

The electron–cyclotron maser for astrophysical application

Rudolf A. Treumann

Received: 15 March 2006 / Published online: 5 July 2006
© Springer-Verlag 2006

Abstract The electron–cyclotron maser is a process that generates coherent radiation from plasma. In the last two decades, it has gained increasing attention as a dominant mechanism of producing high-power radiation in natural high-temperature magnetized plasmas. Originally proposed as a somewhat exotic idea and subsequently applied to include non-relativistic plasmas, the electron–cyclotron maser was considered as an alternative to turbulent though coherent wave–wave interaction which results in radio emission. However, when it was recognized that weak relativistic corrections had to be taken into account in the radiation process, the importance of the electron–cyclotron maser rose to the recognition it deserves. Here we review the theory and application of the electron–cyclotron maser to the directly accessible plasmas in our immediate terrestrial and planetary environments. In situ access to the radiating plasmas has turned out to be crucial in identifying the conditions under which the electron–cyclotron maser mechanism is working. Under extreme astrophysical conditions, radiation from plasmas may provide a major energy loss; however, for generating the powerful radiation in which the electron–cyclotron maser mechanism is capable, the plasma must be in a state where release of susceptible amounts of energy in the form of radiation is favorable. Such conditions are

R. A. Treumann (✉)
Geophysics Section, Ludwig-Maximilians-University Munich, Theresienstr. 37-41,
80333 Munich, Germany
e-mail: art@mpe.mpg.de

Present Address:

R. A. Treumann
The International Space Science Institute Bern, Bern, Switzerland

R. A. Treumann
Department of Physics and Astronomy, Dartmouth College, Hanover, NH 03755, USA

realized when the plasma is unable to digest the available free energy that is imposed from outside and stored in its particle distribution. The lack of dissipative processes is a common property of collisionless plasmas. When, in addition, the plasma density becomes so low that the amount of free energy per particle is large, direct emission becomes favorable. This can be expressed as negative absorption of the plasma which, like in conventional masers, leads to coherent emission even though no quantum correlations are involved. The physical basis of this formal analogy between a quantum maser and the electron–cyclotron maser is that in the electron–cyclotron maser the free-space radiation modes can be amplified directly. Several models have been proposed for such a process. The most famous one is the so-called loss-cone maser. However, as argued in this review, the loss-cone maser is rather inefficient. Available in situ measurements indicate that the loss-cone maser plays only a minor role. Instead, the main source for any strong electron–cyclotron maser is found in the presence of a magnetic-field-aligned electric potential drop which has several effects: (1) it dilutes the local plasma to such an extent that the plasma enters the regime in which the electron–cyclotron maser becomes effective; (2) it generates energetic relativistic electron beams and field-aligned currents; (3) it deforms, together with the magnetic mirror force, the electron distribution function, thereby mimicking a high energy level sufficiently far above the Maxwellian ground state of an equilibrium plasma; (4) it favors emission in the free-space RX mode in a direction roughly perpendicular to the ambient magnetic field; (5) this emission is the most intense, since it implies the coherent resonant contribution of a maximum number of electrons in the distribution function to the radiation (i.e., to the generation of negative absorption); (6) it generates a large number of electron holes via the two-stream instability, and ion holes via the current-driven ion-acoustic instability which manifest themselves as subtle fine structures moving across the radiation spectrum and being typical for the electron–cyclotron maser emission process. These fine structures can thus be taken as the ultimate identifier of the electron–cyclotron maser. The auroral kilometric radiation of Earth is taken here as the paradigm for other manifestations of intense radio emissions such as the radiation from other planets in the solar system, from exoplanets, the Sun and other astrophysical objects.

