The thermal self-focusing instability near the critical surface in the high-latitude ionosphere

P.N. Guzdar and P.K. Chaturvedi
Institute for Plasma Research, University of Maryland, College Park

K. Papadopoulos
Department of Physics, University of Maryland, College Park

S. L. Ossakow
Plasma Physics Division, Naval Research Laboratory, Washington, D.C.

Abstract.
A fully nonlinear development of the thermal self-focusing instability of high power radio waves in the ionosphere in the region near the critical surface is the subject of the present study. In the simulation model studied, a high-powered radio wave in the frequency range 5 - 10 MHz, with a 1% amplitude modulation, is launched vertically. In the high latitude geometry this represents a direction antiparallel to the magnetic field which is almost vertically downwards. The modulated wave undergoes strong self-focusing at the critical surface, where the group velocity of the wave goes to zero. The scale size of the structures transverse to the magnetic field is controlled by the wave intensity and the diffraction effects. The large parallel thermal conduction leads to the diffusion of these irregularities into the underdense and overdense plasma in narrow filaments. The depletion in the density in the overdense plasma allows propagation of the wave to higher altitude above the original critical surface and hence into the overdense plasma.

1. Introduction
The ionospheric modification experiments using RF heater facilities at Arecibo, Platteville, Alaska, the SURA facility in Russia, and in Tromso, provide a rich variety of results related to the spatial structures in the ionospheric medium, the scattered electromagnetic signals, and particle energization [Litvak, 1970; Carlson and Duncan, 1977; Carlson et al., 1982; Stubbe et al., 1982; Fejer et al., 1985; Erakhimov et al., 1987; Wong et al., 1989]. These experiments have shown that increases in HF power results in the predominance of nonlinear effects, for example, the self-focusing instability [Litvak, 1970; Perkins and Valeo, 1974; Cragin et al., 1977; Gurevich, 1978; Perkins and Goldman, 1981; Bernhardt and Duncan, 1982 and 1987]. Observations made with ionosondes [Utlaut, 1970], scintillation studies [Thome and Perkins, 1974; Basu et al., 1983, 1987], radar scattering [Duncan and Behnke, 1978; Frey et al., 1984], in-situ satellite measurements [Farley et al., 1983], optical emissions [Bernhardt et al., 1988], dynasonde HF radar [Wright et al., 1988], and, in-situ rocket measurements [Kelley et al., 1995] have revealed the excitation of irregularities in the medium during the RF (radio wave) heating experiments, which have been attributed to the self-focusing instability (SFI). There are other nonlinear processes that can occur in the vicinity of the critical surface, namely the oscillating two-stream instability (OTSI) or the modulational instability, which leads to strong Langmuir turbulence [Dubois et al., 1990] and the parametric decay instability (PDI) [Fejer, 1979; Stubbe et al., 1984] both driven by the ponderomotive force. These processes generate short scale length (less than or of the order of meters) fluctuations, which occur in the unconditioned ionospheric plasma in the very early phase (< tens of milliseconds). Our focus is on processes that occur on the longer heating timescales (100 ms to few seconds) and generate large as well as intermediate scale structures (tens of meters to kilometers).

