \Psi} \tag{10}$$

i.e.

$$\omega_r = \frac{\boldsymbol{k} \cdot \boldsymbol{V}_d}{1 + \Psi} + \boldsymbol{k} \cdot \boldsymbol{V}_{i0}, \tag{11}$$

and we recall that Ψ is given by Eq. (6) once we take $k = k_\perp$.

While Eq. (11) may look like the standard expression, we remind the reader that there is an important difference with the usual result from the classical dispersion relation in that the frequency and parallel wave number are not assumed to be constant here. These variables are allowed instead to truly change with altitude. Our equations can be viewed as plane wave solutions only when altitude derivatives in Ψ and V_d can be neglected, which is not something that can be done at lower frequencies.

The above consideration also implies that, contrary to the belief rooted in the standard local theory, the parallel wave number is not a free parameter of the problem. Rather, there is a simple connection between the parallel wave number and the frequency, which is trivially based on their local definitions, as posited in Eq. (8). Namely, we must have

$$\frac{\partial k_\parallel}{\partial t} = -\frac{\partial \omega_r}{\partial z}. \tag{12}$$

This constitutes what is sometimes called the "Whitham relation". Its use allows us to tackle our nonlocal problem relatively simply by solving the differential Eq. (12), together with what in effect amounts to be a standard-looking root to the dispersion relation given by Eq. (11).

3.2 Amplitude equation

A second important difference with the standard procedure is the generalization of the simple growth rate expression through a term that includes a convective derivative of the wave amplitude and other quantities as well. As described above the new amplitude balance can be obtained for the case at hand by using the balance of the real terms left from using Eq. (9a) to (9f) into Eq. (1) after dividing by $\exp(iS)$. Alternatively, Eq. (1) can be multiplied by the complex conjugate of $\delta n/n_0$, and added to the equation for the complex conjugate of $\delta n/n_0$ after it has itself been multiplied by $\delta n/n_0$. All quantities involving fast time or spatial changes then cancel out, leaving us with the long time scale, slow spatial scale description.

Either one of these procedures brings the result

$$\frac{\partial A}{\partial t} + V_{g\parallel} \frac{\partial A}{\partial z} = \gamma A$$
$$- \frac{A}{2} \left[\frac{V_{g\parallel}}{k_\parallel} \frac{\partial k_\parallel}{\partial z} - \frac{2 V_{g\parallel}}{1+\Psi} \frac{\partial \Psi}{\partial z} + \frac{2 V_{g\parallel}}{\boldsymbol{k} \cdot \boldsymbol{V}_d} \frac{\partial \boldsymbol{k} \cdot \boldsymbol{V}_d}{\partial z} \right], \quad (13)$$

where

$$\gamma = \frac{(\omega_r'^2 - k^2 c_s^2)\Psi}{(1+\Psi)\nu_i} \quad (14)$$

is the standard classical growth rate. In Eq. (13) we have used the symbol $V_{g\parallel}$ to represent the expression

$$V_{g\parallel} = -\frac{\boldsymbol{k} \cdot \boldsymbol{V}_d}{(1+\Psi)^2} \Psi_0 \frac{2 k_\parallel}{k^2} \frac{\Omega_e^2}{\nu_e^2}. \quad (15)$$

As the symbol suggests, $V_{g\parallel}$ is the parallel component of the group velocity, as can be seen from using the definition

$$V_{g\parallel} = \frac{\partial \omega_r}{\partial k_\parallel}. \quad (16)$$

The fact that we have a convective derivative on the left hand side of the amplitude equation and that the first term on the right-hand side just describes the local growth rate, is very suggestive of a conservation principle being at work here. This is indeed the case, with the conservation principle being the conservation of wave action (e.g. Bretherton and Garret, 1969), of which conservation of energy is just a particular case. Conservation of wave-action states that, if the local frequency changes slowly enough in space and time, we should have

$$\frac{\partial}{\partial t} \left(\frac{U_k}{\omega_r'} \right) + \nabla \cdot \left(\frac{\boldsymbol{V}_g U_k}{\omega_r'} \right) = P, \quad (17)$$

where U_k is the energy of the waves. We have added a source term P to the textbook expressions to account for the sources/sinks of free energy that give/extract energy locally to/from the waves. For instance, in the classical (local) treatment of the Farley-Buneman instability, variations in ω_r' are neglected and a balance is assumed between the source term and the first term on the left-hand-side of Eq. (17).

Clearly from Eq. (17) we could derive an amplitude equation using the conservation of wave energy rather than wave action only if temporal and spatial variations in both $\boldsymbol{V}_d$ and Ψ were to be negligible, as seen from the expression for the frequency, Eq. (10). However, if nothing else, above 110 km, the ion drift does change rapidly with height as the ions start to acquire the $\boldsymbol{E} \times \boldsymbol{B}$ drift, due to their decreasing collision frequencies (see Appendix A for more details). This, according to Eq. (4), rapidly brings $\boldsymbol{V}_d$ to zero. The proper conservation law that we should use must, therefore, be the conservation of wave action and not the conservation of wave energy.

In order to show that Eqs. (13) and (17) are equivalent, we first need an expression for the energy of the waves for the problem at hand. In the ion frame of reference, the wave energy is given by the sum of the electron and ion kinetic energy and the electrostatic potential energy of the waves. Due to its small mass, it is easy to show in the present case that the electron kinetic energy is approximately Ψ_0 smaller than the ion kinetic energy, i.e. much smaller than the ion kinetic energy at most altitudes of interest here. This leaves us with

$$U_k \approx \frac{1}{2} n_0 m_i |\delta \boldsymbol{v}_k^i|^2 + \frac{\epsilon_0 |\delta E_k|^2}{2}$$
$$= \frac{1}{2} n_0 m_i |\delta \boldsymbol{v}_k^i|^2 + \frac{1}{2} n_0 m_i \frac{\Omega_i^2}{\omega_{pi}^2} \frac{|\delta E_k|^2}{B^2}. \quad (18)$$

The second term is normally considerably smaller than the first, while we can also write

$$|\delta \boldsymbol{v}_k^i| \approx \frac{e \delta E_k}{m_i \nu_i}. \quad (19)$$

We can use this expression in the first term of Eq. (18), keep only that first term, and plug the result in Eq. (17). Using γA for the source term P after the equation is expressed in terms of the amplitude of the fluctuations instead of the energy itself, we then obtain from conservation of wave-action the equation

$$\frac{\partial A}{\partial t} + V_{g\parallel} \frac{\partial A}{\partial z} = \gamma A - \frac{A}{2} \left[\frac{V_{g\parallel}}{k_\parallel} \frac{\partial k_\parallel}{\partial z} - \frac{2 V_{g\parallel}}{1+\Psi} \frac{\partial \Psi}{\partial z} \right.$$
$$\left. + \frac{(2+\Psi) V_{g\parallel}}{\boldsymbol{k} \cdot \boldsymbol{V}_d} \frac{\partial \boldsymbol{k} \cdot \boldsymbol{V}_d}{\partial z} + V_{g\parallel} \frac{\partial \ln n_0}{\partial z} \right]. \quad (20)$$

There are two small differences between Eqs. (20) and (13). First, the last term on the right-hand-side of Eq. (13) is $2/(2+\Psi)$ smaller than the same term in Eq. (20). However, this difference is not significant for an instability calculation, because the waves only grow when Ψ is small and, conversely, when Ψ becomes significant, the waves are quickly damped. The second difference is with the last term on the right-hand side of Eq. (20), which is simply not present in Eq. (13). The origin of that term is with n_0 in the energy density expression (18). It is quite difficult to assess from the algebra if this discrepancy between the two results is due to an inconsistency in the assumptions made in the derivation

of the dispersion relation (terms in $\partial \ln n_0 / \partial z$ were explicitly dropped from the algebra) or if it is linked to a violation of the "small spatial variations" requirement (Bretherton and Garret, 1969) in the expression used for the conservation of wave-action. Fortunately, the question is a bit academic in the present context because, within the current formulation, it would be inconsistent for the logarithmic derivative of the background ion density to be able to compete with the other derivatives in Eq. (20). This suggests that Eq. (13) is actually consistent with the approximations that were used to arrive at that point.

The main emphasis to draw from this analysis is that the amplitude equation is indeed an expression of the conservation of wave-action and could be derived on that basis rather than with brute-force algebraic techniques. This is important to note if one were to apply our methodology to other high-latitude irregularity problems. The connection with conservation of wave-action is also useful in order to recognize the physical principles behind an unfamiliar algebraic expression like Eq. (13).

3.3   Impact of the sign of $\mathbf{k}_\parallel$ on the evolution of the wave-trains

In our analysis we will assume that we have positive frequencies. This means from Eqs. (10) and/or (11) that $\mathbf{k} \cdot \mathbf{V}_d$ will be considered positive. According to Eq. (15), this in turn means that the group velocity will have to be negative if $k_\parallel$ is positive. To understand this in a physical context, one could return to Fig. 2 and picture what would happen if the top of the initial shoebox was tilted to the left instead of being aligned with the magnetic field. In that particular case, negative charges would be sitting on top of the positive charges (so that an upward parallel electric field would accompany a positive $k_\parallel$ with our sign convention). The larger electron mobility along the magnetic field at greater altitudes would then mean that the box would erode from the top, at least initially. This would correspond to a negative parallel component of the group velocity. The trend would tend to reverse with time because the shears in the phase velocity would steadily rebuild forward tilts. This would happen first at lower altitudes where these shears are more pronounced.

The effect of the shears on the evolution is confirmed in mathematical terms through Eq. (12), at least if the aspect angles are small. Equation (12) predicts that for positive frequencies that increase with height, $k_\parallel$ decreases monotonically with time, since ω_r increases monotonically with altitude. This is indeed consistent with the evolution of the shoebox that we just presented.

Finally, it is easy to see that the sign of $k_\parallel$ has at least the potential to affect the amplitude results. In particular, if $k_\parallel$ is initially chosen to be positive, it will tend to reverse its sign, and $V_{g\parallel}$ will follow. This sign reversal will affect the propagation of waves in potentially interesting ways due to the influence of $V_{g\parallel}$ on the amplitude evolution through Eq. (13).

4   Solving for the phase and amplitude of nonlocal modes

4.1   Aspect angle as a nondimensional parameter

Before obtaining solutions for the system at hand, it is important to point out that for a fixed flow angle direction (a fixed direction of $\mathbf{k}_\perp$), the evolution equation for the aspect angle $\theta = k_\parallel / k_\perp$ is independent of both components of the wave vector. This can be seen by combining Eq. (12) with (11). We then obtain

$$\frac{\partial \theta}{\partial t} = -\frac{\partial}{\partial z}\left[\frac{\hat{\mathbf{k}} \cdot \mathbf{V}_d}{1 + \Psi} + \hat{\mathbf{k}} \cdot \mathbf{V}_{i0}\right],\tag{21}$$

where $\hat{\mathbf{k}} = \mathbf{k}/k$ and Ψ now takes the form

$$\Psi = \Psi_0\left[1 + \frac{\Omega_e^2}{\nu_e^2}\theta^2\right].\tag{22}$$

From now on, we will deal with the aspect angle θ instead of the parallel wave vector itself. The main advantage is that the aspect angle equation is the same for all magnitudes of the wave number k.

4.2   Finding the characteristics of the waves

While we have already shown that the aspect angle (i.e. the parallel wave vector) has to change with time if the altitude dependence of the frequency cannot be neglected, one advantage of the WKB approximation is that wave propagation and wave growth/decay can be described as separate processes. Clearly from the WKB ordering, the amplitude depends on the aspect angle (see Eq. 13), while the aspect angle evolution does not depend on the amplitude (as can be seen from Eqs. 21 and 13). Therefore, finding the evolution of the waves is simply a matter of solving the aspect angle equation first, and then solving for the amplitude in terms of the aspect angle. However, the aspect angle evolution equation is also clearly nonlinear. Therefore, even though the amplitude equation itself remains linear, nonlinear effects can affect its evolution through the aspect angle behavior. As we show below, the fact that the aspect angle equation is nonlinear has more profound implications for the amplitude evolution than could have been guessed at first, because it often leads to the formation of shocks.

We have explored various ways to solve the nonlinear aspect angle equation. At first we tried finite difference types of schemes, but found that there was a tendency for very large slopes to evolve in the aspect angle. The origin of this steepening in the slope was not clear but it required introducing artificial diffusion in the solver. In turn, this raised some questions about the validity of the solution. In order to gain more insights into what was really going on, we switched to the more analytical method of characteristics. This method has several advantages, as we now hope to show.

The method of characteristics can deal efficiently with the nonlinearity present in Eq. (21), through the Ψ factor. It turns

out that with this method, we can follow not just the propagation paths of a set of an initial aspect angles, but also ("accidentally") the amplitude propagation paths as well.

To start with, we must focus on the evolution of the aspect angle and parallel group velocity. The first step is to find the characteristic paths of the waves. Fortunately, neither aspect angle nor the parallel group velocity depend on the amplitude, so that we can actually go through with this approach. Finally, it should be pointed out that for an inhomogeneous medium, the ray tracing associated with the method of characteristics is associated with lines of constant frequencies rather than lines of constant wave numbers or phase velocities, as would be the case for homogeneous media (e.g. Kundu, 1990, Chapter 7). This is because the rays are following the group velocity, which is the path of propagation of energy. In the present case, a quick glance at our expression for the group velocity shows that it can have directions that are anywhere from parallel to the phase velocity, to practically perpendicular to the same phase velocity.

The first step in finding the wave propagation is to expand the z derivative in Eq. (12) in the form

$$\frac{\partial \omega_r}{\partial z} = \left(\frac{\partial \omega_r}{\partial z}\right)_{k_\parallel} + \frac{\partial \omega_r}{\partial k_\parallel}\frac{\partial k_\parallel}{\partial z}. \tag{23}$$

The notation on the right-hand side means taking the partial derivative of ω_r with respect to z, while keeping $k_\parallel$ constant. Since ω_r depends on z both explicitly through Ψ_0, as well as implicitly through $k_\parallel$ (Eqs. 11 and 6), this notation is necessary to distinguish how we take the partial derivative, i.e. whether we allow $k_\parallel$ to change or not. The plain $\partial \omega_r/\partial z$ will mean the full derivative, including both explicit and implicit dependencies on z.

For the method of characteristics, we start with Eq. (12), write Eq. (23) in terms of θ and $\omega_r/k_\perp$, and then bring all θ derivatives to the left-hand side of Eq. (12) to end up with

$$\frac{\partial \theta}{\partial t} + V_{g\parallel}\frac{\partial \theta}{\partial z} = -\left(\frac{\partial(\omega_r/k_\perp)}{\partial z}\right)_\theta. \tag{24}$$

This way, the right-hand side is independent of $\partial\theta/\partial z$. For example, under the assumption that the ratio of the collision frequencies ν_i/ν_e can be considered constant with altitude in the unstable region, the right-hand-side of Eq. (24) would become

$$-\left(\frac{\partial(\omega_r/k)}{\partial z}\right)_\theta = \frac{\hat{\mathbf{k}} \cdot \mathbf{V}_d}{1 + \Psi}$$

$$\times \left\{\frac{1}{1+\Psi}\frac{\partial \Psi_0}{\partial z} - \frac{\Psi}{1 + \nu_i^2/\Omega_i^2}\frac{\partial \ln(\nu_i^2/\Omega_i^2)}{\partial z}\right\}. \tag{25}$$

With the derivatives of θ fully accounted for on the left-hand-side of Eq. (24), we can now use the method of characteristics to solve for θ. This means solving along a characteristic path i defined by

$$\frac{dz_i(t)}{dt} = V_{g\parallel}\big(z_i(t), \theta_i(t)\big), \tag{26}$$

where we have used the definition given by Eq. (16). Along these characteristics, the equation for the aspect angle evolution becomes

$$\frac{d\theta_i(t)}{dt} = f\big(z_i(t), \theta_i(t)\big), \tag{27}$$

where $f(z, \theta)$ is actually the right-hand-side of Eq. (24) and to a good approximation is given by Eq. (25). In effect, the last two equations are turning the original partial differential equation into a set of two coupled nonlinear ordinary differential equations. The resulting set can easily be solved using an explicit ODE solver, allowing us to find the evolution of the aspect angle. One added advantage of using this method of solution is that there is no need for boundary conditions, in the sense that we only need an initial condition $\{z_i(0), \theta_i(0)\}$ for each of the characteristics $z_i(t)$.

Interestingly enough, the solution to the amplitude Eq. (13) is also easy to obtain within the framework that we have just described, because both aspect angle and amplitude have the same group velocity. Therefore, the amplitude equation can be written in terms of characteristic paths that are identical to the aspect angle ones. Along a $k_\parallel$ characteristic we then have

$$\frac{d \ln A_i(t)}{dt} = \gamma_{eff}, \tag{28}$$

where γ_{eff} is the right-hand side of Eq. (13) divided by A. This effective growth rate includes both local Farley-Buneman growth and convective growth. One important difference between the aspect angle and amplitude equations, however, is that while the aspect angle equation is independent of frequency or wave number, the amplitude equation is not. The latter is due solely to the local classical expressions for Farley-Buneman growth/decay on the right-hand side of Eqs. (13) or (20), and not to wave-action conservation per se. In practice, this means that for a given set of ionospheric conditions, one only needs to solve for the aspect angle once for all possible wave numbers, whereas the amplitude evolution will be different for each wavelength that we wish to consider.

4.3 The handling of shocks associated with intersecting characteristics

We have found that the origin of the large aspect angle derivatives uncovered with other numerical methods was actually linked to the fact that in many cases, characteristics end up intersecting one another if we follow the evolution of the system for long enough periods of time. We now discuss this important issue at length, since it requires taking some action to "fix" the problems this feature entails.

We first provide a specific example of what we mean with Fig. 3. In this case, we calculated the characteristics for an initial aspect angle of zero everywhere, using a 50 mV/m electric field and the modeling parameters described in Appendix A. Figure 3 shows the characteristic paths $z_i(t)$ for various initial positions $z_i(0)$ between 90 and 125 km. After
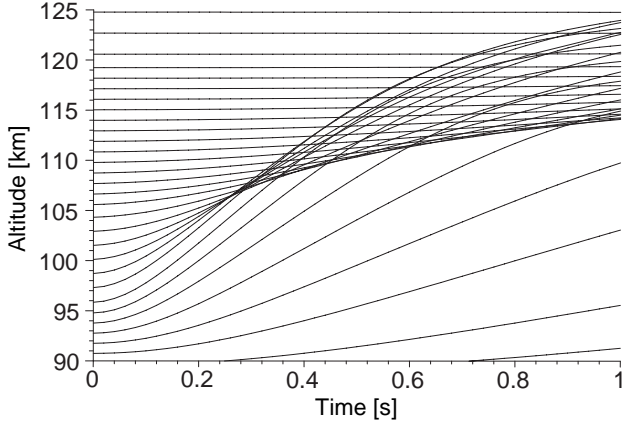
Fig. 3. Characteristic paths $z_i(t)$ obtained by solving the coupled system given by Eqs. (26) and (27) for a 50 mV/m electric field under the assumption of an initially perfectly field-aligned structure.

Fig. 4. Phase fronts calculated from the crossing characteristics of Fig. 3 at intervals of 0.1 seconds. This figure shows that some characteristics are associated with multiple phase values: for example $S(117\,\text{km}, 0.8\,\text{s})$ has values of 395 rad, 415 rad and 425 rad all at the same time and place.

about 0.3 seconds, the paths of several characteristics intersect, for reasons we will explain below. We just note for now that since each of the characteristics propagates information about the aspect angle and amplitude, the implication is that from the point of intersection and beyond, the solution has multiple values for θ and A, and as a result, multiple solutions for ω and $V_{g\parallel}$ as well. We now show that such crossings have to be discarded, and we describe the precise algorithm that allows us to determine which one of the multiple values must be retained.

To understand why characteristics often intersect if left to their own devices, we should first go back to Figs. 1 and 2. These figures illustrate that even if we started with zero aspect angles everywhere (null parallel wave-vector components), the system would quickly evolve nonzero aspect angles, $k_\parallel/k_\perp$, through the systematic increase in ω_r with altitude.

The second element to consider is that the characteristics are actually curves of constant frequency. To show this, first notice that we can write the parallel group velocity in the form

$$V_{g\parallel} = \frac{\partial \omega_r}{\partial k_\parallel} = \frac{\partial \omega_r}{\partial t} \bigg/ \frac{\partial k_\parallel}{\partial t} = -\frac{\partial \omega_r}{\partial t} \bigg/ \frac{\partial \omega_r}{\partial z}. \tag{29}$$

The second equality in the above equation uses the fact that ω_r only depends on time through $k_\parallel$ alone, while for the final equality, we have substituted Eq. (12). Using the last expression for $V_{g\parallel}$, we then find that along a characteristic,

$$\begin{aligned} \frac{d\omega_r}{dt} &= \frac{\partial \omega_r}{\partial t} + V_{g\parallel} \frac{\partial \omega_r}{\partial z} \\ &= \frac{\partial \omega_r}{\partial t} + \left(-\frac{\partial \omega_r}{\partial t} \bigg/ \frac{\partial \omega_r}{\partial z} \right) \frac{\partial \omega_r}{\partial z} = 0. \end{aligned} \tag{30}$$

We can now understand why characteristics have a tendency to intersect. Consider, for example, an initial condition for which the aspect angle is zero everywhere. In that case, the lower altitudes have smaller phase velocities in the
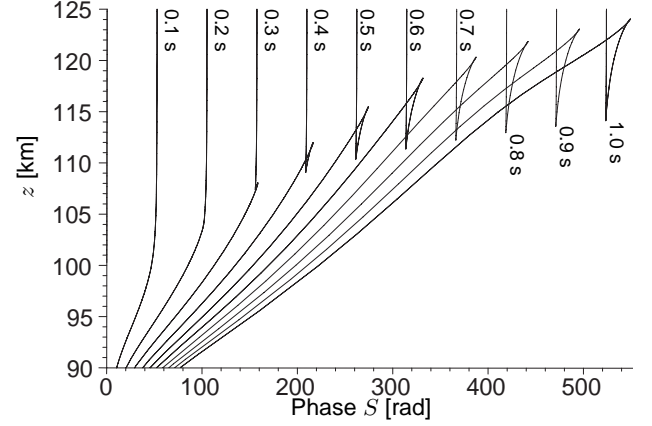
perpendicular direction, simply due to their larger collision frequencies. As a result, however, the aspect angle must increase with time at these lower altitudes (see again Fig. 2). According to Eqs. (11) and (6), this causes an increase in Ψ as well, which implies a decrease in ω_r for a given altitude. But since Eq. (30) requires ω_r not to change, the only solution is to counter the increase in Ψ created by the increase in $k_\parallel$ with a decrease in Ψ_0. This makes it necessary for the characteristic to move to a higher altitude. In other words, the evolution of the aspect angle forces the characteristics to move upwards, consistent with the presence of an upward group velocity. In the example discussed here, the characteristics intersect simply because the characteristics originating at lower altitudes run into characteristics from higher altitudes that have not themselves been moving appreciably upwards due to their much slower aspect evolution.

To understand what must be done with intersecting characteristics, we now need to introduce the phase so as to graphically illustrate how it evolves when characteristics intersect one another. We can easily calculate the phase S from a knowledge of the aspect angle evolution using Eq. (8). Along a characteristic, we simply have

$$\frac{dS}{dt} = \frac{\partial S}{\partial t} + V_{g\parallel} \frac{\partial S}{\partial z} = -\omega_r + V_{g\parallel} k_\parallel. \tag{31}$$

Using Eq. (31) we can now connect lines of equal phases (phase fronts). For the calculations whose results are displayed in Fig. 3 we have plotted in Fig. 4 the phase as a function of altitude for various selected times. Figure 4 shows that, initially, the only effect is a phase front that moves faster at higher altitudes. At later times, however, if we let the system evolve according to our nonlinear characteristic equations, we find that the lines of equal phases start to form a triangular sort of loop. The loops begin at intersection points of lines of equal phase. At these intersection points there is a jump in the derivative of the phase with respect to altitude,

which actually means a jump in $k_\parallel$. Furthermore, beyond that point, that is, once the curves cross, the phase becomes multi-valued and can no longer be thought of as a proper function of altitude. This indicates that something is wrong either with the solution or with the equations themselves.

The hint for what went wrong and how to remedy it comes from noticing that just prior to the crossing, when the parallel wave vector (and the associated parallel group velocity) is about to jump, its altitude derivative (second altitude derivative of the phase) is becoming very large. If we go back to the very beginning of the theory with the original model equation, this gives us a clue as to how to proceed. In other words, consider Eq. (1) again, in the case of very large vertical gradients in the parallel wave vector. When the parallel wave number is about to jump, the second derivative in z has to completely overwhelm the rest of the equation, unless the operator that it acts on is itself zero, or very close to zero. From this we can see that the balance of the equation is no longer a Farley-Buneman wave, but rather a strongly collisionally damped ion-acoustic wave (if we have a wave at all) described by the equation

$$\left(\frac{\partial^2}{\partial t'^2} + \nu_i \frac{\partial}{\partial t'} + k_\perp^2 c_s^2 \right) \frac{\delta n}{n_0} = 0. \quad (32)$$

The implication is that as the characteristics come close to crossing and $\partial^2/\partial z^2$ becomes large (this is labeled a "shock" in the pertinent literature; e.g. Logan, 1994), the waves couple into strongly damped ion-acoustic waves. The characteristic damping time is $1/\nu_i$, and has a typical value of a few milliseconds. For the long wavelengths of concern in the present work, this time could, therefore, be considerably shorter than $1/kc_s$. Therefore, in effect, as soon as two characteristics are about to collide, we have to terminate the evolution of both the fast and slow parts of the wave train, given the immediate damping the waves suffer at that point.

In spite of the reasonable mathematical understanding, the physical interpretation behind the crossing paths and the subsequent conversion to heavily damped modes is not entirely clear to us. We can focus either on the fact that just near the moment and place when and where some characteristics intersect, the wave number (altitude derivative of the phase) is about to jump in space, or we can say that there is a very fast change in the frequency (time derivative of phase) with time. Using the second viewpoint we can say that near the intersecting region, low frequency waves arriving from lower altitudes ram into higher frequency waves triggered higher up. With the phase velocity being horizontal (that is, the phase fronts moving horizontally), one can see that the phase and any wave-like structure it represents have to be destroyed near the characteristics' intersection points. The fact that under these conditions the only possible wave-like solution is represented by heavily damped ion-acoustic waves (if the ion-collision frequency is too large, there might not even be oscillations) suggests that some form of mandatory mode coupling may be taking place under these circumstances.

No matter what the physical interpretation might be for shock formation, we had to adopt the following procedure in order to avoid unphysical results associated with the limitations of the model near intersecting characteristics singularities: at the crossing points, given that all waves are turned into heavily damped modes, the energy of any structure associated with the colliding characteristic paths was removed from the structures (implicit in this choice is the fact that the wave energy had to be given back to particles through some heating of the plasma; however, we did not pursue the plasma heating aspects of this problem). This procedure was particularly easy to follow, given that the amplitudes can be described in terms of the same characteristic paths as the aspect angle itself. A second part of the procedure was that in order to be able to follow the characteristics of the part of waves that were unaffected by the shock condition, we also imposed that the phase remain continuous, as well as single-valued, and we joined the lines of unequal slopes without going through a loop. This choice ensured that we could continue to describe self-consistently the behavior of the waves on each side of the shocks.

We have also uncovered conditions for which the characteristics jumped instead of crossing over. In such cases, the arguments regarding the fate of the waves and the continuity of trajectories over the shocks are similar to those we have presented here for intersecting characteristics, even though some details obviously differ. Since the procedure we used in such cases involves, in many ways, more of the same arguments presented in the present section, we have simply outlined its main features in Appendix B.

4.4 Initial conditions

For the method of characteristics, initial conditions are simple to take care of. We must simply specify a starting value of the function we are solving for each characteristic we wish to study. For a specific starting altitude, this amounts to specifying the initial starting aspect angle if we are solving Eq. (27), and the initial amplitude if we are solving the amplitude equation given by Eq. (28). Of course, the real physical initial condition is on the initial density perturbation $(\delta n/n_0)$ (0), which must be described in terms of an initial amplitude A_0 and an initial phase S_0. In order not to violate the assumption that the derivatives of the phase are larger than those of the amplitude, we chose to describe all initial perturbations in terms of an initial change of phase with altitude, and selected a uniform A_0 as our initial condition.

Given that the amplitude part of the problem is linear, all our amplitude solutions could be expressed in terms of an amplitude ratio $A_i(t)/A_i(0)$. Our solutions will, therefore, be presented in terms of this ratio, rather than in terms of the amplitude itself. Since, to start with, we chose an initial amplitude that was uniformly distributed, the initial conditions on the new unknown, $A_i(t)/A_i(0)$ have, therefore, been described as being equal to unity at all altitudes.

We studied two types of initial conditions. For the first type, we modeled the real part of the initial density pertur-

bation with an impulse described by a Gaussian of a given width and peak height using the relation

$$\frac{\delta n}{n_0}(0)/A_0 = \cos S(0) = \exp\left[-\left(\frac{z - z_0}{H}\right)^2\right]. \qquad (33)$$

Solving for the phase $S(0)$, and taking $k_\parallel(z, 0) = \partial S/\partial z(0)$, this meant having

$$k_\parallel(z, 0) = \pm\frac{2(z - z_0)}{H^2}\frac{\cos S(0)}{\left[1 - \cos^2 S(0)\right]^{1/2}}, \qquad (34)$$

where we picked one sign or the other according to the particular run being chosen, while H described the Gaussian width and z_0 described the altitude of the peak. For many of our calculations, we chose $H = 5$ km and $z_0 = 95$ km, which translates into a peak value for the initial $k_\parallel$ values of $\pm 2.828 \times 10^{-4} \mathrm{m}^{-1}$. As we shall see below, this latter choice turned out to produce results that are essentially equivalent to a zero initial aspect angle condition.

For the second type of initial phase conditions, we imposed a fixed "tilt" to the perturbation, i.e. a uniform, nonzero aspect angle. For these latter runs, we did not look at the combined effect of pulses and tilts, since the list of possibilities can in fact become endless. Our present goal is physical insights, rather than a study of all possible cases, and the two types of initial conditions seem to be sufficient to meet this goal.

Specifying the initial phases was only one part of the parameters making up the initial conditions. Most importantly, we also had to vary the electron drift velocity, i.e. the dc electric fields. It was also necessary to change the wavelength since the absolute growth rate of the waves is affected by the wave number for a given electric field strength. Therefore, we studied the response of the system to wavelengths in the range of 12 to 50 m.

5 Results

5.1 Benchmark case: 12 m Farley-Buneman waves with 50 mV/m field

We have chosen 12 m waves as our benchmark to explore nonlocal effects. As described in the Introduction, this kind of wavelength is now probed by several HF radars at high latitudes, through the SuperDARN experiment. In order to have non-negligible growth rates for this case, we have used a fairly strong electric field, of the order of 50 mV/m (1000 m/s $\boldsymbol{E} \times \boldsymbol{B}$ drift). For weaker fields, destabilization through the gradient-drift mechanism would have to be considered. We are leaving a study of this process for the future, however.

Having chosen an $\boldsymbol{E} \times \boldsymbol{B}$ drift velocity equal to 1000 m/s, we then selected $T_e = T_n$ (electron heating is known to become important only above 50 mV/m). We have also used standard ion frictional heating theory to produce a temperature given by $T_i = T_n + (m_n V_i^2)/(3k_B)$. The initial conditions were those discussed in Sect. 4.4 in relation to Eq. (34).

The maximum aspect angle was, therefore, 0.0310 degrees initially. In all calculations, the $\boldsymbol{k}_\perp$ direction was held fixed and pointing towards the $\boldsymbol{E} \times \boldsymbol{B}$ drift.

In Fig. 5 we display image plots of the results of our benchmark run calculations through six panels. All panels have contours superimposed, with the contour levels indicated by the corresponding horizontal lines on the color bar.

Figure 5a shows how the aspect angle evolves as a function of altitude (y axis) and time (x axis). Initially, the angle is almost zero, as prescribed by our choice of initial conditions. It then grows with time to reach a value of several degrees. The exception is above 120 km, where the aspect angle does not evolve at all. This is because in that case, $\partial k_\parallel/\partial t = -\partial \omega_r/\partial z \approx 0$.

Figure 5b displays the phase S, obtained by integrating Eq. (31) in time along the characteristics. In this panel, the contours are equivalent to wave fronts, because they indicate curves of constant phase. This clearly shows how the wave train curves upward due to the aspect angle becoming more negative. In addition, a thin black line indicates the location where the characteristics would cross, but are instead terminated, as discussed in Sect. 4. This happens in such a way that the phase remains continuous across that region (the appearance of steps being solely due to the limitations of the plotting package), so that Eq. (12) remains valid.

Figures 5c and 5d show the phase speeds in the neutral and ion frames of reference, respectively. In Fig. 5c, we present the phase speed seen from the ground, ω_r/k. This speed increases monotonically with altitude (see Eqs. 11, A5 and A4), starting with a value of zero below 90 km, where Ψ exceeds 10, and ending with E/B above 120 km, where Ψ is zero for all practical purposes. Similarly, Fig. 5d shows the phase speed ω_r'/k, that is, the phase speed seen from the ion frame of reference. This variable also starts at zero, but reaches a maximum at 105 km and then decreases again, because the ions steadily accelerate toward the $\boldsymbol{E} \times \boldsymbol{B}$ drift above 105 km, so that the relative electron-ion speed V_d decreases above that height. We recall that ω_r is important for the wave propagation parameters like the phase and group velocities, whereas ω_r' determines the local growth or decay of the wave. Since we expect the largest amplitudes where the growth rate is zero, this means that these would occur where $\omega_r' = kc_s$. However, ω_r' is measured in the ion frame, so that the speed with respect to the ground would be larger, with the exact amount depending on the height.

Next, Fig. 5e shows the effective growth rate, γ_{eff}, defined in Eq. (28). In this panel, as well as in panel (f), white areas are in regions for which the growth rates are too negative to bother showing them. Notice how there is almost no growth in regions where the aspect angle is either very small or very large. The reason is that the growth rate from Eq. (14) is a dominant contribution and has a peak between 0.5 and 1 degrees. For smaller values the factor Ψ in the numerator is too small, while for large values the denominator is too large, and, more importantly, ω_r' becomes too small.

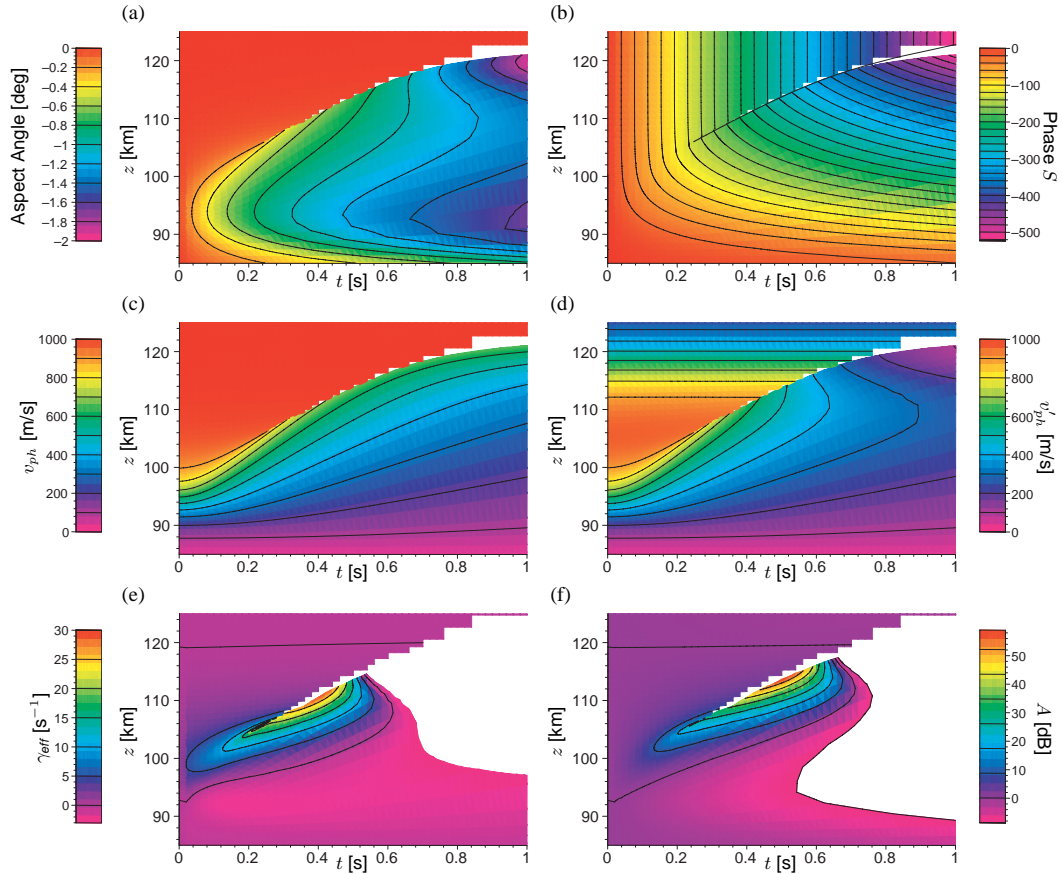Finally, by integrating the effective growth rate along a

Fig. 5. Results from the benchmark run at $\lambda = 12$ m, $V_d = 1000$ m/s, and $T_e = T_n$. Contour levels are given by the corresponding horizontal lines on the color bar. The white spaces associated with a step function-like staircase are produced when characteristics are being terminated to avoid crossing. The limitations of the plotting package are responsible for the dented look. The rest of the white areas associated with growth rate plots are in regions for which the growth rate is too negative to matter.

characteristic, we have obtained the amplitude, which is shown in Fig. 5f. In this panel we have used a logarithmic scale, with 0 dB representing the initial amplitude of 1. Figure 5f illustrates that while the wave train travels through the unstable region and grows initially, it also starts decaying when it reaches the upper altitudes, where the aspect angles become so large that local growth is replaced by decay. Additionally, the upper fraction of the wave train "crashes" into parts of the wave that have not evolved at all, and where $k_{\parallel}$ is, therefore, still zero, so that the vertical group velocity is also zero as well. As explained in Sect. 4.3, when these two parts of the wave train are about to interact, the oscillations suddenly decay and disappear from the region.

A final parameter retrieved from our calculations is the parallel group velocity. For our benchmark case, it can be as large as 40 km/s (not shown) for our electron drift velocity of 1000 m/s. Furthermore, the largest group velocities are located at lower altitudes, as expected from earlier discussions. Also, as mentioned earlier, our group velocities confirm that convective processes would lead to a lifetime of, at most, a few seconds for our 12 m wave train (20 km/s vertical speeds in a region about 20 km thick). This confirms our earlier

inference that in order for the amplitudes to be themselves involved in nonlinear processes (not the case in the present work), the unstable waves must have growth rates well in excess of 10 s^{-1}.

Our calculations can also be used to infer some of the spectral properties that would be observed with radars or even rockets, since one added benefit of our approach is that we know the frequency of the wave train at any point in time and space. This makes it easy to find the amplitude in frequency space, by plotting, for example, $A(z, \omega_r(z, t))$.

Various forms of amplitude computations can be made. In Fig. 6a we show $A(z, \omega_r'/k)$, the amplitude as a function of altitude, and the Doppler-shifted frequency ω_r', in terms of a phase speed v_{ph}'. This phase speed in the ion frame of reference determines the growth or decay of the wave, depending on whether it is larger or smaller than the local ion-acoustic speed c_s. From this, one would expect to see the largest amplitudes right at the threshold condition where $\omega_r'/k = c_s$, and Fig. 6a does in fact show a maximum near 500 m/s, a typical value for c_s. However, the ion-acoustic speed is a function of altitude, due to the change in ion and electron temperatures. Therefore, a clearer picture is obtained by
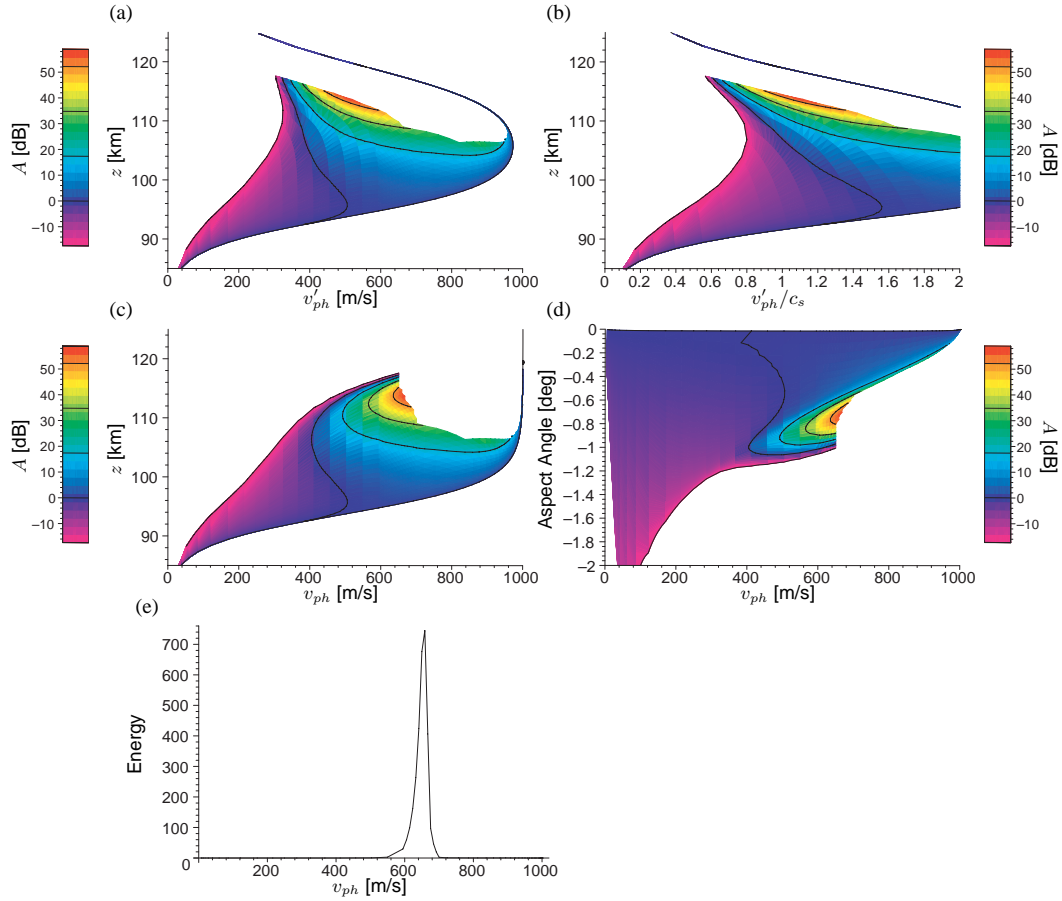
Fig. 6. Frequency and aspect angle dependence of the amplitude (panels a to d) and energy (panel e) for the benchmark run shown in Fig. 5.

comparing ω_r'/k and c_s directly; this is done in Fig. 6b. Here, the horizontal axis has been replaced by ω_r'/kc_s, meaning that a value of 1 is equivalent to the threshold between growth and decay. This panel also agrees with the notion that the largest amplitude should be found near the threshold conditions.

However, from an observational point of view, the value of ω_r' is not meaningful, because it would be quite difficult to attempt measurements in the ion frame of reference. A more useful number is the frequency in the neutral frame of reference, ω_r, which, in our calculations, has been assumed to be at rest with respect to the ground. Directly related to ω_r is the phase velocity of the wave, $v_{ph} = \omega_r/k$, shown in Fig. 6c. Here, the maximum amplitude is no longer close to c_s, but rather is between 650 and 700 m/s. This is due to the ion motion, which gradually increases with increasing height. As a result, a ground-based observer would see the largest amplitudes not at c_s, but rather at a higher speed of $v_{ph} = c_s + \hat{\mathbf{k}} \cdot \mathbf{v}_i$. Obviously, at least for the flow direction that we have selected (along the electron $\mathbf{E}_0 \times \mathbf{B}$ direction), the difference between Fig. 6a and Fig. 6c shows that the change in speed can be substantial.

Next, we show in Figure 6d how the amplitude varies as a function of phase speed and aspect angle. The maximum

amplitude occurs at an angle of about 0.8 degrees. Here, the correlation between aspect angle and phase speed becomes obvious, since an increase in aspect angle leads to a decrease in the phase speed through an increase in Ψ.

Finally, in Fig. 6e we show the energy integrated over all aspect angles as a function of phase speed. We have obtained this figure by summing $A^2\Delta\theta$ along each characteristic, where $\Delta\theta$ is the change in θ, which varies from one time step to the next. The result is a useful value, because the frequency ω is constant along a characteristic, as proven in Eq. (30). Therefore, the sum over $A^2\Delta\theta$ is a measure of the total energy for all aspect angles at one frequency, or equivalently, at one phase speed. Essentially, this is equivalent to taking the integral of Fig. 6d over all aspect angles at a particular phase speed, except that the energy is proportional to A^2, not A, and is shown here using a linear scale.

As with the previous plots, the energy reaches a maximum between 650 and 700 m/s. The energy distribution also exhibits a certain skewness as well, which is not a result of growth or decay as such, but rather of the shock that occurs when two parts of a wave train collide with one another. The resulting cutoff is visible in all panels as it leaves open contour regions. This happens at altitudes above 110 km and phase speeds greater than 650 m/s. Since the amplitude im-

mediately drops to zero at that point, the higher frequencies are cut off, leading to the skewness seen in Fig. 6e. This result means that the amplitudes are largest for the part of the wave that just barely misses the shock. For just slightly higher phase speeds, that part of the wave train will hit the shock and be removed from the wave train before it can grow to larger amplitudes.

The benchmark case that we have just described is one for which the local Farley-Buneman growth rate was deliberately chosen to be fairly large, so that local growth affected the wave train evolution in a measurable way and yet was small enough to allow convection to affect the amplitudes. This being said, convective effects must always impact the final stages of the evolution, because the advective terms on the right-hand-side of Eq. (13) become highly competitive when the aspect angle reaches a degree or more. Quite independently from this, however, the nonlocal treatment plays an additional important role in our benchmark run because the wave trains do not grow enough while they propagate through the E-region as a result of the convective term $V_{g\parallel}\partial A/\partial z$. This restricts the amplitude gains in the wave train by limiting the time during which growth can occur. The formation of a shock puts a further limit on the amplitude gain. To assess the interplay between these various processes, we now examine other cases through which we can study the relative importance of the various terms.

5.2 Results from other case studies

5.2.1 Longer wavelengths

Figure 7 shows growth and amplitude results for a wavelength of 50 m, again for a 50 mV/m field. Since the aspect angle, group velocity and phase speeds are independent of the wavelength, the behavior of these parameters are the same as that shown in the first four panels of Fig. 5. These plots are, therefore, not repeated here. The single difference between 12 m and 50 m waves has to do with the growth rate, which is shown in Fig. 7a and should be compared with Fig. 5e (as with other similar figures, the white areas in panel *a* are regions for which the growth rates are too small and negative). For 50 m waves, the Farley-Buneman growth rate, γ_{FB}, is zero or negative almost everywhere, except for a small region in front of the shock. Even when positive, γ_{FB} is much smaller than in the 12 m case. Consequently, the advective terms are now everywhere more important than local Farley-Buneman growth, and the waves cannot grow by much, even with a small aspect angle. A comparison between Fig. 7b and Fig. 5f makes that point clearly enough. Not even the growth seen just before the shock is significant, since it disappears almost instantaneously in the shock. We conclude that even a 50 mV/m field would not be large enough to support the growth of 50 m waves of any significant amplitude without a gradient-drift mechanism.

In order for 50 m waves to grow significantly with a Farley-Buneman mechanism, a larger field is necessary. For this reason, we present for our next case in Fig. 8 the results from a 100 mV/m field with a 50 m wave train.

Figure 8a shows that the growth rate is similar to the previous case for the part right in front of the shock. It also shows a much larger Farley-Buneman growth phase below the shock region. This leads to a significant amplitude gain between 0.2 and 0.4 s in Fig. 8b. The gain is, however, still less than in the 12 m waves and 50 mV/m case. Comparing Fig. 8c with Fig. 6b also shows a broader distribution of amplitudes over a larger range of velocities for the 50 m, 100 mV/m case. However, the maximum is still near the ion-acoustic speed. The broadening is a result of a larger range of c_s values near the shock, because the ion temperature profile is increasing rapidly with altitude due to additional frictional heating induced by the larger electric field.

Finally, as was found for the benchmark case, Fig. 8d shows that even though the amplitude is maximal at the ion-acoustic speed in the ion frame of reference, moving to the neutral frame of reference shifts the peak significantly, as it is now found at roughly 1200 m/s. The peak is also much broader than in the benchmark case, because there is a larger range of ion-acoustic speeds at various heights. The skewness is much more pronounced as well, with a very sharp cutoff at higher phase speeds, which occurs again at the point where part of the wave just misses the shock. Lower phase speeds can grow for as long as the growth rate permits, whereas higher phase speeds collide with the shock and disappear instantaneously. The additional bumps at 1800 m/s and 1900 m/s are from the initial growth in front of the shock.

5.2.2 Variations in initial conditions

So far in our presentation we have used an initial condition associated with an initial zero aspect angle (no tilt) superimposed with a small aspect angle "pulse" excited from an altitude region 5 km wide and centered at 95 km. As discussed earlier, this meant very small initial aspect angles everywhere, even for a 50 m wavelength.

We have studied the effect of changing the initial phase pulse on the evolution of the wave trains by varying the altitude, sign and magnitude of the initial impulse. We of course restricted the range of aspect angles to values small enough for the wave trains to be able to grow in the first place. In all cases studied we have found that the end result for the amplitude was very similar to the near-zero aspect angle initial condition runs. The reason for this lack of sensitivity is associated with the development of shocks, which is new to this work. In general, subsets of wave trains that initially move upward develop their own shocks before reaching significant amplitudes. When, by contrast, parts of the wave trains are made to move downward, they either collide with upward moving trains or create their own shocks on their way back up after reflection. Even when our initial conditions were fine-tuned to avoid early shocks, the change in the final wave amplitudes were not large compared to other pulsed phase runs; the reason for this is that in order to avoid shock conditions we had to use initial aspect angles that were too small
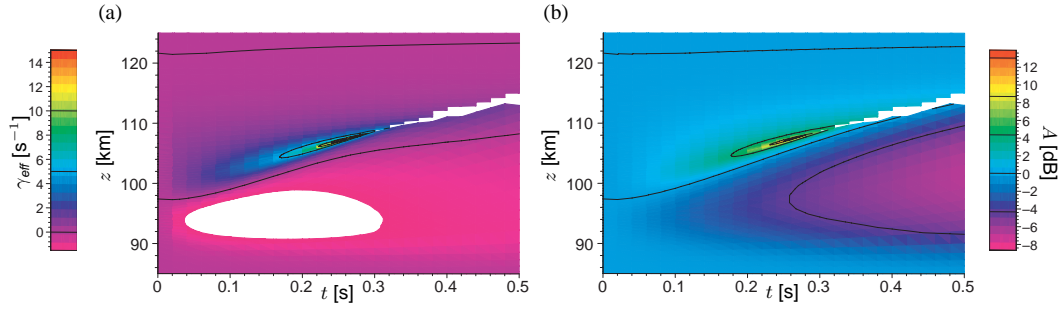
Fig. 7. Results for the same parameters as in Fig. 5 but with $\lambda = 50$ m instead of 12 m. Panel **(a)**: effective growth rate. Panel **(b)**: logarithm of amplitude.
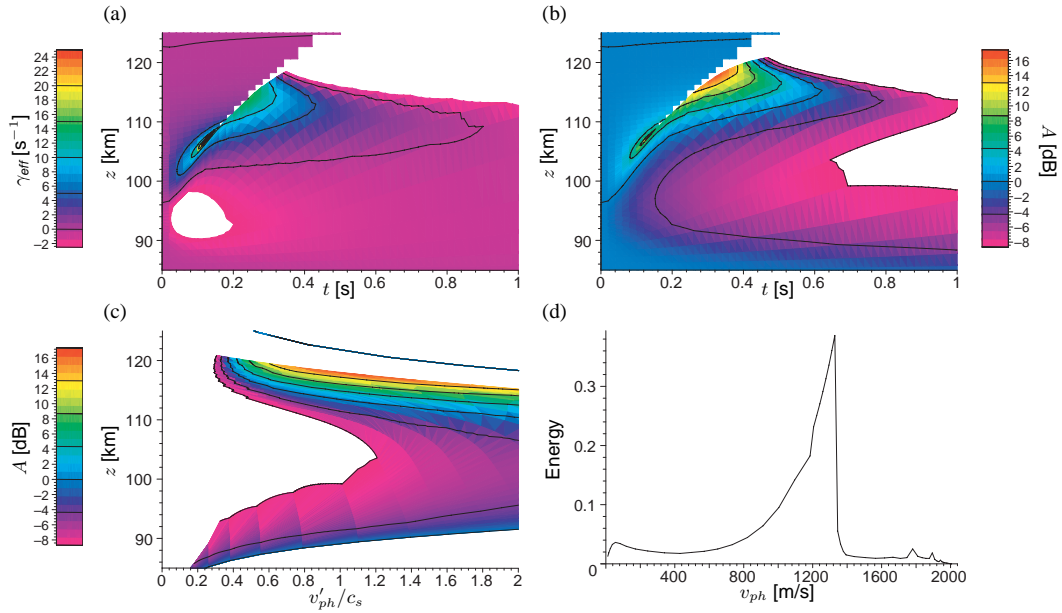

Fig. 8. Same case as in Fig. 7 but for a 100 mV/m field instead of a 50 mV/m field. Panel **(a)** and **(b)** same as in Fig. 7. Panels **(c)** and **(d)**: same as panels (a) and (e) of Fig. 6, respectively.

for the waves to have a chance to grow during the initial stages. At the later stages of evolution, shocks would still be present with our pulsed initial conditions, which would prematurely terminate the growth process.

We also studied the effects of varying the tilt or phase angle of the initial object as a whole, using our benchmark case. The tilts produced interesting modifications for the wavetrain solutions, even though figuring out just what would create such initial conditions is not easy to see. Nevertheless, more insights into the evolution of the nonlocal behavior can be gained from studying such cases. In Fig. 9 we first present the results obtained with a positive tilt as our initial condition. This was done by selecting an initial aspect angle of $-0.5°$ throughout the entire E-region, and getting rid of any additional phase angle pulse. The most important difference with the previous figures is that we found no shocks anywhere in this case: due to their initially large aspect angles, the characteristics originating at lower altitudes bent quickly enough

to avoid the formation of the shocks.

A second consequence of starting at $-0.5°$ aspect angle is that the Farley-Buneman growth rates were larger in the 110 to 120 km altitude region. When this is coupled to the absence of shocks, it allows the waves to grow as long as the aspect angle is not too large (recall that the aspect angle decreases monotonically with time in the absence of pulses, as seen from Panel a). As expected from local growth arguments, Fig. 9 shows that the amplitude in this case is a maximum when the effective growth rate is zero, that is, when it is about to turn negative due to aspect angles that have become too large. Since this is the case, we can also see from the figure that, as expected, v'_{ph} is equal to the local ion acoustic speed at peak amplitude. However, because the location of the amplitude peak is relatively high, 115 km, the phase velocity of the large amplitude waves is still of the order of 700 m/s, with approximately 250 m/s coming from the drift of the reference frame, namely the ions themselves.
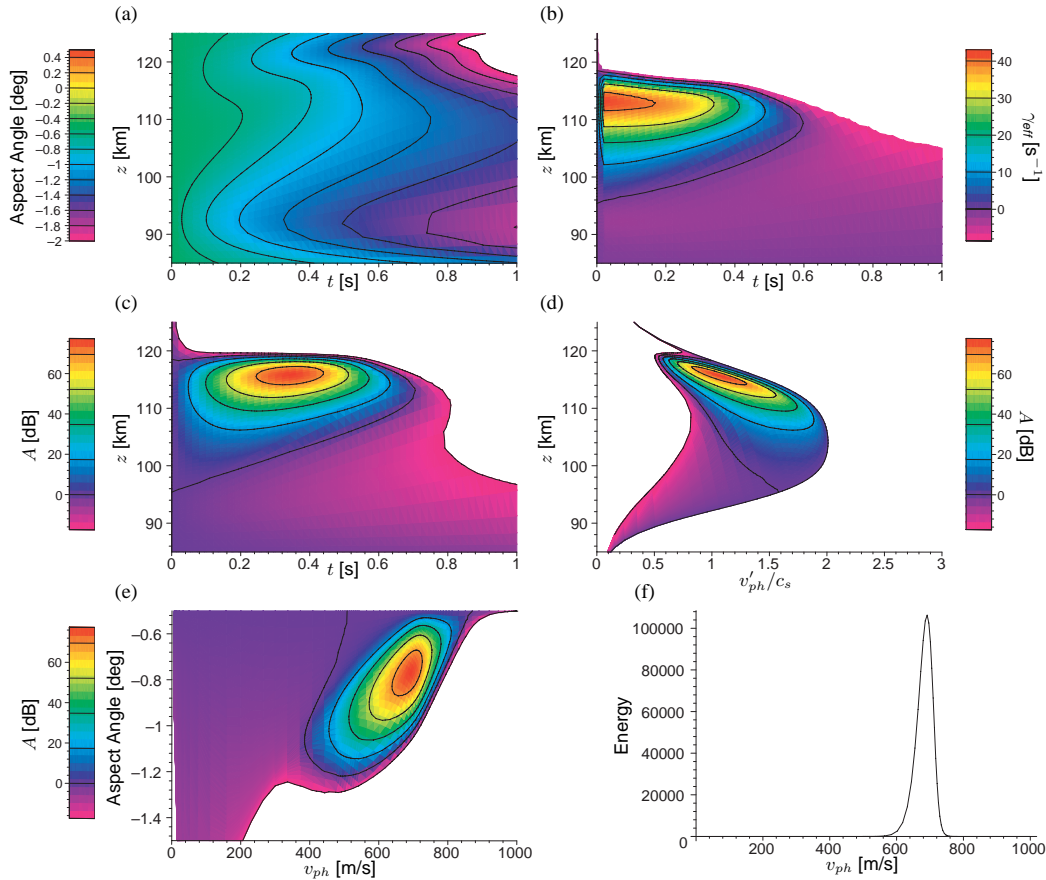
Fig. 9. Same case as in Fig. 5 but for an initial condition associated with a uniform aspect angle of $-0.5°$. Panels **(a)**, **(b)** and **(c)** are the same as panels (a), (e) and (f) in Fig. 5, respectively. Panels **(d)**, **(e)** and **(f)** are the same as panels (b), (d) and (e) in Fig. 6.

In Figure 10 we show the results associated with a negative initial tilt, that is, an initial aspect angle of $+0.5°$ throughout the entire E-region. This case is quite different in that shocks play a leading role in the evolution of the object. Not only that, the shocks are also different from those discussed so far in relation to the zero initial tilts, because both the phase and its slope are now discontinuous. The conditions that have to be applied to such cases are more technical than what we have discussed thus far and, therefore, have been moved to Appendix B.

In spite of the fact that we now have a shock produced by the initial tilt itself, Fig. 10 clearly shows that the final amplitude of the waves is actually greater than in our previous case (Fig. 9), which itself was also considerably larger than the benchmark case. The reason for the large amplitudes of the two tilted cases is rooted, as usual, in the fact that the Farley-Buneman waves actually grow more in the presence of a $0.5°$ aspect angle of either sign than at a $0°$ aspect angle. However, even though the positively tilted case is suddenly interrupted by shocks, it ends up with a larger amplitude than the non-shocked negatively tilted case of Fig. 9. A study of the growth rate plots shows why: in both cases most of the growth takes place well above the shock region of Fig. 10. However, for the positive initial aspect angles, the growth

continues for much longer than it does with the negative initial aspect angles. The reason is simply that the aspect angle monotonically moves toward negative values in both cases. Having an initial positive value ensures a longer period of time during which the magnitude of the aspect angle is less than $1°$. The longer growth time is obviously the single most important factor controlling the amplitudes in this case, because the shock happens only after a prolonged growth period. A second point of interest is that the mean phase speed of the shocked waves (Fig. 10) is actually larger than that of the unshocked ones (Fig. 9), in spite of the fact that the altitude of the shocked waves is quite a bit lower at maximum amplitude than for the unshocked case. The key factor in this case is the aspect angle, which is of the order of 0.25 degrees in the shocked case and 0.8 degrees in the unshocked case. The aspect angle affects the Doppler shift of the waves more strongly in the unshocked case, since it is actually responsible for bringing the growth rate to zero. By contrast, it is the shock in Fig. 10 that determines the maximum amplitude, and at that point the aspect angle is nowhere near a value capable of slowing the waves down to c_s.

One point that the different initial tilts brings forth is the sensitivity to initial conditions. The sensitivity is relatively easy to understand mathematically, in the sense that the as-
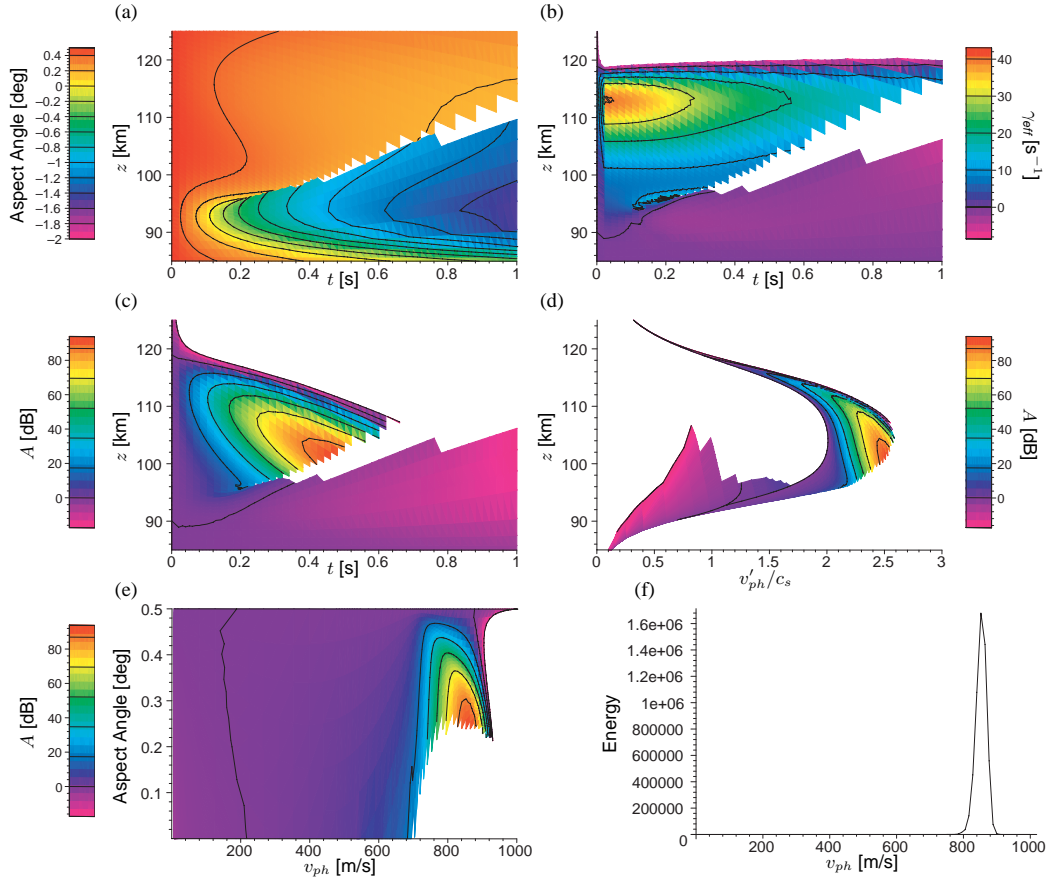
Fig. 10. Same as in Fig. 10, but with a uniform aspect angle of $+0.5°$ initially.

pect angle problem is actually strongly nonlinear. Nonetheless, the real problem is in just what triggers the Farley-Buneman waves. If one cannot answer this question, one cannot know what the initial conditions are and, therefore, what the waves' properties will be for the moderate growth conditions discussed in the present work. We surmise that a good seed of Farley-Buneman waves should be wave steepening (or mode-coupling) associated with large-scale gradient drift instabilities. These gradient-drift waves tend to grow at smaller aspect angles than Farley-Buneman waves. Therefore, they should seed Farley-Buneman waves with rather small aspect angle magnitudes. If so, the properties of waves with moderate growth conditions should be closer to our benchmark case than to the last two cases presented in this section.

5.2.3 Smaller electric fields or slower growth rates

When the plasma is just barely above threshold conditions, the growth is taken over by purely convective effects and the amplitude gain becomes quite modest. To illustrate this we have chosen to run our standard case with a smaller 35 mV/m electric field. At this point Farley-Buneman growth plays a role, but not a role that is large enough to make a big difference on the final amplitude. This case is illustrated in

Fig. 11. In Fig. 11a we first show the ratio of the Doppler-shifted phase speed, v'_{ph}, to the ion-acoustic speed, c_s, for our choice of a relatively weak electric field. The main difference with the standard run is that the peak amplitude now occurs when v'_{ph} is significantly larger than c_s. This is happening because there is very little Farley-Buneman growth in the present case. The amplitude evolution is, therefore, dominated by convective effects and conservation of wave action, which allows things to grow even when local growth is weak or even slightly negative.

To help visualize the process taking place at small growth rates, we are also showing, in Fig. 11b, how the altitude profile of the amplitude changes with time. For the first 0.3 s, convective effects simply propagate the wave train to higher altitudes. At the same time, the amplitude increases through the principle of conservation of wave action. At 0.4 s, a shock has formed, resulting in a decrease in amplitude where the characteristics are absorbed in the shock. After this time, the amplitude is controlled by the lower parts of the wave train. These lower parts arrive later at higher altitudes and do not become involved in a shock because they are associated with characteristics that evolve slowly enough with time to avoid crashing. However, this part of the wave train is also one for which conservation of wave action plays a lesser role. This means that in a relative sense, local Farley-Buneman
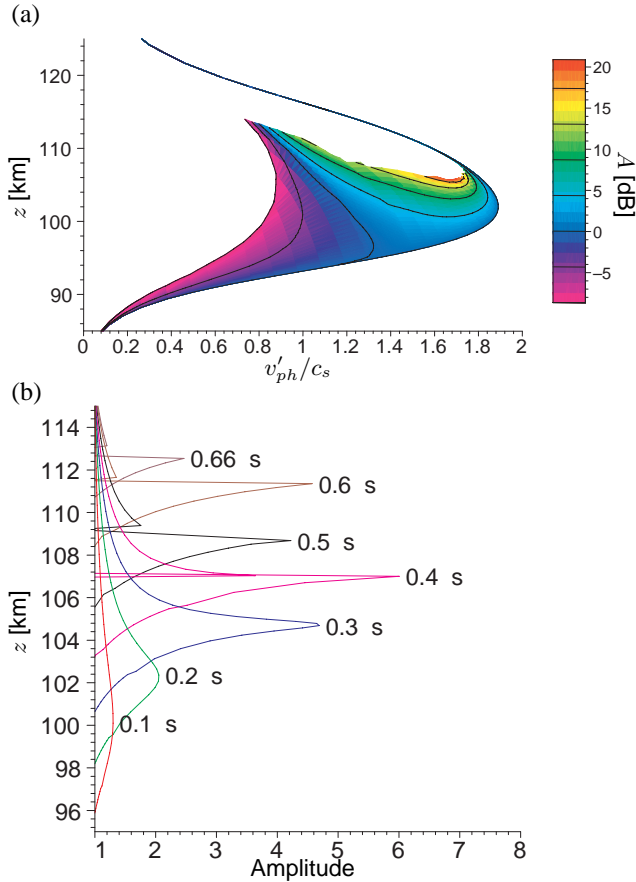
(a)

(b)

Fig. 11. Same as in Fig. 5 but for a 35 mV/m field. Panel **(a)**: same as in Fig. 6b. Panel **(b)**: profile of the amplitude as a function of altitude at various time intervals.

growth is now influencing the evolution. This does produce a growth, as illustrated by the fact that following the shock, the amplitude grows back until about 0.6 s. However, this growth is short-lived because it occurs at a relatively "large" (greater than 0.5°) aspect angle, and the aspect angle keeps increasing with time at an increasingly fast rate. In fact, due to this increasingly fast evolution in the aspect angle, not only is the growth period short-lived, but it is also followed by a steep decay. Indeed, by the time the simulation reaches 0.7 s (not shown), there is no appreciable amplitude left.

In the end, therefore, the maximum amplitude reached during the Farley-Buneman-dominated growth phase is less than the gain reached during the convective-dominated growth phase, because the wave, after the shock, does not exist for long enough in conditions exceeding the Farley-Buneman threshold conditions.

6 Discussion and conclusion

The central goal of this work has been to develop a tool for the nonlocal study of high-latitude field-aligned irregularities. We have demonstrated that the altitude dependence of

the angular frequency from the dispersion relation is responsible for the creation of parallel electric fields that evolve from any initial condition, including perfectly zero aspect angle (purely perpendicular wave vector) conditions. This is another way of saying that the aspect angle cannot, in principle, be treated as a free parameter of the problem. We have focused on E-region instabilities here and examined the evolution of high-latitude Farley-Buneman waves initially excited in the direction of the electron $E \times B$ drift. We have found that, in addition to the expected increase in the aspect angle with time, the nonlinearity of the aspect angle dependence is responsible for the formation of aspect angle (and amplitude) shocks, which we have vaguely referred to as a "crashing" mechanism. These shocks further constrain the evolution of the wave packets that we studied by limiting the amplitude of growing wave trains sometimes precisely in regions where the local growth should have been expected to be the largest. At other times, their most important role is to stop the wave train evolution before the phase speed goes down to the ion-acoustic speed.

We have compared relatively large local growth rate conditions with locally weakly linearly unstable cases. For relatively large local growth rates we have found that a ground-based observer would see the structures moving at a speed between the ion-acoustic speed and the electron drift. This is because the waves actually tend to move at a speed close to the ion-acoustic speed of the medium, but in the ion frame of reference. Since, shocks notwithstanding, the convected waves still reach their largest amplitudes near the upper portions of the unstable layer, the ion component of motion in the $E \times B$ drift direction is not negligible and the waves are, therefore, moving at a speed that is measurably greater than the ion-acoustic speed, but also measurably less than the $E \times B$ drift. These results have to be contrasted with situations for which local growth rates are modest, in which case a ground-based observer would see a phase speed that approaches the electron $E \times B$ drift. Growth in this latter case is quite reduced and is due mostly to convective effects.

In order to explore the often-neglected physics associated with convective effects, we had to neglect nonlinearities associated with large gains in amplitude, even though most of the cases we studied dealt with gains in amplitude that were several orders of magnitude strong. Such gains may well be large enough to trigger density fluctuation levels in excess of 1%, even if the structures were to be generated by thermal noise background. However, in most of the cases that we have explored, the amplitudes were found to be by far their largest in a restricted region of space and for a restricted range of phase velocities very near the ion-acoustic speed. With this in mind, consider that one could argue that perpendicular mode-coupling should control the evolution of the waves when the amplitude becomes too large. It has been argued (Hamza and St.-Maurice, 1993a) that for the geometry that we have studied in the present work, the mode-coupling mechanism would produce waves that would be seen to drift at the ion-acoustic speed of the medium (as seen from the ion frame of reference). In the end, therefore, there would be

no real change with what we have presented, except for the last case that we have shown through Fig. 10. For that latter case the final phase velocities at "saturation" were about 2.5 times the ion-acoustic speed, and the amplitudes were very large over an extended altitude regime, strongly suggesting that convective effects were probably unable to compete with nonlinear amplitude saturation effects. An interesting consequence could be that an asymmetry between positive and negative aspect angles maybe present in cases such as this, as seen by the contrast between the situations depicted in Fig. 9 and Fig. 10. At any rate, we could state, in general, that if the growth rates are so large in a particular situation that substantial growth occurs while a wave packet moves upward or downward over a distance of a few km, nonlinear amplitude saturation effects should dominate the wave evolution so that the parallel wave-vector evolution discussed here should not matter so much. Given magnitudes in the parallel group velocity that are in excess of 10 km/s, a wave train would exist for a time scale of the order of 1 or 2 s. Therefore, a rule of thumb would be that if the waves grow at a rate that exceeds a few tens of s^{-1} over an extended altitude region, convective effects would not control the amplitude evolution. This is the situation that was in fact met in Fig. 10, which means that the results could well be strongly modified by nonlinear amplitude saturation effects in that case.

While we have concentrated on Farley-Buneman waves, we could have equally well considered gradient-drift waves simply by changing the expression for the local growth rate in Eq. (14). The latter would not be very different from the former in the sense that the most important consideration remains the comparison between the value of the inverse of the largest local growth rate and the time it takes for a wave packet to move through the unstable region at the parallel group velocity. Two new factors would, nevertheless, have to be considered. First, there is the fact that gradient-drift waves have their largest growth rates at zero aspect angles, whereas Farley-Buneman waves grow the fastest at about $0.5°$ off perpendicularity. While this may not seem significant, we should recall that the waves undergo most of their evolution while the aspect angle varies between 0 and $1°$ off perpendicularity. A second factor would need to be added for gradient-drift waves excited by a sharp bottomside E-region. In that case the variation in the background electron density with respect to height would have to be considered given that its gradients would be comparable to the variations in the background neutral properties. In that case, furthermore, the favorable gradient could be limited in height, thus changing in a significant way the convective growth phase of the waves. For these reasons, we are leaving a study of gradient-drift waves per se to future work.

Our present results can also be compared with those of St.-Maurice (1985) and Moorcroft (1984). St.-Maurice (1985) had found that a nonlocal eigenfrequency decomposition could yield a negative nonlocal growth rate (i.e. an absolutely stable plasma), while the amplitude of an initial wave train could actually grow while propagating upwards at a speed of several km/s. A wave train would reach its maximum ampli-

tude before decaying on its way out of the unstable layer. The wave trains would change shape as they evolved: some of the waves in an initial pulse appeared to move down first before being reflected, and sometimes catching up with the rest of the train. St.-Maurice (1985) used an eigenfrequency decomposition with a numerical Laplace transform technique to obtain these results. For the convectively unstable situations studied in the present work, the amplitudes obtained by St.-Maurice (1985) were similar to those we have shown here for initial impulses from the lower altitudes. However, St.-Maurice (1985) obtained larger amplitudes for impulses started at higher altitudes. The reason appears to be related to the shock conditions met in the present work, which were overlooked until now, and allowed the waves to continue their growth instead of crashing.

It is more difficult to compare our results with those of Moorcroft (1984) because Moorcroft (1984) did not attempt to compute wave amplitudes. He showed instead that downward moving rays with small enough initial aspect angles would be invariably reflected from the E-region, while initially upward moving rays would continue to move upward and out of the region. On that basis he concluded that waves with too small a local growth rate would not be able to gain large amplitudes before exiting the E-region and should not be observed with radars. The present calculations quantify these conclusions and agree with the general features of Moorcroft's (1984) calculations.

We conclude that the approach proposed here for the study of nonlocal (convective) high-latitude instabilities offers a powerful new tool for the study of weakly growing field-aligned modes at high latitudes. Our WKB approach shies away from an eigenfrequency decomposition and its underlying implicit mode-coupling triggered by coupling with the inhomogeneity of the medium. Instead, it allows the frequency to depend on position as long as the wave vector is allowed to change with time through the Whitham relation (Eq. 12). For the calculation of amplitudes the method uses the conservation of wave-action in the presence of a source term. In general, this would amount to finding an expression for the wave energy for the particular system under consideration.

While our work should not be applied to fast growing 3 m waves due to their very fast local growth, the extension of our work to gradient-drift waves should have important implications, particularly for the interpretation of HF radar spectra obtained from high-latitude E-region echoes. The Super-DARN radars operate at a frequency such that the inverse growth rates are usually small by comparison to the travel time of wave packets through the E-region. Typical frequencies would match closely the 12-m wavelength situation that we have studied here, albeit with bottomside E-region gradient-drift effects needing to be added, particularly for the evening sector. For the reasons stated earlier in the present discussion, this problem needs to be studied in more detail in future work. Suffice it to say that E-region spectra obtained from HF radars have properties that differ from other radars. For instance, the Doppler width in velocity units is

often smaller than at other frequencies (e.g. Hanuise et al., 1991; Eglitis et al., 1995; Eglitis and Robinson, 1998), indicating that the waves are only weakly turbulent, at least by comparison to what is found at higher frequencies.

One aspect of the observations that our ongoing study might already be able to address is the fact that under weaker electric field conditions, the Doppler shift of HF waves is suggestive of a line-of-sight velocity matching the electron line-of-sight velocities (Villain et al., 1987; Jayachandran et al., 2000). This seems to match what we have found for our weakly growing cases, particularly the one presented in Fig. 11. In this particular situation, growth turns out to be dominated by convective processes in what could be best described as an absence of local decay (rather than a presence of local growth). This remains only a tentative result, however, because we need to study this issue in more detail by introducing effects due to the gradient-drift growth mechanism.

Appendix A Altitude parameterization used in the calculations

In order to move to a more quantitative analysis of the solution, we have introduced a very basic model of the unstable E-region. To begin with, we have assumed a vertical and constant magnetic field. The perpendicular wave vector $k_\perp$ was also taken to be parallel to V_{e0}, that is to say, we have only studied here the wave response in the electron $E \times B$ drift direction. This means that the x-component of $k_\perp$ is equal to zero since E is taken to be in the x-direction. In such a case, the scalar product $k \cdot V_{e0}$ can be written as kV_{e0}.

The neutrals are furthermore assumed to be at rest. The relative electron-ion drift consequently depends on the neutral atmospheric model through the equation (e.g. Schlegel and St.-Maurice, 1981; St.-Maurice et al., 1981).

$$V_d = V_{e0} - V_{i0}$$
$$= \frac{1}{1 + v_i^2/\Omega_i^2} \left(-\frac{v_i}{\Omega_i} \frac{E}{B} + \frac{v_i^2}{\Omega_i^2} \frac{E \times B}{B^2} \right) \tag{A1}$$

Here, the electrons have been assumed to be $E \times B$ drifting, due to their large cyclotron frequency. As a result, the magnitude of V_d is given by

$$V_d = \left[\frac{1}{(1 + v_i^2/\Omega_i^2)^2} \left(\frac{v_i^2}{\Omega_i^2} \frac{E^2}{B^2} + \frac{v_i^4}{\Omega_i^4} \frac{E^2}{B^2} \right) \right]^{1/2}$$
$$= \frac{1}{\sqrt{1 + v_i^2/\Omega_i^2}} \frac{v_i}{\Omega_i} V_{e0}. \tag{A2}$$

By the same token, the ion drift velocity is expressed by

$$V_{i0} = \frac{1}{1 + v_i^2/\Omega_i^2} \left(\frac{v_i}{\Omega_i} \frac{E}{B} + \frac{E \times B}{B^2} \right). \tag{A3}$$

The above then yields

$$k \cdot V_d = \frac{1}{1 + v_i^2/\Omega_i^2} \left[-\frac{v_i}{\Omega_i} \frac{k \cdot E}{B} + \frac{v_i^2}{\Omega_i^2} \frac{k \cdot (E \times B)}{B^2} \right]$$
$$= \frac{1}{1 + v_i^2/\Omega_i^2} \frac{v_i^2}{\Omega_i^2} k V_{e0} \tag{A4}$$

and

$$k \cdot V_{i0} = \frac{1}{1 + v_i^2/\Omega_i^2} \left[\frac{v_i}{\Omega_i} \frac{k \cdot E}{B} + \frac{k \cdot (E \times B)}{B^2} \right]$$
$$= \frac{1}{1 + v_i^2/\Omega_i^2} k V_{e0}. \tag{A5}$$

The neutral temperature T_n was taken from the MSIS-90 neutral model (Hedin, 1991). For the electrons, we assume that the temperature is equal to the neutral temperature. Ion temperatures, on the other hand, are affected by frictional heating (e.g. Schunk, 1975; Schlegel and St.-Maurice, 1981), and we obtain

$$T_i = T_n + \frac{m_n}{3k_B} V_{i0}^2. \tag{A6}$$

Finally, the collision frequencies were chosen to be the same as those used by Schlegel (1983), namely

$$v_e = 2.33 \times 10^{-11} n(N_2) \left(1 - 1.21 \times 10^{-4} T_e \right) T_e$$
$$+ 1.82 \times 10^{-10} n(O_2) \left(1 + 3.6 \times 10^{-2} T_e^{1/2} \right) T_e^{1/2}$$
$$+ 8.9 \times 10^{-11} n(O) \left(1 + 5.7 \times 10^{-4} T_e \right) T_e^{1/2} \tag{A7}$$

$$v_i = 4.24 \times 10^{-12} n(N_2) + 4.18 \times 10^{-10} n(O_2)$$
$$+ 2.38 \times 10^{-10} n(O), \tag{A8}$$

where the densities of N_2, O_2 and O were also taken from the MSIS-90 neutral model. These two expressions for v_e and v_i are actually given by Schunk and Nagy (1978) and Schunk and Walker (1973), respectively.

For completeness, also note that the perpendicular component of the group velocity is given by

$$V_{g\perp} = \frac{\partial \omega_r}{\partial k_\perp}$$
$$= \frac{\omega_r}{k} + \frac{k \cdot V_d}{(1 + \Psi)^2} \Psi_0 \frac{k_\perp}{k^2} \frac{2k_\parallel^2}{k^2} \frac{\Omega_e^2}{v_e^2}$$
$$= V_p - V_{g\parallel} \frac{k_\parallel}{k}. \tag{A9}$$

Comparing the parallel group velocity with the perpendicular group velocity indicates that the parallel group velocity is much larger than the perpendicular group velocity in the range of interest, namely for $k_\parallel \neq 0$. This is consistent with the statement made by Moorcroft (1984). By contrast, when $k_\parallel = 0$, we immediately obtain $V_{g\parallel} = 0$ and $V_{g\perp} = V_p$. In such a case the total group velocity is equal to the (perpendicular) phase velocity.
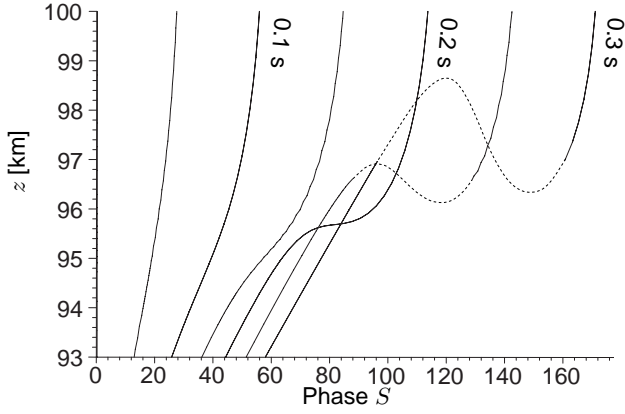
Fig. B1. Evolution of the phase S. Same as Fig. 4 but for a structure with a uniform initial aspect angle of $+0.5$ degrees. The wave tears apart, as opposed to crashing into itself. The dotted parts of the lines correspond to the multi-valued part of the solution and have to be removed according to the algorithm described in Appendix B.

Appendix B   Handling of shocks for the general case

Finding the location of the shock in the case shown in Fig. 4, where the lower parts of the wave crash into the upper part, is relatively straightforward: we only need to remove the loops in the phase. However, this treatment does not work in the case of a wave which has a positive initial tilt. A wave like that has a larger perpendicular group velocity at higher altitudes than at lower altitudes. As a result, instead of the wave crashing into itself, it is torn apart, with the upper regions moving ahead of the lower regions.

In Fig. B1 we show the phase evolution for such a case. Here, the difficulty rests with the determination of the altitude at which the tear occurs, because there is no way to make the phase continuous.

Obviously, Whitham's relation (12) no longer applies across the shock because the derivatives are undefined. However, we can write it in integral form, which is valid even across the shock. To do this, we determine the phase difference between two altitudes z_1 and z_2 on opposite sides of the shock, following the formalism presented by Logan (1994, ch. 3). Even with a jump in $k_\parallel$, it is valid to write

$$S(z_2, t) - S(z_1, t) = \int_{z_1}^{z_2} k_\parallel(z, t)\, dz \qquad (B1)$$

which follows trivially from the definition of $k_\parallel$. Taking the time derivative of this, and splitting the integral into the two branches above and below the shock $s(t)$, we obtain, using the definition of ω_r,

$$-\omega_r(z_2, t) + \omega_r(z_1, t) = \frac{d}{dt}\int_{z_1}^{s(t)} k_\parallel dz + \frac{d}{dt}\int_{s(t)}^{z_2} k_\parallel dz. \quad (B2)$$

Then, we apply Leibniz' rule for differentiating an integral whose integrand and limits depend on a parameter (here t), and take the limits $z_1 \to s(t)$ and $z_2 \to s(t)$, to find

$$-\omega_{r+} + \omega_{r-} = (k_{\parallel-} - k_{\parallel+})\frac{ds(t)}{dt}, \qquad (B3)$$

where the subscripts $+$ and $-$ refer to the values on either side of the shock. From this we obtain a differential equation for $s(t)$:

$$\frac{ds(t)}{dt} = \frac{\omega_{r+} - \omega_{r-}}{k_{\parallel+} - k_{\parallel-}}. \qquad (B4)$$

Not too surprisingly, this equation gives the parallel group velocity in the absence of a jump while providing a proper generalization in the case of a shock. To find the shock's position $s(t)$ with Eq. (B4), we also need an initial condition. We simply start the shock at the first intersection of characteristics or at the first point where a jump is encountered, depending on the kind of shock condition we face. Knowing $s(t)$, we then simply remove the parts of the solution that are multi-valued. We do this for every time step by following one solution until we reach the altitude given by $s(t)$, where we make the solution jump to the value on the other side of the shock. In Fig. B1, this means removing the dotted parts of the curves. The exact altitude where we remove the multiple values is given by $s(t)$. Comparing this to the cases where we had loops in the phase (e.g. Fig. 4), the $s(t)$ we find using this procedure is at exactly the same place where the loops start, so both procedures, loop removal and finding $s(t)$, give the same result.

References

Bretherton, F. P. and Garret, C. J. R.: Wave trains in inhomogeneous moving media, Proc. Roy. Soc. A, 302, 529–554, 1969.

Eglitis, P., Robinson, T. R., McCrea, I. W., Schlegel, K., Nygren, T., and Rodger, A. S.: Doppler spectrum statistics obtained from three different-frequency radar auroral experiments, Ann. Geophysicae, 13, 56–65, 1995.

Eglitis, P. E. and Robinson, T. R.: Investigations of the phase speed and spectral width of high latitude plasma irregularities, Adv. Space Sci., 22, 1361–1364, 1998.

Gagnepain, J., Crochet, M., and Richmond, A. D.: Comparison of equatorial electrojet models, J. Atmos. Terr. Phys., 39, 1119–1124, 1977.

Hamza, A. M. and St.-Maurice, J.-P.: A turbulent theoretical framework for the study of current-driven E region irregularities at high latitudes: Basic derivation and application to gradient-free situations, J. Geophys. Res., 98, 11 587–11 599, 1993a.

Hamza, A. M. and St.-Maurice, J.-P.: A self-consistent fully turbulent theory of auroral E region irregularities., J. Geophys. Res., 98, 11 601–11 613, 1993b.

Hanuise, C., Villain, J. P., Cerisier, J. C., Senior, C., Ruohoniemi, J. M., Greenwald, R. A., and Baker, K. B.: Statistical study of high-latitude E-region doppler spectra obtained with the SHERPA HF radar, Ann. Geophysicae, 9, 273–285, 1991.

Hedin, A. E.: Extension of the MSIS thermospheric model into the middle and lower atmospheres, J. Geophys. Res., 96, 1159–1172, 1991.

Hu, S. and Bhattacharjee, A.: Two-dimensional simulations of gradient-drift turbulence in the daytime equatorial electrojet, J. Geophys. Res., 103, 20 749–20 759, 1998.

Janhunen, P.: Three-dimensional stabilization mechanism for the auroral Farley- Buneman instability, J. Atmos. Terr. Phys., 54, 1633–1643, 1992.

Janhunen, P.: Perpendicular particle simulation of the E region Farley-Buneman instability, J. Geophys. Res., 99, 11 461–11 473, 1994.

Jayachandran, P., St.-Maurice, J.-P., MacDougall, J. W., and Moorcroft, D. R.: HF detection of slow long-lived E-region plasma structures, J. Geophys. Res., 105, 2425–2442, 2000.

Kelley, M.: The earth's ionosphere. Plasma physics and electrodynamics, Academic Press, 1989.

Kundu, P. K.: Fluid Mechanics, Academic Press, 1990.

Lacroix, P. J. and Moorcroft, D. R.: Ion acoustic HF radar echoes at high latitudes and far ranges, J. Geophys. Res., 106, 29 091–29 104, 2001.

Logan, J. D.: An Introduction to Nonlinear Partial Differential Equations, Wiley-Interscience, 1994.

Machida, S. and Goertz, C.: Computer simulation of the Farley-Buneman instability and anomalous electron heating in the auroral ionosphere, J. Geophys. Res., 93, 9993–10 000, 1988.

Milan, S. E., Yeoman, T. K., Lester, M., Thomas, E., and Jones, T. B.: Initial backscatter occurrence statistics from the CUTLASS HF radars, Ann. Geophysicae, 15, 703–718, 1997.

Moorcroft, D. R.: Propagation of plasma wave energy in the auroral E region, J. Geophys. Res., 89, 2963–2970, 1984.

Oppenheim, M., Otani, N., and Ronchi, C.: Hybrid simulation of the saturated Farley-Buneman instability in the ionosphere, Geophys. Res. Lett., 22, 353–356, 1995.

Otani, N. F. and Oppenheim, M.: A saturation mechanism for the Farley-Buneman instability, Geophys. Res. Lett., 25, 1833–1836, 1998.

Riggin, D. and Kadish, A.: Nonlocal theory of long-wavelength plasma waves associated with sporadic E-layers, J. Geophys. Res., 94, 1495–1500, 1989.

Robinson, T. R.: Towards a self-consistent nonlinear theory of radar auroral backscatter, J. Atmos. Terr. Phys., 48, 417–423, 1986.

Ronchi, C., Sudan, R., and Farley, D. T.: Numerical simulations of large-scale plasma turbulence in the daytime equatorial electrojet, J. Geophys. Res., 96, 21 263–21 279, 1991.

Schlegel, K.: Interpretation of auroral radar experiments using a kinetic theory of the two-stream instability, Rad. Sci., 18, 108–118, 1983.

Schlegel, K. and St.-Maurice, J.-P.: Anomalous heating of the polar E-region by unstable plasma waves, 1. Observation, J. Geophys. Res., 86, 1447–1452, 1981.

Schunk, R. W.: Transport equations for aeronomy, Planet. Space Sci., 25, 437, 1975.

Schunk, R. W. and Nagy, A. F.: Electron temperatures in the F-region of the ionosphere: Theory and observations, Rev. Geophys. Space Phys., 16, 355–399, 1978.

Schunk, R. W. and Walker, J. C. G.: Theoretical ion densities in the lower ionosphere, Planet. Space Sci., 21, 1875–1896, 1973.

St.-Maurice, J.-P., A nonlocal theory of the high-latitude Farley-Buneman instability, J. Geophys. Res., 90, 5211–5225, 1985.

St.-Maurice, J.-P.: A unified theory of anomalous resistivity and Joule heating effects in the presence of E-region irregularities, J. Geophys. Res., 92, 4533–4542, 1987.

St.-Maurice, J.-P.: Wave-induced diffusion in the turbulent E-region, in Polar cap Dynamics and High Latitudes turbulence: SPI Conference proceedings and Reprint series, No. 8 (1988), pp. 323–348, Scientific Publishers, Cambridge, Mass., 1990.

St.-Maurice, J.-P. and Hamza, A. M.: A new nonlinear approach to the theory of E-region irregularities, J. Geophys. Res., 106, 1751–1759, 2001.

St.-Maurice, J.-P. and Laher, R.: Are observed broadband plasma wave amplitudes large enough to explain the enhanced electron temperatures of the high-latitude E-region?, J. Geophys. Res., 90, 2843–2850, 1985.

St.-Maurice, J.-P., Schlegel, K., and Banks, P. M.: Anomalous heating of the polar E-region by unstable plasma waves, 2. Theory, J. Geophys. Res., 86, 1453–1462, 1981.

St.-Maurice, J.-P., Cussenot, C., and Kofman, W.: On the usefulness of E-region electron temperatures and lower F-region ion temperatures for the extraction of thermospheric parameters: a case study., Ann. Geophysicae, 17, 1182–1198, 1999.

Sudan, R. N.: Nonlinear theory of type I irregularities in the equatorial electrojet, Geophys. Res. Lett., 10, 983–986, 1983a.

Sudan, R. N.: Unified theory of type I and type II irregularities in the equatorial electrojet, J. Geophys. Res., 88, 4853–4860, 1983b.

Sudan, R. N. and Keskinen, M. J.: Theory of strongly turbulent two-dimensional convection of low-pressure plasma, Phys. Fluids, 22, 2305–2314, 1979.

Thiemann, H. and Schlegel, K.: Particle-in-cell plasma simulations of the modified two stream instability, Ann. Geophysicae, 12, 1091–1100, 1994.

Villain, J. P., Greenwald, R. A., Baker, K. B., and Ruohoniemi, J. M.: HF radar observations of E-region plasma irregularities produced by oblique electron streaming, J. Geophys. Res., 92, 12 327–12 342, 1987.

Villain, J. P., Hanuise, C., Greenwald, R. A., Baker, K. B., and Ruohoniemi, J. M.: Obliquely propagating ion acoustic waves in the auroral E-region: Further evidence of irregularity production by field-aligned electron streaming, J. Geophys. Res., 95, 7833–7846, 1990.