Keywords Electron–cyclotron maser · Non-thermal radiation · Coherent radiation · Radio emissions from magnetized planets · Auroral kilometric radiation · Jupiter radio bursts · Solar radio bursts/spikes · Coherent radiation from stars · Coherent radiation from Blazar jets

1 Introduction

It is common place that astronomy and astrophysics live from the observation of radiation and that there is no window anywhere in astronomy for making observations in situ. Nevertheless, space plasma physics opens up a window in near-Earth space, where on the one hand, some astrophysical concepts of

processes, small-scale models, and theories can be put to test and, on the other hand, new concepts can be developed. These concepts range from the physics of collisionless shocks, over the acceleration of particles, to the merging, breaking-off, and reconnection of magnetic lines of force. And that these concepts have not yet been ultimately settled shows the enormous complexity of the physics involved.

The generation of radiation, spanning the whole range from the longest radio waves through the infrared, optical, ultraviolet, X- and γ -rays up to cosmic rays, where the physics of particle acceleration enters directly, has always been thought of as being quite well understood. This holds certainly true for most of the optical emission and absorption processes including coherent emissions. It holds also true, though in a more complicated way, for the infrared and sub-millimeter bands, for bremsstrahlung generation of X-rays and for the nuclear processes involved in the γ -ray emission. However, at radio wavelengths the situation has become much more complex. Indeed, gyro- and synchrotron emission processes in magnetic fields, although providing excellent results on mapping large-scale structures, have turned out to be of value only on large spatial and long temporal scales. Depending on the magnetic field strength and the energy of the radiating plasma component, radio emission processes may even 'pollute' the electromagnetic spectrum deep into the optical and even into the γ -ray domains. When it comes to the observation of short-time radiation events, which for the time being can be detected only from objects located close-by in our galaxy or from extremely violent remote emitters, these simple classical theories fail. One is then thrown back to the consideration of plasma emission processes, which are accessible either under grossly non-astronomical conditions in the laboratory or with somewhat better correspondence in near-Earth space. This is where progress in understanding those mechanisms has been achieved and with what we will deal in this review.

Over the last few decades, it has been realized in observational space plasma physics that radio emission from the magnetized planets in the solar system can occasionally become so intense that the conventional theories of emission at radio wavelengths are unable to account for it. The concerted efforts of theorists to explain the extraordinarily high inferred emissivity (or in the language of radio astronomy: T_B , the brightness temperature¹) of these sources initially concentrated on non-linear wave-evolution which, in a perturbational approach, becomes important at large wave amplitudes, and which includes non-linear wave-particle interaction processes. However, these non-linear approaches, as well as those based on non-perturbational techniques invoking the non-linear evolution of the waves, failed in many respects to provide the proper understanding of how natural plasmas could generate the excessively high observed radiative powers. The situation became even more mysterious when high time- and frequency-resolution techniques became available, and showed that the radiation occurred in narrow bands, varied with time and reached peak

¹ For the definition of the brightness temperature in plasma see, e.g., Melrose (1980).

power levels of up to several percent of the available kinetic plasma energy. Indeed, brightness temperatures many orders of magnitude higher than plasma temperature and sometimes reaching values above 10^{20} K were observed. Such brightness temperatures pointed to an unknown coherent emission mechanism² that had to be responsible for the generation of such radiation. The mechanism ultimately identified was the direct linear amplification of the free-space electromagnetic modes by a non-thermal (energetic) electron population present in the plasma. In some of its properties this kind of direct amplification of radiation resembles the well-known molecular masers or lasers where population inversion leads to stimulated emission of coherent radiation of high yield, a phenomenon that has found wide application not only in technology but also in astrophysics. In fact, molecular maser processes have been identified in many astrophysical objects such as the environment of oxygen-rich late-type giants, dust clouds and H-II regions where various rotational lines of OH, H₂O, SiO ... have been attributed to molecular masers. In these masers the pumping is provided by either chemical processes, collisions, or infrared and ultra-violet radiation. (For a recent review the reader may consult the proceedings volume in Migones and Reid 2002.) The resemblance between molecular masers and plasma emissions of high brightness temperature led to the somewhat unfortunate name ‘electron-cyclotron maser’ for the radiation mechanism. This name is misleading, since neither quantum effects, nor energy levels, nor elementary population inversions are involved, and no electrons jump coherently down from an excited energy level onto some lower level.³ In fact, the emission mechanism of the electron-cyclotron maser is simply based on a linear instability of free-space modes in the presence of a non-thermal electron population, which under some special conditions can be excited in a plasma. Such special conditions are, however, not so rare. They occur in many space plasmas, and it is reasonable to assume that they also occur in astrophysical plasmas. The main remaining similarity to a maser is that the plasma must contain a substantially abundant non-thermal electron component as the carrier of an excess of free energy.

Depending on whether this electron component is mono-energetic or has a broad velocity spread corresponding to high temperature, one distinguishes between ‘bunching maser’ and ‘genuine maser’ mechanisms, respectively. For

² ‘Coherent emission’ is defined as emission with brightness temperature $T_B > T$ or $T_B > W_e/k_B$ that is much larger than the temperature T of the plasma or electron kinetic energy $W_e/m_e c^2 = \gamma - 1$, where γ is the relativistic factor. Such high brightness temperatures cannot be explained in terms of thermal emission. Thermal emission, on the other hand, has brightness temperature $T_B < T, W_e/k_B$, where k_B is Boltzmann’s constant.

³ Le Quéau (1988) has attempted a quantum mechanical treatment of the electron-cyclotron maser emission based on the pumping of Landau levels $W_L = \hbar \omega_{ce}(l + \frac{1}{2})$ in magnetized plasma. However, this analogy can be considered only metaphorically since under conditions in space and in most astrophysical plasmas, which are subject to electron-cyclotron maser emission, the prevalent magnetic field strengths are so weak that the Landau levels form a featureless continuum. For example, in the auroral kilometric radiation — the paradigm of an electron-cyclotron maser discussed below in depth — the energy spacing between adjacent Landau levels is $\Delta W = \hbar \omega_{ce} \approx 10^{-29}$ J or $\approx 10^{-10}$ eV which when compared to the auroral electron temperature, where $k_B T_e$ lies between 1 and 10 keV, is entirely negligible.

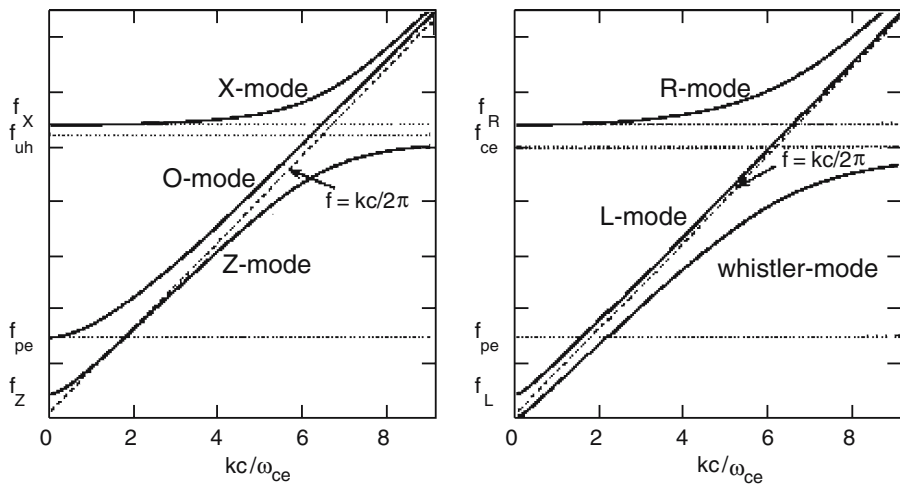


Fig. 1 The dispersion relations of the magneto-ionic wave modes in a plasma for perpendicular (*left*) and parallel (*right*) wave propagation with respect to the ambient magnetic field. f is the frequency given on the ordinate, and $k = 2\pi/\lambda$ is the wave-number vector component given on the abscissa in terms of the product kc of wave number and velocity of light c . The representation with respect to kc (instead of k) has been chosen since then the dispersion relation of light waves becomes the diagonal of the figure. For perpendicular propagation the X-mode is the ‘fast’ branch of the right-hand polarized wave, the Z-mode the ‘slow’ branch of the right-hand polarized wave, and the O-mode the left-hand polarized wave. The X-mode propagates only above the X-mode cutoff f_X , the Z-mode between the Z-mode cutoff f_Z and the upper hybrid frequency f_{uh} , and the O-mode above the plasma frequency f_{pe} . For parallel propagation the X-mode becomes the right-circular polarized R-mode, the O-mode the left-circular polarized L-mode, and the Z-mode the right-circular polarized whistler mode which is confined below the electron–cyclotron frequency f_{ce}

a recent review of the former one may consult Chu (2004). It forms the basis for laboratory attempts of generating intense slow-wave⁴ cyclotron emissions. Application to astrophysical problems is still missing as one does not expect the special conditions of strong bunching to occur under most incoherent natural plasma conditions. Therefore, we will discuss the mechanism of bunching only briefly in context with cyclotron radiation. Nevertheless, the distinction between the bunching maser instability and the genuine maser instability is not a fundamental one: Winglee (1983) has shown that they are just two different limiting cases of the same basic cyclotron instability.

When one calculates the absorption coefficient of a plasma that contains a non-thermal component, negative values are obtained in this volume. Thus, the plasma turns into a macroscopic radiator, which is just the global property of a maser. Melrose and Dulk (1982) have formulated the theory in terms of the absorption coefficient in order to demonstrate the similarity to quantum masers. Loosely spoken, the electron–cyclotron maser is a classical maser for which we will, therefore, retain the now generally adopted name of a maser. In fact, the

⁴ For the definition of ‘fast’ and ‘slow’ waves in plasma, see Fig. 1 and the caption to that figure.

electron–cyclotron maser is a linear plasma instability which directly pumps the free-space electromagnetic modes in the presence of a non-thermal electron population; and this is the view we adopt in the presentation of the theory below rather than the more ambitious absorption-coefficient representation. We note, however, that both representations are equivalent.

Generation of intense radiation is one way, though not the only way, for the plasma to get rid of a substantial part of the excess energy stored in the non-thermal electron component. This happens, when other conventional ways of dissipating the free electron energy by heating, particle acceleration, and generation of anomalous transport are slow. Such conditions are best satisfied when the thermal plasma component is dilute or absent and the non-thermal component dominates, or when the plasma is immersed into a very strong magnetic field. In the latter case, all plasma waves are confined to low frequencies far below the cyclotron frequency. They are not involved in the radiation.

Provisionally, we can provide a rule of thumb: with n_e the plasma density, and $B = |\mathbf{B}|$ the strength of the ambient magnetic field \mathbf{B} , the necessary condition for plasma hosting the electron–cyclotron maser can be written as $f_{pe}^2/f_{ce}^2 \ll 1$. Here $f_{pe} = e\sqrt{n_e/m_e\epsilon_0}/2\pi$, $f_{ce} = eB/2\pi m_e$ are, respectively, the electron plasma and the electron–cyclotron frequencies. Handy formulas for these two frequencies are $f_{pe} \approx 9\sqrt{n_e}$ kHz and $f_{ce} \approx 28B$ kHz, where the density n_e is in cm^{-3} , and the magnetic field B is measured in gauss. The above inequality is, in fact, slightly too strong. The electron–cyclotron maser will also work under the weaker condition $f_{pe}^2/f_{ce}^2 \leq 1$, when it will only be less efficient.

However, only under the most extreme conditions would all the free energy go into radiation. Such conditions are barely realized in magnetized plasmas. Efficiencies of up to several percent are already very close to being extreme. The remaining $\sim 90\%$ free energy may still be dissipated in the plasma in other ways. How this can be achieved will be mentioned when discussing the electron–cyclotron maser mechanism in the main body of this review.

The review is structured as follows: in the next section we briefly discuss the history of the electron–cyclotron maser theory, thereby we include parts of the history of the bunching theory, since the two are closely related. The main in situ observations of the electron–cyclotron maser emission in near-Earth space are reviewed in Sect. 3, and Sect. 4 deals with the elementary theory of the electron–cyclotron maser mechanism in a cold-plasma fluid-approximation including the various hitherto proposed models. