JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 103, NO. A4, PAGES 7017–7026, APRIL 1, 1998

A z mode electron-cyclotron maser model for bottomside ionospheric harmonic radio emissions

A. J. Willes and S. D. Bale[1]
Astronomy Unit, Queen Mary and Westfield College, University of London, London

Z. Kuncic
Department of Physics and Astronomy, University of Victoria, British Columbia, Canada

Abstract. A model is proposed for recent ground-based observations of auroral roar emissions, detected at $2\Omega_e$ and $3\Omega_e$, where Ω_e is the local electron cyclotron frequency in the source region, between 200 and 500 km above the Earth's surface. Electron cyclotron maser emission is a likely mechanism to account for these emissions because it naturally produces coherent radiation at harmonics of Ω_e. A theory for auroral roar emissions has already been proposed, whereby maser-generated second (X2) and third (X3) harmonic x mode radiation is amplified in the source region by multiple reflections off the walls of the density cavity in which they are produced. After many reflections the X2 and X3 waves propagate along the density cavity to a ground-based observer. However, it is demonstrated here with ray-tracing calculations that it is highly probable that maser-generated X2 and X3 radiation is reabsorbed at lower altitudes and thus cannot be detected at the ground. An indirect maser mechanism is proposed instead, where maser-generated z mode waves at Ω_e grow to high levels in the source region and then undergo repeated nonlinear wave-wave coalescence to produce second- and third-harmonic waves that propagate directly to the ground. The z mode waves must satisfy the necessary kinematic constraints to produce observable second- and third-harmonic radiation. The dependence of the z mode maser on the temperature and functional form of the unstable electron distribution is discussed, along with the conditions required for the coalescence processes to proceed and produce the observed levels of radiation.

1. Introduction

Ground-based observations in the auroral zone of narrowband radiation at twice the (local ionospheric) electron cyclotron frequency [*Kellogg and Monson*, 1979, 1984] have recently been supplemented by the detection of third-harmonic emissions [*Weatherwax et al.*, 1993a, 1995]. These "auroral roar" emissions occur in the frequency range 2500–3000 kHz (for the second harmonic) and 3800–4300 kHz (third harmonic), and the harmonic components can be observed either simultaneously or individually. However, when observed simultaneously, the two harmonic bands typically do not exhibit similar features or obey an exact harmonic (3:2) frequency relationship [*Weatherwax et al.*, 1993a], implying that they can only correspond to second- and third-harmonic emission if each component is produced at different source heights (with slightly variant values of the local electron cyclotron frequency Ω_e). A cyclotron maser theory has been proposed [*Yoon et al.*, 1996] to explain these observations. In this model the observed emissions are produced by direct electron-cyclotron maser emission in the second-harmonic (X2) and third-harmonic (X3) magnetoionic x modes. The free energy for the maser is provided by an unstable, trapped electron population in the lower ionosphere. Energetic electrons are trapped at these altitudes between the opposing effects of magnetic mirroring at lower altitudes and parallel electric fields at higher altitudes. Comparison of the relative temporal growth rates for the fundamental and harmonic x and o modes reveals that X2, X3 and O1 (fundamental o mode) waves can all be produced for typical parameter values in the bottomside ionosphere. The fundamental x mode is suppressed because the required low densities are not attained in the ionosphere.

[1]Now at Space Sciences Laboratory, University of California, Berkeley.

Copyright 1998 by the American Geophysical Union.

Paper number 97JA03601.
0148-0227/98/97JA-03601$09.00

In this paper we draw attention to growth of fundamental z mode waves (Z1). The z mode is the "slow" branch of the magnetoionic extraordinary mode. Fundamental z mode waves were not included in the analysis of *Yoon et al.* [1996] because they focused on direct maser emission that generates waves that can escape the ionosphere to reach a ground-based observer. The z mode waves are usually trapped in the ionosphere and can only propagate to the ground in special circumstances [Horne, 1995]. Previous analyses of maser growth rates comparing growth rates in the various competing modes [*Melrose et al.*, 1984; *Winglee*, 1985; *Benson and Wong*, 1987] have demonstrated the dominance of the z mode over fundamental and harmonic x and o modes over a broad range of parameters, particularly when the spatial growth rate is considered (which favors z mode waves because of their slow group speed).

Here we argue the case for a z mode-driven maser mechanism to explain the auroral roar emissions. First, we show that the third-harmonic x mode (X3) direct maser spatial growth rates are too low (relative to the second harmonic) to explain the observed comparable intensities of the two harmonic components. We then demonstrate, with a ray-tracing code, that it is highly probable that direct, maser-generated X2, X3, and O1 (fundamental o mode) radiation is reabsorbed at lower altitudes in the source region. The ray paths of z mode waves, on the other hand, are restricted to a small range of altitudes about the source height and can thus continue to grow to high levels without encountering strong absorption. The proposed mechanism is indirect because the trapped z mode waves must undergo repeated nonlinear wave coalescences in order to produce second- and higher-harmonic electromagnetic waves which then propagate to the ground. A theory for emission at cyclotron harmonics due to the repeated coalescence f z mode waves was first proposed in a model to explain second-harmonic [*Melrose and Dulk*, 1984] and higher-harmonic emission bands [*Melrose*, 1991] in solar microwave spike bursts generated in the solar corona. A later model for solar spike bursts, based on the nonlinear coalescence of electron-cyclotron maser emission (ECME)-generated electrostatic Bernstein waves, better accounts for the observed noninteger ratios between the multiple-frequency bands and their comparable intensities [*Willes and Robinson*, 1996, 1997]. Indirect generation mechanisms to produce second- and third-harmonic auroral roar emissions have also been proposed by other authors. *Weatherwax et al.* [1995] ruled out nonlinear wave-wave interactions between lower hybrid and upper hybrid waves to produce electromagnetic waves because of the low power produced. A speculative remark on the possibility of repeated z mode wave coalescence being responsible for these emissions was also made by Weatherwax et al.

The following sections we calculate the spatial growth rates for the competing wave modes and demonstrate the viability of a z mode maser for plasma parameters relevant to the bottomside ionosphere. The conditions under which z mode growth is favored are discussed, along with the dependence on the functional form and temperature of the unstable electron population. We then perform ray-tracing calculations in a model density cavity in the lower ionosphere, for a particular auroral roar event in which both second- and third-harmonic emissions are present. We demonstrate both the reabsorption of maser-generated X2 waves and the "trapping" of maser-generated z mode waves in the source region. Finally, we discuss the nonlinear coalescence processes required to produce the observed radiation, in particular, the requirements of energy and momentum conservation and a comparison of the inferred brightness temperatures with observations.

2. Comparison of Spatial Growth Rates

In this section, cyclotron-maser growth rates are compared for X2, X3, O1, and Z1 waves. We choose a ring-like, hot electron distribution to model the trapped electron population as the source of free energy for maser emission. By comparing the X2 and X3 spatial growth rates, we argue that the X3 growth rate is too low to adequately explain observations of simultaneous second- and third-harmonic auroral roar.

The temporal growth rate of the emitted waves in mode M, defined so that the energy in the waves grows as $e^{\Gamma_M(\mathbf{k})\,t}$, is given by [*Melrose*, 1986]

$$\Gamma_M(\mathbf{k}) = \sum_{s=-\infty}^{\infty} \int d^3\mathbf{u}\, \frac{\hbar\, w_M(\mathbf{k},\mathbf{u},s)}{mc} \times \left(\frac{s\Omega_e}{cu_\perp} \frac{\partial}{\partial u_\perp} + k_\parallel \frac{\partial}{\partial u_\parallel} \right) f(\mathbf{u}) \,, \quad (1)$$

where $f(\mathbf{u})$ is the distribution function of electrons and the normalized momentum $\mathbf{u}$ is defined in terms of the velocity $\mathbf{v}$ and momentum $\mathbf{p}$ of the electrons by

$$\mathbf{u} = \frac{\mathbf{p}}{mc} = \frac{\gamma \mathbf{v}}{c} \,, \quad (2)$$

and the probability function $w_M(\mathbf{k},\mathbf{u},s)$ satisfies

$$w_M(\mathbf{k},\mathbf{u},s) = \frac{2\pi\gamma q^2 R_M(\mathbf{k})}{\varepsilon_0 \hbar |\omega_M(\mathbf{k})|} \left|\mathbf{e}^*_M(\mathbf{k}).\mathbf{V}(\mathbf{k},\mathbf{u},s)\right|^2 \times \delta\left(\gamma\omega - s\Omega_e - k_\parallel c\, u_\parallel\right) \,, \quad (3)$$

for the Lorentz factor $\gamma = (1+u^2)^{1/2}$, $R_M(\mathbf{k})$ is the ratio of electric to total energy (for mode M), $\omega_M(\mathbf{k})$ is the wave frequency, $\mathbf{e}_M(\mathbf{k})$ is the polarization vector, and $\mathbf{V}(\mathbf{k},\mathbf{u},s)$ is the velocity function for a spiraling charge [*Melrose*, 1986].

The strongest contribution to the growth rate is provided by the $\partial f(\mathbf{u})/\partial u_\perp$ term in (1), and positive wave

growth thus requires a positive slope in perpendicular velocities. *Omidi et al.* [1984] identified three possible sources of free energy in the measured electron distribution in the source region of maser-produced, auroral kilometric radiation (at significantly higher altitudes than for the auroral roar sources). The two main sources are a loss cone and a "bump" feature. The loss cone distribution is produced by magnetic mirroring and therefore should also exist at lower altitudes (provided that the source is above the mirror point). The bump component is composed of electrons trapped between the upward directed mirror and the downward force on electrons exerted by parallel electric fields [Louarn et al., 1990]. This is modeled here by the ring-like distribution used by *Yoon et al.* [1996], with

$$f(u,\beta) = A \exp\left[-\frac{(u-u_0)^2}{\alpha^2} - \frac{\cos^2\beta}{\delta^2}\right], \qquad (4)$$

where β is the pitch angle, and A is the normalization factor, with $A^{-1} = \pi^{3/2}\alpha^3\,\delta\,\mathrm{erf}(1/\delta)\,P(u_0/\alpha)$, for $P(x) = x\exp(-x^2) + \sqrt{\pi}(x^2+1/2)[1+\mathrm{erf}(x)]$, α is the thermal spread parameter, u_0 defines the perpendicular offset of the distribution, and δ defines the pitch angle width of the distribution.

Yoon et al. [1996] derived a growth rate expression assuming a ring-like, hot electron distribution and a thermal background electron population (which contributes to damping of waves) and assuming the semirelativistic approximation to the Lorentz factor, with $\gamma = 1 + u^2/2$. Their calculation assumed a ratio of hot to thermal electrons of $n_{\mathrm{hot}}/n_{\mathrm{th}} = 10^{-3}$. In this regime, the wave dispersion is dominated by the thermal plasma and cold plasma magnetoionic modes may be assumed. For comparison with the results of *Yoon et al.* [1996],

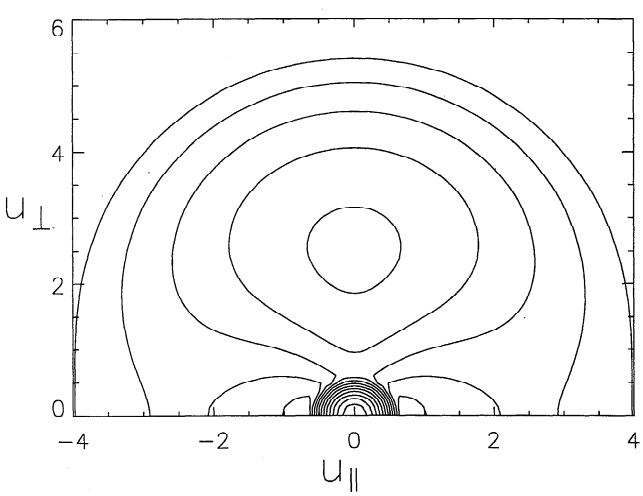

Figure 1. Contour plot (log) of the total electron distribution function, comprising thermal electrons ($\alpha_{\mathrm{th}} = 5.8\times10^{-3}$) and a ring-like component ($\alpha = 0.045, \delta = 0.4, u_0 = 2.5\alpha$). Both $u_\parallel$ and $u_\perp$ (components of $\mathbf{u}$ parallel and perpendicular to $\mathbf{B}$) axes are scaled in units of α. See text for variable definitions.

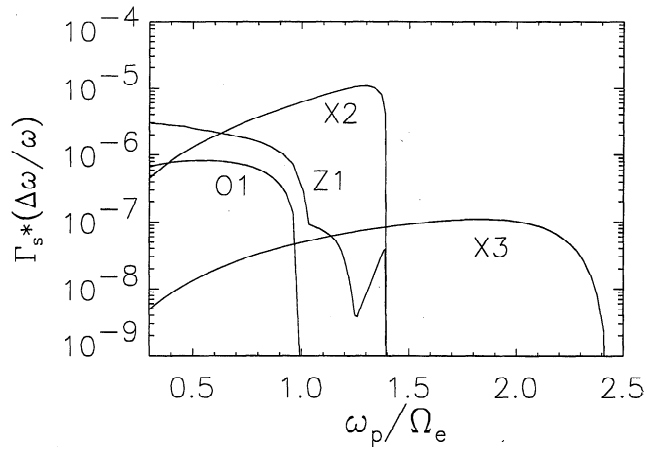

Figure 2. Comparison of the product of spatial growth rate $\Gamma_s = \Gamma c/\Omega_e v_g$, and bandwidth $\Delta\omega/\omega$ for the X2, X3, Z1, and O1 wave modes, maximized over wave frequency and angle, as a function of ω_p/Ω_e.

we assume the same electron distribution, with similar parameters: $\alpha_{\mathrm{hot}} = 0.06$ (2 keV electrons), $\delta = 0.4$, $u_0 = 2.5\alpha$, and $\alpha_{\mathrm{th}} = 5.8\times10^{-4}$ (8.6×10^{-2} eV electrons). For the Maxwellian component, $u_0 = 0$ and the limit $\delta \to \infty$ is taken. The total electron distribution (thermal and trapped) is shown in Figure 1 (with an exaggerated thermal temperature for clarity).

The growth rates for each mode vary most markedly with the ratio of frequencies, ω_p/Ω_e. Figure 2 shows the product of the spatial growth rate (temporal growth rate divided by group speed of waves) and the bandwidth $\Delta\omega/\omega$, maximized over wave frequency and angle, as a function of ω_p/Ω_e. The product of the spatial growth rate and the bandwidth gives the best measure of the amplification rate for when the maser does not saturate and the electron cyclotron frequency varies with height [*Hewitt et al.*, 1983; *Melrose et al.*, 1984; *Melrose*, 1991]. The spatial growth is more relevant in this context because z mode waves have significantly lower group speeds than X2, X3, or O1 waves (where the group speed is close to c). This is important if growth occurs in small, localized regions or if the maser operates close to a state of marginal stability, where growth rates are small and the maser does not saturate [*Robinson*, 1991; *Melrose*, 1991], so that the z mode waves spend more time in the growth region and hence experience more amplification. It is also necessary to take into account the bandwidth because the local cyclotron frequency varies with altitude and the waves can move out of the region in which they can grow. Additionally, wave absorption from the background Maxwellian is included in the growth rate calculation. For $\omega_p/\Omega_e \approx 0.1$, the fundamental x mode has the highest growth rate (not shown in Figure 2). However, in the bottomside ionosphere, only values for ω_p/Ω_e above 0.3 are observed [*Benson and Wong*, 1987; *Yoon et al.*, 1996]. Figure 2 shows that Z1, O1, and X2 waves have comparable growth rates for $0.3 \lesssim \omega_p/\Omega_e \lesssim 0.8$. The reason

for the discontinuity in the slope of the Z1 curve near $\omega_p/\Omega_e \approx 1$ is discussed in Section 3. *Weatherwax et al.* [1995] showed that all these modes have growth rates that exceed the collisional damping rate in the lower ionosphere. Results similar to Figure 2 are obtained if a loss cone electron distribution is assumed in place of the ring distribution. In both cases the X3 spatial growth rate is at least 2 to 3 orders of magnitude below the X2 growth rate. Ray-tracing calculations of the path of propagation of maser-emitted X2 and X3 waves in a model cavity density structure [*Yoon et al.*, 1996] predict comparable path lengths (to within a factor of 2) for the propagating X2 and X3 waves in the source region. It is thus unlikely that the third-harmonic component of auroral emission is produced by direct X3 emission because it is often observed at comparable intensities to the second-harmonic component [*Weatherwax et al.*, 1995], whereas the theory suggests that the product of spatial growth rate and path length is considerably lower for X3 than X2 waves. The X2 spatial growth rate is itself quite low, and many passes through the source region (after reflections from the cavity walls) are required to produce several e-foldings of growth [*Yoon et al.*, 1996].

If the assumption $n_{\rm hot}/n_{\rm th} \ll 1$ is not valid in the source region, then the X2 and O1 spatial growth rates are significantly higher (relative to the Z1 and X3 rates) owing to slower group speeds as a result of splitting of the dispersion relation [*Winglee*, 1985]. This further widens the gap between the spatial growth rates of the X2 and X3 waves.

3. Ray Tracing of Maser-Generated Waves

In this section we report on the results of ray-tracing calculations, modeling a particular auroral roar event in which simultaneous second- and third-harmonic components were observed. We first show that X2 rays encounter a strong absorption region just below their source altitude and thus cannot propagate to a ground-based observer. This conclusion extends also to X3 and O1 waves. We then demonstrate that, for the appropriate parameters, z mode waves can be trapped in the source region and thus grow to high levels before being damped. These z mode waves can coalesce to produce second- and third-harmonic electromagnetic waves, as discussed in section 4.

The numerical ray-tracing calculations discussed in this section are based on the geometric optics approximation, which is valid in our model because the wavelength of the propagating waves is well below the other spatial scales in the source region. It is possible, however, that in the auroral ionosphere, smaller spatial

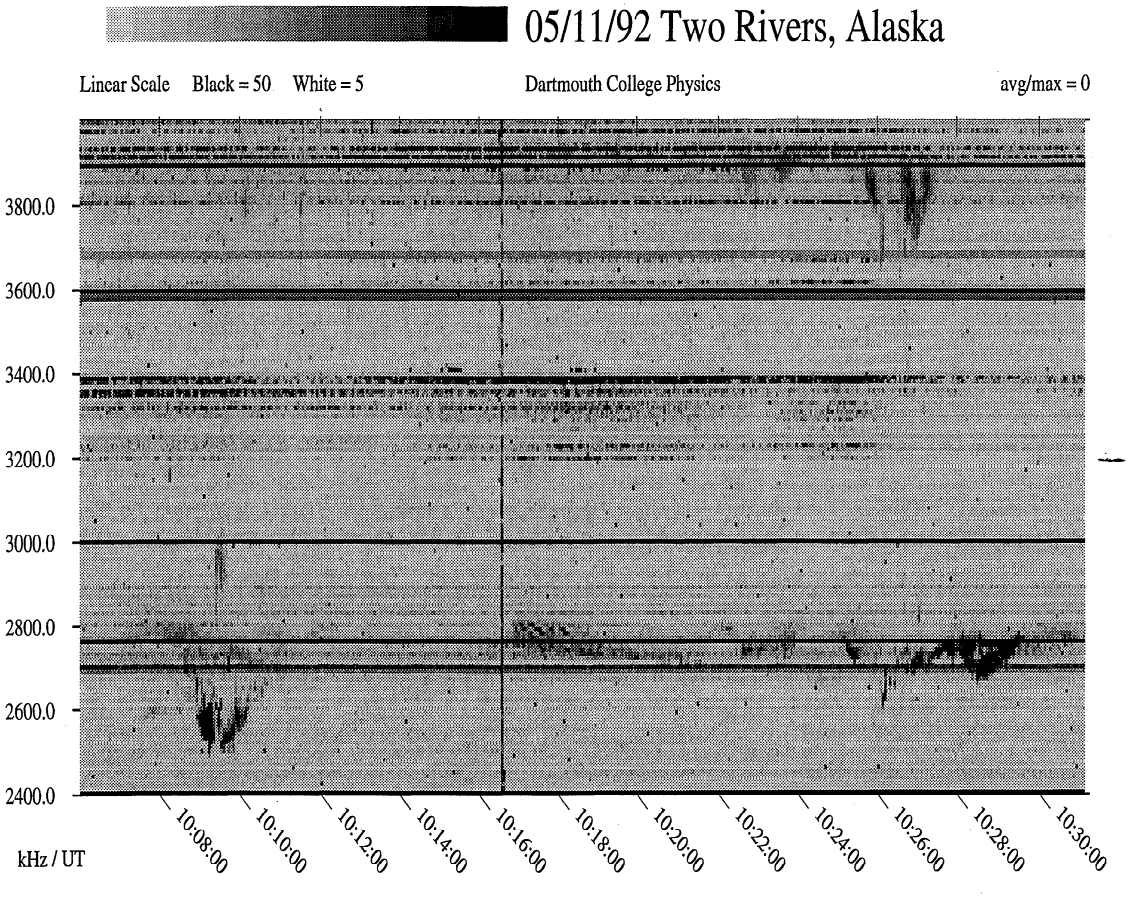

Figure 3. Dynamic spectrum for an auroral roar event with both second- and third-harmonic components [from *Weatherwax et al.*, 1993a,b].

structures may exist where the geometric optics approximation is not valid. The ray equations are

$$\frac{d\mathbf{r}}{dt} = \frac{\partial\omega(\mathbf{k},\mathbf{r})}{\partial\mathbf{k}}, \quad (5)$$

$$\frac{d\mathbf{k}}{dt} = -\frac{\partial\omega(\mathbf{k},\mathbf{r})}{\partial\mathbf{r}}, \quad (6)$$

where t is a parametric representation of the points along the wave path, defined so that the tangent to the ray path is equal to the group velocity, with $\mathbf{v}_g(\mathbf{k},\mathbf{r}) := \partial\omega(\mathbf{k},\mathbf{r})/\partial\mathbf{k}$. The ray-tracing code used here is an adaptation of an earlier code formulated to study the propagation and absorption effects of maser-generated fundamental x and o mode radiation in a coronal flux loop [*Kuncic and Robinson*, 1993].

In our ray-tracing analysis we choose to model a particular auroral roar event (May 11, 1992, between 1020 and 1030 UT [*Weatherwax et al.*, 1993a, b, Figure 3]) that exhibits simultaneous second and third harmonics. The dynamic spectrum for this event is shown in Figure 3. Assuming dipolar variation in the magnetic field strength and that the two components are emitted at twice and three times the local electron cyclotron frequency, the second-harmonic component (2.75 MHz) is emitted approximately at a height of 200 km, and the third harmonic (3.9 MHz) is emitted at a height of 300 km. It then follows that the source region (where the trapped energetic electrons exist) extends at least 100 km in the vertical direction. A schematic representation of the source region for auroral roar emissions is shown in Figure 4.

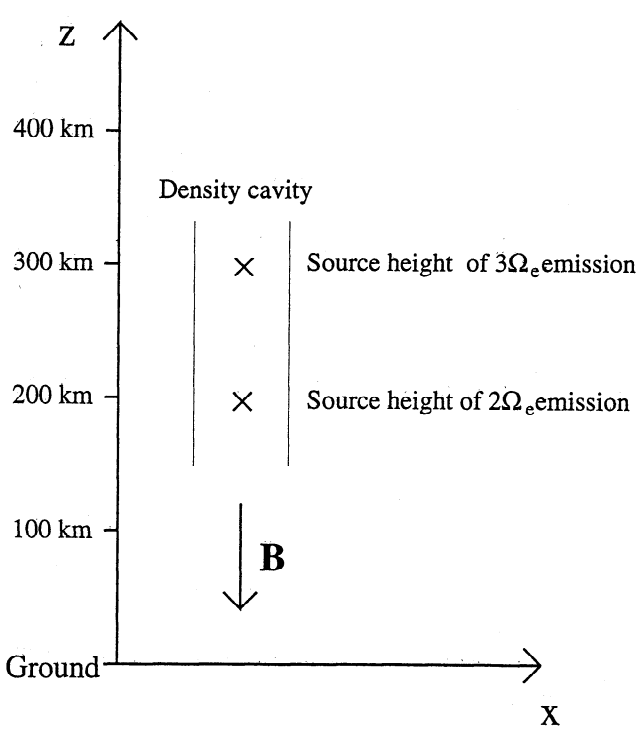

Figure 4. Schematic picture of the auroral roar source region, with the second- and third-harmonic components generated at different altitudes in a field-aligned density cavity.

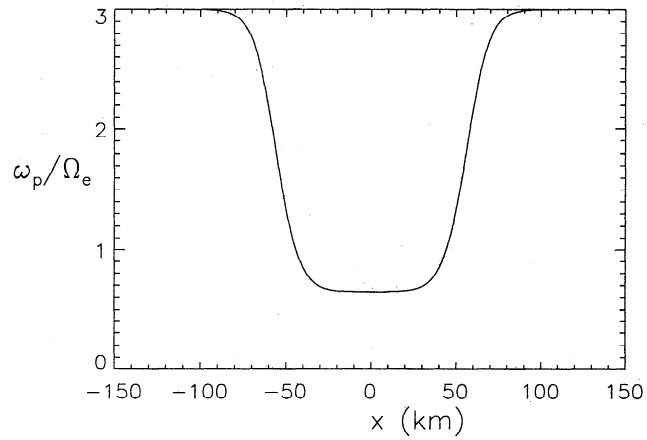

Figure 5. Model density cavity profile, using (7) with the following parameters: $a = 0.65, b = 3, w = 6 \times 10^4$ m, and $\Delta = 1.2 \times 10^4$ m.

In the lower ionosphere the ratio ω_p/Ω_e exceeds unity (typically $3-5$). For such high values of ω_p/Ω_e, none of the X2, O1, or Z1 modes is unstable, as shown in Figure 2. However, lower values of ω_p/Ω_e exist in field-aligned density cavities, where ω_p/Ω_e can be low enough for any of the X2, X3, O1, or Z1 waves to grow. In the model of *Yoon et al.* [1996], such cavities also act as a waveguide, allowing X2 and X3 waves to propagate to the ground. We adopt a model similar to that of *Yoon et al.* [1996] for the density cavity profile, based on ground-based incoherent scatter radar observations [*Doe et al.*, 1993, 1995], with

$$\omega_p = \left[\frac{b^2-a^2}{2}\left(\tanh\frac{|x|-w}{\Delta} + \frac{b^2+a^2}{b^2-a^2}\right)\right]^{1/2}, \quad (7)$$

where b is the value of ω_p outside the cavity (which varies with the ionospheric density profile; see Figure 6), a is the value of ω_p at the center of the cavity (which remains constant), w is the width of the cavity, and Δ determines the slope of the cavity. Equation (7) is only well defined for $b > a$. Figure 5 shows a typical density profile for $b = 3$, $a = 0.65$, $w = 6 \times 10^4$ m, and $\Delta = 1.2 \times 10^4$ m. In addition, we incorporate into our model the variation of the local cyclotron frequency, which slowly increases with decreasing altitude, and the variation of the plasma frequency with height (which determines parameter b in (7)) using the international reference ionosphere [*Bilitza et al.*, 1993] for the time and coordinates of the *Weatherwax et al.* [1993a] event. Figure 6 displays the variation of electron density with height used in our ray-tracing model. In the ray-tracing calculations presented below we assume that the hot, trapped electron distribution only exists in the inner region of the density cavity, for $|x| \leq 3 \times 10^4$ m.

3.1. X2 Waves

Figure 7 shows contours of constant growth (solid contours) and absorption (dashed contours) rate for maser emission in the second-harmonic x mode, assum-

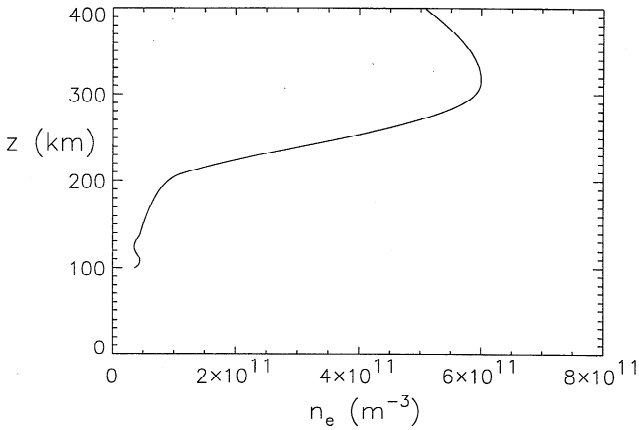

Figure 6. Variation of electron density with height at the time and coordinates of the May 11, 1992, event, calculated using the international reference ionosphere [*Bilitza et al.*, 1993].

ing $\omega_p/\Omega_e = 0.65$ in the source region at the center of the cavity ($a = 0.65$ in (7)). Each growth contour is separated by $0.2\Gamma_{\max}$ (where $\Gamma_{\max}$ is the maximum growth rate), and each absorption contour is separated by $\Gamma_{\max}$. The fastest growing waves are produced perpendicular to the ambient magnetic field. We launch X2 waves at $\omega/\Omega_e = 1.99$ and $\theta = 90°$ (i.e., waves with the maximum growth rate in Figure 7) at an altitude of 200 km (the approximate source height of X2 waves) and trace the ray path in Figure 8. The ray path is solid where the integrated spatial growth rate $G(r) = \int_{r_0}^{r} dl\, \Gamma/v_g$ is positive and dashed where it is negative. The ray paths are terminated where the magnitude of negative G exceeds $\max(G)$. In Figure 8, $\max(G) = 4 \times 10^{-2}$. The most striking feature of Figure 8 is that the X2 waves are strongly absorbed within tens of kilometers

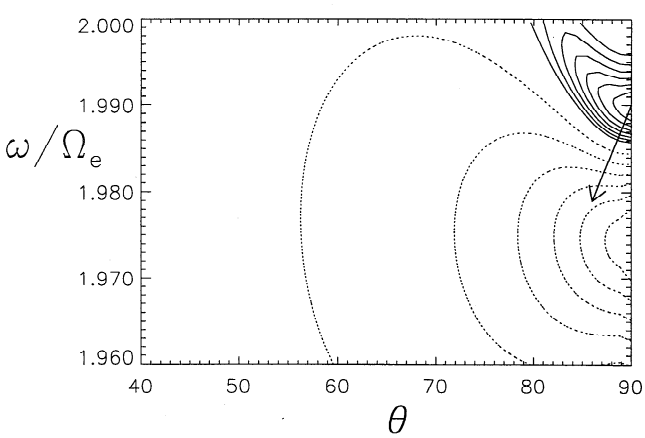

Figure 7. Contour plot of temporal growth rate for second-harmonic x mode waves ($\omega_p/\Omega_e = 0.65$, $\alpha = 0.045$, $\delta = 0.4$, $u_0 = 2.5\alpha$, and $\alpha_{\rm th} = 5.8 \times 10^{-4}$). The contour spacing is set at $0.2\Gamma_{\max}$, where $\Gamma_{\max}$ is the maximum growth rate, for the growth (solid curves) contours and $\Gamma_{\max}$ for the absorption (dashed curves) contours.

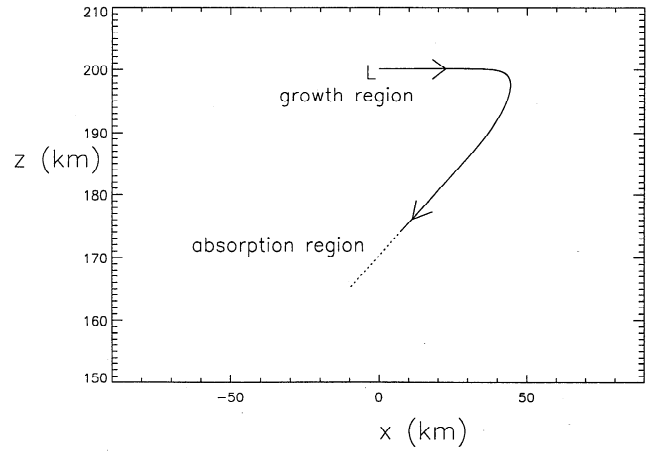

Figure 8. Ray path of X2 waves in a density cavity. The X2 ray is launched at an altitude of $z = 200$ km (position labeled L) at the center of the cavity ($x = 0$) at an angle $\theta = 90°$ to $\mathbf{B}$, and $\omega/\Omega_e = 1.99$. The ray path is solid where the integrated spatial growth rate is positive and dashed where it is negative.

below the launch height. This can easily be explained with reference to Figure 7. As the waves propagate to lower altitudes they effectively move in the direction of the arrow drawn in Figure 7 because ω/Ω_e decreases as Ω_e increases with decreasing altitude and the rays become more field-aligned (decreasing θ) as they propagate down the density cavity. Thus they soon enter a strong absorption region. A semiquantitative estimate of the vertical distance traveled by the X2 waves before they are absorbed is made by taking

$$\frac{\Delta\Omega_e}{\Omega_e} \approx \Delta\left(\frac{\omega}{\Omega_e}\right), \tag{8}$$

where $\Delta\Omega_e/\Omega_e$ is the change in local cyclotron frequency with height and $\Delta(\omega/\Omega_e)$ is the change in effective frequency required for waves to be reabsorbed,

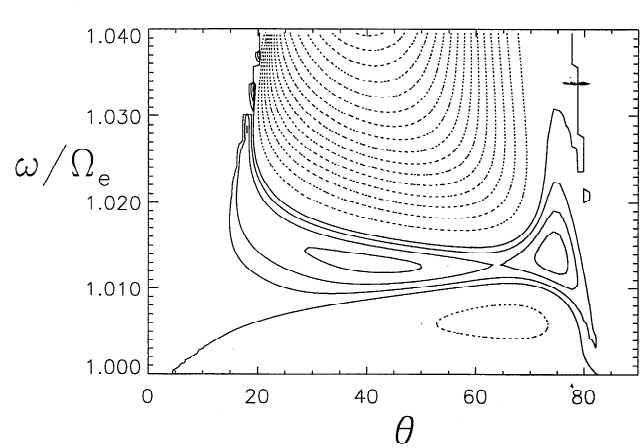

Figure 9. Contours of constant temporal growth rate for fundamental z mode waves. Parameters and contour spacings are the same as in Figure 7.

which is less than or equal to the cyclotron line width (in Figure 7, $\Delta(\omega/\Omega_e) \approx 0.005$). Assuming a dipolar magnetic field, this gives

$$\frac{3\Delta z}{R_E + z} \approx 0.005\,, \tag{9}$$

where R_E is the Earth's radius, z is the altitude, and Δz is the change in altitude. This yields $\Delta z \approx 10$ km, in approximate agreement with the ray-tracing calculations. X3 and O1 waves are similarly reabsorbed at lower altitudes. Absorption from the thermal (background) plasma has a negligible effect here because the absorption band is extremely narrow (a result of the relatively low temperature of the background electrons).

For the second- and third-harmonic event considered here, the source extends over a vertical range of greater than 100 km. Hence any radiation observed at the ground would need to be generated close to the bottom of the source region because radiation generated at higher altitudes would be reabsorbed. There are two problems associated with this conclusion. The first is that it is unlikely that X2 waves generated at the bottom of the source region can "bounce" back through the source region a sufficient number of times to grow to observable levels. In addition, X3 radiation is produced at (significantly) higher altitudes (in the present example, 100 km above the X2 source). The X3 radiation would be totally absorbed in this case. Hence it is very difficult to explain simultaneous second- and third-harmonic auroral roar emissions with a direct X2 and X3 maser model.

It is an interesting point that wave damping due to the fundamental absorption band is not usually an inhibitor to ECME. In other applications, such as solar microwave spike bursts [*Holman et al.*, 1980; *Melrose*, 1994], jovian decametric radiation [*Hewitt et al.*, 1982], and auroral kilometric radiation [*Wu and Lee*, 1979; *Melrose et al.*, 1982], the waves propagate away from the source in the direction of decreasing Ω_e (with decreasing magnetic field strength). Hence ω/Ω_e increases and the waves do not encounter the fundamental absorption band (at lower values of ω/Ω_e than the growth band). Absorption of radiation at higher harmonics is possible, however. It is a well recognized problem that fundamental x and o mode radiation generated in the solar corona to produce solar microwave spike bursts encounters a strong second-harmonic absorption layer where the local cyclotron frequency is half the wave frequency (Ω_e at the source) [*Holman et al.*, 1980; *Melrose and Dulk*, 1982, 1984; *Kuncic and Robinson*, 1993].

3.2. The z Mode Waves

Figure 9 shows growth rate contours for Z1 waves, for the same parameters and with similar contour spacings as Figure 7. In this case, maximum wave growth occurs for $\omega/\Omega_e \approx 1.13$ and over a wide range of angles. There are two maxima as a function of angle, one near $\theta = 80°$ and the other near $\theta = 40°$. The discontinuity in the Z1 maximum growth rate curve as a function of ω_p/Ω_e in Figure 2 is a result of the absolute maximum shifting from the low-angle maximum for $\omega_p/\Omega_e < 1$ to the higher-angle maximum (close to 90°) for $\omega_p/\Omega_e > 1$. For a loss cone electron distribution, nearly perpendicular z mode waves are produced, but are not produced at lower angles [*Hewitt et al.*, 1983]. The ray-tracing results for Z1 waves launched close to 90° to **B** are similar to the X2 results presented above; the Z1 rays soon propagate into a strong absorption region. However, at lower angles the propagation properties of z mode waves are somewhat different because the group velocity vector $\mathbf{v}_g$ points in a different direction to **k**. Figure 10 displays θ_g (the angle between $\mathbf{v}_g$ and **B**) as a function of θ, for $\omega/\Omega_e = 1.13$ in Figure 9. Note that waves launched at $\theta = 40°$ actually propagate nearly perpendicularly to **B**. In this way the waves can propagate and grow in the source region and not significantly vary in altitude, and hence they avoid regions of strong damping. This is illustrated in Figure 11 for the ray path of z mode waves with $\omega/\Omega_e = 1.13$, $\theta = 32°$, and $z = 200$ km (note that the angles are not accurately represented in Figures 8 and 11 owing to the different x and y axis scales). For this particular example the rays pass through the center of the cavity 10 times before they enter a region of strong damping (at $z \approx 280$ km), with $\max(G) = 3 \times 10^{-2}$. Thus, in order to achieve several e-foldings of growth, hundreds of crossings are required. This is certainly possible but requires the right conditions, which perhaps may explain why auroral roar events occur only at particular times (for example, at the onset of auroral substorms). It is important to also note that the ray paths of z mode waves (and hence the number of crossings in the growth region) are sensitively dependent on the source parameters and the frequency and launch angle of the z mode waves. In Figure 11 the z mode waves maintain an angle similar to **B** as their initial value of θ throughout the source region (except

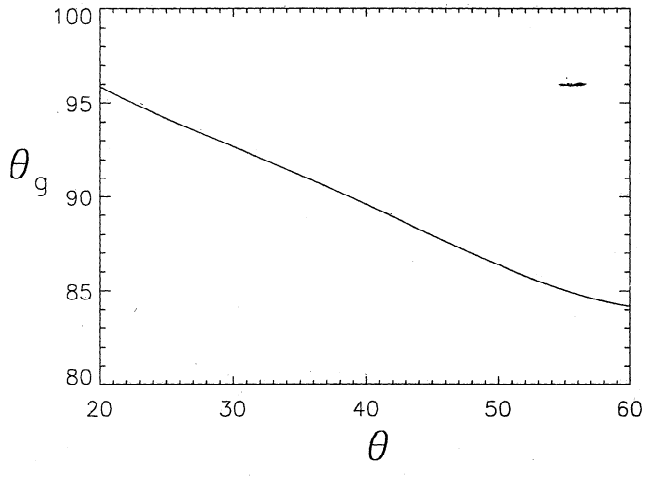

Figure 10. Group velocity angle θ_g (the angle between $\mathbf{v}_g$ and **B**) for Z1 waves (with $\omega/\Omega_e = 1.013$ in Figure 9).

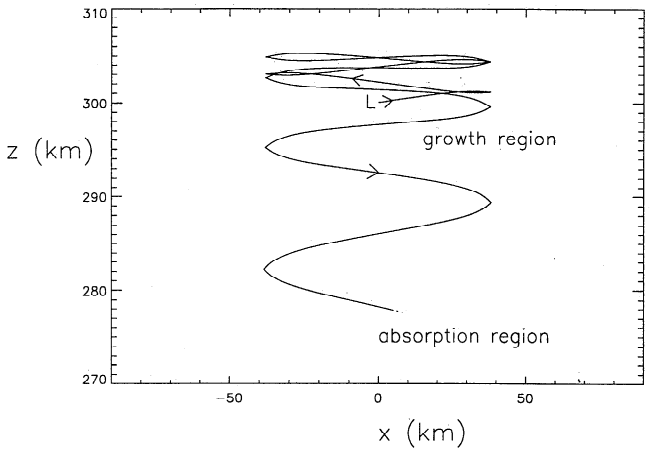

Figure 11. Ray path of Z1 waves launched (at position L) at an altitude of 300 km, with $\omega/\Omega_e = 1.012$ and $\theta = 29°$. Note that the ray trajectory initially is directed in an upward direction, even though the wave vector is directed downward. The waves eventually propagate to a strong absorption band below $z = 280$ km.

at the edges of the cavity, where the waves are reflected) and remain within the range of angles of $25° \lesssim \theta \lesssim 45°$. In section 4 we show that these z mode waves can satisfy the kinematic constraints required for two z mode waves to coalesce to produce higher harmonic waves. For lower temperatures of the ring-like electron distribution, the frequency of the maximum growth rate in Figure 9 moves down to lower values of ω/Ω_e, and the corresponding angles at which $\theta_g \approx 90°$ ($\approx 40°$ in Figure 10) move down to slightly lower values of θ. In this way, the mechanism works over a broad range of temperatures and produces z mode waves over a broad range of angles.

Melrose [1991] discussed the group properties of nearly perpendicular, upward directed z mode waves generated from an upward directed loss cone distribution in the context of solar microwave spike bursts in the solar corona. For these waves the group velocity vector has a small downward component that decreases as the waves propagate toward higher values of Ω_e (and approaches zero as the wave frequency approaches the local cyclotron frequency). In this way the z mode waves collect in a layer just below their generation height. However, it is unlikely that loss cone-generated z mode waves contribute toward ground-directed auroral roar emissions because the wave vectors always have an upward directed component, and hence any product wave from nonlinear wave coalescence is also directed away from the ground.

4. Coalescence of z Mode Waves

The mechanism by which second- and third-harmonic emission is produced from a trapped source of z mode waves is via two nonlinear wave coalescence processes [*Melrose*, 1991]. First, two z mode waves coalesce to produce a second-harmonic (x mode) wave: $Z + Z' \rightarrow X_2$. Then, a second-harmonic (x mode) wave coalesces with a z mode wave to produce a third-harmonic (x mode) wave: $Z + X_2 \rightarrow X_3$. There are two possible sources of X_2 waves for this process: from direct maser emission and from the first coalescence process.

Conservation of energy and momentum must be satisfied for the wave quanta in each process. Conservation of energy is equivalent to conservation of wave frequency, with $\omega_Z + \omega_{Z'} = \omega_{X2}$ for the first three-wave process. This is easily satisfied, where two z mode waves generated at the cyclotron frequency coalesce to produce an x mode wave at $2\Omega_e$. Conservation of momentum is equivalent to conservation of wave vectors, with $\mathbf{k}_Z + \mathbf{k}_{Z'} = \mathbf{k}_{X2}$. For the example given in section 3.2, z mode waves are typically produced in the range of wave numbers $\Omega_e/c \lesssim k_Z \lesssim 2.5\Omega_e/c$ and lie in the range of angles $25° \lesssim \theta_Z \lesssim 45°$ to $\mathbf{B}$. Two z mode waves within these ranges can coalesce to produce an X2 wave, with $k_{X2} \approx 2\Omega_e/c$, as shown in the first wave vector addition diagram in Figure 12. For the third-harmonic three-wave process, $\mathbf{k}_{X2} + \mathbf{k}_Z = \mathbf{k}_{X3}$, with $k_{X2} \approx 2\Omega_e/c$ and $k_{X3} \approx 3\Omega_e/c$. The second wave vector diagram in Figure 12 illustrates the wave vector addition for z mode waves satisfying the kinematic constraints for the third-harmonic process. In both cases, the product second- and third-harmonic waves make relatively small angles with $\mathbf{B}$ and can propagate directly to the ground with zero or few bounces required from the cavity walls.

Further harmonics may be generated by coalescence of an nth harmonic electromagnetic wave with a z mode wave to produce a $(n+1)$th harmonic transverse wave, namely, $X_n + Z_1 \rightarrow X_{n+1}$. However, production of higher harmonics beyond the third is unlikely owing to the finite energy available in the maser source [*Melrose*, 1991]. The condition for saturation of the second-

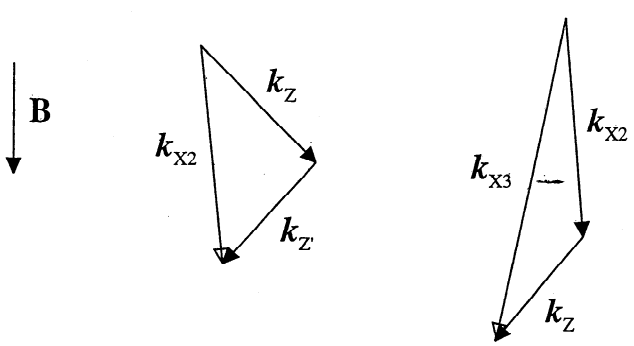

Figure 12. Wave vector addition for the three-wave coalescence processes to produce second- and third-harmonic waves. (left) Coalescence of two z mode waves (solid arrows), with $k_Z = 1.5\Omega_e/c$, $\theta_Z = 45°$, $k_{Z'} = 1.25\Omega_e/c$, and $\theta_{Z'} = -42°$, to produce a second-harmonic x mode wave (open arrow), with $k_{X2} = 2\Omega_e/c$, $\theta_{X2} = 5°$. (right) Coalescence of an X2 wave, with $k_{X2} = 2\Omega_e/c$, and $\theta_{X2} = 5°$, and a z mode wave, with $k_Z = 1.25\Omega_e/c$, and $\theta_Z = -40°$ to produce a third-harmonic wave, with $k_{X3} = 3\Omega_e/c$, and $\theta_{X3} = -12°$.

harmonic coalescence process is determined by comparison of the time for saturation of the three-wave process with the time taken for the second-harmonic ray to propagate out of the region in which it is generated. [*Melrose*, 1991], with

$$\frac{\pi^2 \alpha\, n_{\rm hot} r_0 L_B\, v_{\rm hot}^2 \omega_p}{\sqrt{2}\,\Omega_e^2 c} \gtrsim 1\,, \qquad (10)$$

where α is the maser efficiency, r_0 is the classical radius of the electron, L_B is the characteristic length over which B changes in magnitude (here $L_B = 3 \times 10^5$ m), and $v_{\rm hot}$ is the characteristic speed of the hot electrons. The free energy available for maser emission is the fraction α of the energy density of the hot electrons and is equal to $\alpha n_{\rm hot} m_e v_{\rm hot}^2/2$. For $\omega_p/\Omega_e = 0.5$, $\Omega_e = 10^7\,{\rm s}^{-1}$, $v_{\rm hot} = 4 \times 10^7\,{\rm m\,s}^{-1}$, and $n_{\rm hot}/n_{\rm th} = 10^{-3}$, the left-hand side of (10) equals $\alpha \times 10^{-2}$. The maser efficiency α is, at most, a few percent. Hence the condition for saturation (10) is not satisfied. It follows that the third-harmonic process also does not saturate. We derive the brightness temperature from auroral roar electric field observations as the radio flux density per solid angle of antenna beam, with

$$k_B T = \frac{c^2}{2f^2}\frac{F(f)}{\Delta\Omega}\,, \qquad (11)$$

where f is the frequency; $F(f)$ is the radio flux density, with $F(f) = \frac{1}{2}\epsilon_0 E^2 c/\Delta f$; and $\Delta\Omega$ is the solid angle subtended by the (unresolved) radio source, with $\Delta\Omega \approx x^2/(4\pi r^2)$, for a source of dimension x and distance r from the ground. *Weatherwax et al.* [1993a] reported amplitudes of 20 dB above a 5×10^{-9} V/m$\sqrt{\rm Hz}$ threshold. Assuming $x = 10^4$ m and $r = 4 \times 10^5$ m, then $T \approx 1.5 \times 10^{13}$ K. Source sizes larger than 10^4 m would yield lower temperatures. From the above parameters the saturation brightness temperature for z mode wave coalescence satisfies [*Melrose*, 1991] $T_2 = T_z = 5\alpha \times 10^{21}$ K, which well exceeds the temperature inferred from observations. This is consistent with the coalescence process not reaching saturation.

5. Discussion and Summary

In this paper we have presented a model for second- and third-harmonic auroral roar emissions, where z mode waves are generated in field-aligned density cavities in the lower ionosphere via the process of electron cyclotron maser emission. The following conclusions are made: (1) In the range of values of ω_p/Ω_e relevant to the bottomside ionosphere ($\omega_p/\Omega_e \geq 0.3$), the most strongly growing modes are the fundamental z mode (Z1), fundamental o mode (O1), and the second-harmonic x mode (X2). The third-harmonic x mode (X3) growth rate is significantly weaker than for these modes, with spatial growth rates 2 to 3 orders of magnitude lower (for both types of electron distribution considered and over a wide range of electron energies). This is not consistent with the comparable path lengths for X2 and X3 waves in a density cavity determined from ray-tracing calculations and the observed comparable intensities of the second- and third-harmonic components. (2) Ray-tracing calculations of second harmonic x mode (X2) waves in a model density cavity reveal that these waves are reabsorbed at lower altitudes, where the local electron cyclotron frequency is higher. This is also the case for X3 and O1 waves. Such reabsorption of maser-generated radiation provides a strong argument against direct maser emission in X2 and X3 waves propagating from the source region to a ground-based observer. (3) Maser-generated z mode waves can remain at the source altitude over many "bounces" between the cavity walls. Several hundred passes through the source region are required in order to grow to significant levels. Trapping of z mode waves could be enhanced by local irregularities in the ionospheric plasma density (on the scale of tens of kilometers) that act as a "density trap". (4) For typical parameters relevant to the source region in the lower ionosphere, the trapped z mode waves can satisfy the kinematic constraints required for the nonlinear wave-wave coalescence processes ($Z + Z' \rightarrow X_2$ and $X_2 + Z \rightarrow X_3$) to proceed. This is not the case for z mode waves generated from an upward directed loss cone distribution, which only produces upward directed product waves. (5) Semiquantitative calculations suggest that neither of the wave coalescence processes reaches saturation levels.

The case for a z mode maser being the responsible mechanism for the auroral roar emissions is strengthened by low-altitude rocket observations of z mode waves in the vicinity of the auroral roar source region [*Benson and Wong*, 1987] between the cyclotron and upper hybrid frequencies.

One question that remains to be answered is why the third-harmonic source is statistically at higher altitudes than the second-harmonic source over many auroral roar events [*Weatherwax et al.*, 1995]. It is also pertinent to comment on the fine-frequency structure in auroral roar emissions. Higher-resolution frequency observations of auroral roar events [*LaBelle et al.*, 1995] reveal that auroral roar emissions comprise many narrowband features that can drift in frequency (in either direction). Narrowband emissions are consistent with small source sizes, which is the case if z mode waves are trapped in a small range of altitudes in the source region. It may also be that only z mode waves in an even narrower range of altitudes satisfy the necessary kinematic conditions to achieve nonlinear wave coalescence. The observed drift in frequency in the narrowband features may be due to moving density structures in the lower ionosphere [*La Belle et al.*, 1995] that favor the trapping of z mode waves.

Acknowledgments. This work was supported by a PPARC grant. The authors thank Allan Weatherwax for Figure 3 and for helpful comments and Peter Robinson for provision of computer resources at the University of Sydney.

The Editor thanks J. LaBelle, I. H. Cairns, and A. T. Weatherwax for their assistance in evaluating this paper.

References

Benson, R. F., and H. K. Wong, Low-altitude ISIS 1 observations of auroral radio emissions and their significance to the cyclotron maser instability, *J. Geophys. Res., 92*, 1218-1230, 1987.

Bilitza, D., K. Rover, L. Bossy, and T. Gulyaeva, International Reference Ionosphere - Past, present, future, *Adv. Space Res., 13* (3), 3-23, 1993.

Doe, R. A., M. Mendillo, J. F. Vickrey, L. Zanetti, and R. Eastes, Observations of nightside auroral cavities, *J. Geophys. Res., 98*, 293-310, 1993.

Doe, R. A., J. F. Vickrey, and M. Mendillo, Electrodynamical model for the formation of auroral ionospheric cavities, *J. Geophys. Res., 100*, 9683-9696, 1995.

Hewitt, R. G., D. B. Melrose, and K. G. Rönnmark, The loss-cone driven electron-cyclotron maser, *Aust. J. Phys., 35*, 447-471, 1982.

Hewitt, R. G., D. B. Melrose, and G. A. Dulk, Cyclotron maser emission of auroral Z mode radiation, *J. Geophys. Res., 88*, 10,065-10,071, 1983.

Holman, G. D., D. Eichler, and M. R. Kundu, An interpretation of solar flare microwave spikes as gyrosynchrotron masering, in *Radio Physics of the Sun*, edited by M. R. Kundu and T. E. Gergely, p. 457, Int. Astron. Union, 1980.

Horne, R., B., Propagation to the ground at high latitudes of auroral radio noise below the electron gyrofrequency, *J. Geophys. Res., 100*, 14,637-14,645, 1995.

Kellogg, P. J., and S. J. Monson, Radio emissions from the aurora, *Geophys. Res. Lett., 6*, 297-300, 1979.

Kellogg, P. J., and S. J. Monson, Further studies of auroral roar, *Radio Sci., 19*, 551-555, 1984.

Kuncic, Z., and P. A. Robinson, Propagation and absorption of cyclotron maser radiation in solar microwave spike bursts, *Sol. Phys., 145*, 317-338, 1993.

LaBelle, J., M. L. Trimpi, R. Brittain, and A. T. Weatherwax, Fine structure of auroral roar emissions, *J. Geophys. Res., 100*, 21,953-21,959, 1995.

Louarn, P., A. Roux, H. de Féraudy, D. Le Quéau, M. André, and L. Matson, Trapped electrons as a free energy source for the auroral kilometric radiation, *J. Geophys. Res., 95*, 5983-5995, 1990.

Melrose, D. B., *Instabilities in Space and Laboratory Plasmas*, Cambridge Univ. Press, New York, 1986.

Melrose, D. B., Emission at cyclotron harmonics due to coalescence of z mode waves, *Astrophys. J., 380*, 256-267, 1991.

Melrose, D. B., Cyclotron masers and solar spike bursts, *Space Sci. Rev., 68*, 159-170, 1994.

Melrose, D. B., and G. A. Dulk, Radio wave heating of the corona and electron precipitation during flares, *Astrophys. J., 259*, L41-L44, 1982.

Melrose, D. B., and G. A. Dulk, Radio frequency heating of the coronal plasma during flares, *Astrophys. J., 282*, 308-315, 1984.

Melrose, D. B., K. G. Rönnmark, and R. G., Hewitt, Terrestrial kilometric radiation: The cyclotron theory, *J. Geophys. Res., 87*, 5140-5150, 1982.

Melrose, D. B., R. G. Hewitt, and G. A. Dulk, Electron-cyclotron maser emission: Relative growth and damping rates for different modes and harmonics, *J. Geophys. Res., 89*, 897-904, 1984.

Omidi, N., C. S. Wu, and D. A. Gurnett, Generation of auroral kilometric and Z-mode radiation by the cyclotron maser mechanism, *J. Geophys. Res., 89*, 883-895, 1984.

Robinson, P. A., Electron-cyclotron maser emission in solar microwave spike bursts, *Sol. Phys., 134*, 299-314, 1991.

Weatherwax, A. T., J. LaBelle, M. L. Trimpi, and R. Brittain, Ground-based observations of radio emissions near $2f_{ce}$ and $3f_{ce}$ in the auroral zone, *Geophys. Res. Lett., 20*, 1447-1450, 1993a.

Weatherwax, A. T., J. LaBelle, M. L. Trimpi, and R. Brittain, Correction to "Ground-based observations of radio emissions near $2f_{ce}$ and $3f_{ce}$ in the auroral zone", *Geophys. Res. Lett., 20*, 2143, 1993b.

Weatherwax, A. T., J. LaBelle, M. L. Trimpi, R. A. Treumann, J. Minow, and C. Deehr, Statistical and case studies of radio emissions observed near $2f_{ce}$ and $3f_{ce}$ in the auroral zone, *J. Geophys. Res., 100*, 7745-7757, 1995.

Willes, A. J., and P. A. Robinson, Electron-cyclotron maser theory for noninteger-ratio emission frequencies in solar microwave spike bursts, *Astrophys. J., 467*, 465-472, 1996.

Willes, A. J., and P. A. Robinson, Electron-cyclotron maser theory for extraordinary Bernstein waves, *J. Plasma Phys. 58*, 171-191, 1997.

Winglee, R. M., Fundamental and harmonic electron cyclotron maser emission, *J. Geophys. Res., 90*, 9663-9674, 1985.

Wu, C. S., and L. C. Lee, A theory of the terrestrial kilometric radiation, *Astrophys. J., 230*, 621-626, 1979.

Yoon, P. H., A. T. Weatherwax, T. J. Rosenberg, and J. LaBelle, Lower ionospheric cyclotron maser theory: A possible source of $2f_{ce}$ and $3f_{ce}$ auroral radio emissions, *J. Geophys. Res., 101*, 27,015-27,025, 1996.

S. D. Bale, Space Sciences Laboratory, University of California, Berkeley, CA 94720.

Z. Kuncic, Department of Physics and Astronomy, University of Victoria, Victoria, British Columbia, BC V8W 3P6, Canada.

A. J. Willes, Astronomy Unit, School of Mathematical Sciences, Queen Mary and Westfield College, London, E1 4NS, England (e-mail: A.Willes@qmw.ac.uk)

(Received June 2, 1997; revised December 3, 1997; accepted December 4, 1997.)