Theories have often been successful in explaining certain observational features. There are, however, several puzzling observations and associated scientific controversies. For example, in the context of ionospheric modification, the observations often indicate the generation of medium scale size (hundreds of meters) irregularities for typical pump powers, while the theoretical estimates predict longer scales (> kilometers) for these power levels. Current theoretical models overestimate the thresholds for excitation of the observed scales by an order of magnitude [Farley et al., 1983; Frey et al., 1984; Frey and Duncan, 1984].
Current theories emphasize identification of the physical processes leading to self-focusing in homogeneous plasmas. They have often not considered many realistic effects, such as density irregularities in the medium. These nonideal factors affect the self-focusing instability, because they affect the local refractive index of the plasma. In a recent paper we [Guzdar et al., 1996], studied the self-focusing of radio waves in the ionosphere for the underdense (pump frequency, $\omega_0$, much greater than the plasma frequency, $\omega_{pe}$) case in two dimensions (2-D). The 2-D geometry captures the essential physics aspects of the self-focusing instability. We found that the presence of preformed density irregularities in the medium results in simultaneous excitation of the self-focusing instability at short and long wavelengths due to mode coupling. This broadband excitation occurs for threshold pump powers for which the homogeneous plasma theory predicts excitation of long wavelengths only. The mechanism is akin to the quasi-resonant mode coupling excitation of high $k$ wavenumbers by long wavelength Langmuir waves in the presence of sinusoidal density perturbations [Kaw et al., 1973]. These results provide a natural explanation for the observations of small scale sizes at low pump powers [Frey et al., 1984; Frey and Duncan, 1984]. Bernhardt and Duncan [1982] have carried out 2-D numerical simulations of the self-focusing instability. They included an initial sinusoidal perturbation of the density in the case of underdense plasma when the pump wave was represented by a plane wave. This density perturbation leads to the linear growth of the self-focusing instability (SFI) at the same wavelength. Their results for this case showed that the initial pattern evolved into a distorted pattern (due to nonlinear effects) with smaller-scale sizes. The fundamental limitation of most of these studies (including our earlier work) is that they focus on the underdense plasma, where there is no clear observational evidence of structuring. Few of the theoretical studies have extrapolated the results to the critical density and overdense region, where there is a preponderance of evidence for a broad spectrum of scale sizes, from kilometers to tens of meters. Cragin et al. [1977] and Gurevich, [1978] have developed theories for the linear stability of the self-focusing instability in the vicinity of the critical surface and have found that it is an absolute instability, unlike the underdense case. Motivated by these works we have developed a two dimensional nonlinear code to study the full nonlinear evolution of the thermal self-focusing instability. Our simulations show that for a pump with a frequency of 5 – 10 MHz, at an intensity of 100 – 300 $\mu W/m^2$ the critical surface becomes unstable to the thermal filamentation. In the time scale of a few seconds a broad spectrum of scale sizes develop nonlinearly with a distinct peaking at short scale lengths, determined by diffraction. As the parallel thermal conduction transports the heat into the overdense region, a depletion of the density in the overdense region occurs. This allows the radio wave to propagate upward beyond the original critical surface. Thus because of the nonlinear development of the instability, initially scale lengths of the order of kilometers grow and by the process of thermal self-focusing collapse to shorter scale lengths of the order of tens to hundreds of meters. Also, this process leads to the formation of field-aligned filaments which extend into the overdense as well as underdense region of the plasma by transport processes.

The paper is organized as follows: We first derive the basic nonlinear system of equations for the thermal filamentation in the region of the critical surface. This is followed by a presentation of the results obtained from the solution of these equations. We finally discuss the implications of these results to some of the high latitude heating facilities.

2. Basic Equations

The basic geometry for the wave propagation and the orientation of the magnetic field in the high latitude is shown in Figure 1. The wave propagates vertically upward along the magnetic field along the z axis. The plasma density is assumed to increase linearly as a function of $z$. The surface $z = z_c$ is the critical surface where the local plasma frequency $\omega_{pe}(z_c)$ matches the heater wave frequency $\omega_0$. The basic equations representing the wave propagation and the plasma are [Gurevich, 1978]

\[
\frac{\partial^2}{\partial t^2} - c^2 \nabla^2 + \omega^2_{pe}(z_c) \left( 1 + \frac{z - z_c}{L} + \frac{\delta n}{n_0} \right) \nabla \phi = 0
\]

\[
\frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left( \chi_{\perp} \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial z} \left( \chi_{\parallel} \frac{\partial T}{\partial z} \right) + Q - \delta \nu (T - T_b)
\]

and an isothermal equation of state $P = nT$, where $P$ is the pressure, $n$ is density, and $T$ is the temperature. Here $L$ is the scalelength of the density inhomogeneity, $\chi$ is the parallel thermal conduction coefficient, $\chi_{\perp}$ is the perpendicular thermal conduction coefficient, and $Q = (v_0^2/2\nu)\nu T$ is the time-averaged ohmic heating by the wave, with $v_0 = eE_0/(m\omega_0)$ the jitter ve-

![HIGH-LATITUDE GEOMETRY](image)

Figure 1. Schematic of the high-latitude geometry for the self-focusing instability.
locity of the electron in the wave, \( v_e \) is the electron thermal velocity, \( \nu \) the electron-ion collision frequency, \( \delta = m_e/2m_i \), and \( T_b \) the background ion temperature. The specific expressions for the thermal conduction coefficients are \( \chi_\parallel = 3.93\nu^2/v \) and \( \chi_\perp = 1.2\rho^2v \), where \( \rho_e \) is the electron gyroradius. The perturbed density \( \delta n \) represents the change in the density caused by the presence of the wave. It is related to the change in the electron temperature from the background temperature by assuming that the electron pressure remains constant. This is achieved by dynamics along the field line. If we write \( T = T_b + T_1 \), then from the constancy of the pressure, we get that \( \delta n/n_0 = -T_1/(T_b + T_1) \). This simple relationship between the perturbed density and temperature shows that the heating leads to density depletion. Furthermore, we also see that for very high intensities, such that the heating raises the temperature significantly higher than the original background temperature, the perturbed density would at best be equal to \( n_0 \). Now if we substitute \( T = T_b + T_1 \) in equation (2), the equation for \( T_1 \) becomes

\[
\frac{\partial T_1}{\partial t} = \frac{\partial}{\partial x}\left( \chi_\perp \frac{\partial T_1}{\partial x} \right) + \frac{\partial}{\partial z}\left( \chi_\parallel \frac{\partial T_1}{\partial z} \right) + Q - \delta n T_1
\]

Thus equations (1), (3), and the local relationship \( \delta n/n_0 = -T_1/(T_b + T_1) \) is our basic set of equations. To further make the system of equations amenable to an efficient numerical algorithm, we first remove the very fast timescale associated with the electromagnetic pump wave frequency by representing it as

\[
E_0 = E_{\infty}e^{-i\omega_0 t}
\]

The “slow” wave equation now becomes

\[
\left[ \frac{\partial}{\partial t} - i\frac{\omega^2}{2\omega_0}\left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} \right) \right] V_0 + \frac{i\omega\rho_e(x,z)}{2}\left( z - \frac{z_c}{L} + \frac{\delta n}{n_0} \right) E_{\infty} = 0
\]

We now introduce the following normalizations to derive a set of dimensionless equations. We normalize the spatially independent variables \( x \) and \( z \) to the Airy length, \( z_0 = (c^2L/\omega_0^2)^{1/3} \), time \( t \) to \( t_0 = (2\omega_0^2\alpha^2/c^2) \), and define a new temperature variable as \( \theta = \alpha T_1/T_b \) with \( \alpha = \omega_0^2\rho_e^2/c^2 \). With these normalizations the two dimensionless equations for thermal filamentation in the vicinity of the critical surface are

\[
\left[ \frac{\partial}{\partial t} - i\left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} \right) \right] V_0 = 0
\]

\[\frac{\partial \theta}{\partial t} = \kappa_\perp \left( f_1(\theta)\frac{\partial \theta}{\partial x} \right) + \kappa_\parallel \left( f_2(\theta)\frac{\partial \theta}{\partial z} \right) + \gamma f_3(\theta)(|V_0|^2 - \theta)
\]

with \( f_1(\theta) = 1/(1 + \theta/\alpha)^{3/2} \), \( f_2(\theta) = (1 + \theta/\alpha)^{7/2} \) and \( f_3(\theta) = 1/(1 + \theta/\alpha)^{5/2} \). The dimensionless parameters are (1) \( \alpha = \omega_0^2\rho_e^2/c^2 \), (2) \( \kappa_\perp = 3.9\nu^2/v\rho_e^2 \), (3) \( \kappa_\parallel = 1.2\rho^2\nu/v\rho_e^2 \), (4) \( \gamma = 2m_e\nu\tau_0/m_i \), and (5) \( |V_0|^2 = |v_0|^2/v_0^2 \) with \( v_0^2 = 6c_0^2/\alpha \) and \( c_0 \), the ion acoustic speed with electron temperature. These 2-D coupled system of equations are solved numerically using a split-step, pseudospectral algorithm [Press et al., 1986; Guzdar et al., 1993]. The results of these studies are presented in the section 3.

3. Numerical Results

We first choose a typical set of plasma parameters which are needed to evaluate the various dimensionless parameters derived in section 2. We choose high-latitude \( F \) region parameters. The incident wave is assumed to have a frequency \( f_0 = 5 \text{ MHz} \). The magnetic field \( B_0 = 0.5 \text{ Gauss} \), and the electron temperature \( T_e = 0.1 \text{ eV} \). The plasma is assumed to have a linear density profile with the scale length \( L = 200 \text{ km} \). We will hold these parameters as constant for the runs we discuss in this section. The parameters we will vary will be the intensity of the incident radiation and the length of the box in the direction transverse to the magnetic field. We will discuss this at length later in this section.

For the values chosen for the plasma parameters, the Airy length \( z_0 \) is 263 ms, and the normalization timescale \( t_0 = 483 \mu\text{secs} \). The various dimensionless parameters can be evaluated: \( \alpha = 758.7, \kappa_\perp = 0.123, \kappa_\parallel = 7.2 \times 10^{-11}, \gamma = 1.2 \times 10^{-6} \). Also for these normalizations, the parameter \( |V_0|^2 = 1 \) corresponds to an intensity = 0.067 \( \mu\text{W/m}^2 \). We now choose for our first run, the incident intensity as \( 0.067 \text{ mW/m}^2 \). This makes our dimensionless parameter \( |V_0|^2 = 10^4 \).

The next important aspect of the simulations to discuss are the boundary conditions. At the lower boundary, \( z = 0 \), we specify \( V_0 \) and at the top boundary, \( z = L = 15\pi \), the wave amplitude is chosen to be zero. This is because beyond the critical surface at \( z = 7.5\pi \), the wave is evanescent. For the temperature the boundary conditions in \( z \) are that the derivative with respect to \( z \) be zero. This is because the parallel thermal conduction transports the energy away from the region of heating and is allowed to escape through the boundary. In the transverse direction for both the em wave and the temperature we use periodic boundary conditions. The choice of the size of the box in this direction needs to be addressed in some detail. The characteristic scale size in this direction should be the basic size of the heating region, which is typically about 30 km. Because of the self-focusing instability we anticipate the shortest scale lengths to be of the order of 10 to 100 meters. This would require at least \( 10^4 \) grid points in this direction. However, for the present intensities, even scale lengths of the order of a kilometer are linearly unstable. Thus in our normalized unit we choose the size in the \( x \) direction to be \( L_x = 2\pi \). This will allow us to
Wave amplitude is basically an Airy-like pattern. The "collapse" due to the self-focusing instability. Finally, plasma in that region, as seen in the temperature contours, the second term in equation (5) be-

In Figure 2 we show the amplitude contours at time $t = 10$. It shows that the wave which starts at $z = 0$ gets reflected at the critical surface at $z_c = 25$. The wave amplitude is basically an Airy-like pattern. The swelling factor increases the amplitude at the critical surface by a factor of 3. In this very early phase the temperature has not evolved since the heating time is 2 orders of magnitude longer (for the given intensity) than the wave propagation time. Also, at $z = 0$, we have introduced a 1% modulation in the amplitude of the incident wave, with a spatial dependence given by $\sin 2\pi x/L_x$. With the chosen normalizations the wavelength of this perturbation is 1.6 km. As time progresses, the region of higher intensity heats up faster than the loss due to parallel thermal conduction along the field lines. Also, since the wave group velocity is zero near the critical surface, the stationarity of the wave reduces wave convective losses. As a consequence thermal self-focusing of the incident wave starts near the critical surface and is an absolute instability [Cureich, 1978]. Shown in Figures 3a and 3b are the contours of the wave amplitude and the temperature at $t = 200$. In real-time units this is about 50 ms. The thermal self-focusing has created two hot spots at the critical surface, which is clearly seen both on the wave amplitude and on the temperature contours. The wave in the underdense plasma ($z < 25$) has also heated the plasma in that region, as seen in the temperature contours. Also, the spatial structure for the wave amplitude is on a shorter scale length (Airy length) compared to the temperature. The large parallel thermal conduction diffuses the heat along the field line wiping out any temperature structures on the Airy scale. However, since the swelling occurs over many Airy scale lengths, there is a variation in the temperature along the field line. Another interesting feature to observe is that the original critical surface was at $z = 25$, beyond which the wave became evanescent as shown in Figure 2. If we look at the contours for the wave amplitude at $t = 200$ (Figure 3a), we observe that the wave has propagated beyond the original critical surface. In fact, if we look at the temperature contours, we see that the heat has diffused to heights beyond $z = 30$. It is this diffusion of the heat which leads to a reduction in the density in the overdense region, which allows for the propagation of the wave beyond the original critical height. At a much later time $t = 2000$, which in real time is about half a second after the heater was turned on, the wave amplitude and temperature perturbations have undergone significant evolution (Figures 3c and 3d). The original two hot spots in the vicinity of the critical surface have now evolved into two field-aligned filaments, as seen clearly in the temperature contours in Figure 3d. Although the diffusion of the temperature occurs above and below the critical surface, the additional heating of the wave in the underdense region accounts for the higher temperature in the filaments in the underdense plasma. What is also most interesting to see is the generation of new filaments on either side of the two original filaments. During the nonlinear evolution of the hot spots the thermal self-focusing leads to channeling of the wave into the filament, which further heats the filament and facilitates the self-focusing into a more localized channel. As the filament collapses to smaller transverse scale-lengths, the second term in equation (5) becomes important. This term, which causes diffraction, leads to transverse propagation of the wave. The transverse propagating wave packet seeds new sites where thermal filamentation is initiated and new secondary filaments are spawned. Furthermore, since the heat has diffused much farther into the overdense region, the wave propagation has progressed significantly farther into the overdense plasma. The two original filaments of enhanced wave intensity have propagated higher in altitude than the new filaments, as is to be expected. Also, at this stage the size of the filaments does not decrease any further, and the original filaments have reached a saturated state.

What determines the transverse scalelength of the filaments? The self-focusing instability is finally stopped by the combined effects of the increase in the parallel thermal conduction (because of the increase in the temperature) and the transverse diffraction of the wave. The transverse size of the filament $\Delta z$ can be estimated by balancing the diffraction term (second term in equation (5)) with the density depletion caused by the self-focusing (the last term on the left hand side of equation (5)). It is given by
The density perturbation $\delta n$ is determined by the temperature perturbation which is controlled by the parallel thermal conduction loss and the collisional loss to the background ions. Thus the smallest size that a filament can achieve is of the order of the electron skin depth.

In Figures 4a, 4b, and 4c we show the temperature, density, and wave amplitude respectively, at the original critical surface at $z = 25$, for three different instants of time $t = 200$, $t = 600$, and $t = 2000$ as a function of $x$. The plots show an interesting feature for the temperature and density. Besides the development of the instability which leads to the localized structures, the average temperature as well as the average density also evolve. Thus temperature increases, while the density decreases. The maximum amplitude of the density and temperature fluctuations is about $5 - 10\%$.

4. Discussion

We have investigated the full nonlinear two-dimensional development of the thermal self-focusing instability in the high-latitude F-region ionosphere near the critical surface, where the wave frequency matches the local plasma frequency. We find that an absolute instability develops at this surface. The typical power densities which destabilize modes having wavelengths of a few kilometers, are $100 - 300 \mu W/m^2$. The local heating in the vicinity of the critical surface leads to density depletion, which then spreads along the field lines.

\[
\Delta x = \frac{c}{\omega_p^2} \sqrt{\frac{n_0}{\delta n}} \quad (8)
\]
thermal diffusion process allows for the heat to be transported into the overdense plasma. The reduction in the local density there leads to propagation of the waves and the field-aligned irregularity to extend into the overdense region. As the instability develops, the structures collapse in the direction transverse to the direction of the magnetic field. The smallest scale lengths obtained are in the range of tens of meters for initial perturbations in the range of kilometers. The associated average amplitude of the density irregularities are typically 2–5%. In the recent work by Basu et al. [1997], electron density irregularities in the F region, excited by the European Incoherent Scatter (EISCAT) high-power facility and observed by scintillations of the 250 MHz satellite signals, show that the spectra are in the range of kilometers to tens of meters. Furthermore, the saturation amplitudes are attained in times less than tens of seconds, which is also consistent with observations. An interesting feature from our simulation, is that the original long wavelength instability grows at a rate slower than the short wavelengths, since the short wavelength occurs due to a nonlinear collapse which is on a faster timescale. Thus our present model has many features consistent with observational data.

**Acknowledgments.** This work was supported by NSF under the Grant Number ATM-9412848, and by the ONR. The Editor thanks R. P. Drake and another referee for their assistance in evaluating this paper.

**References**


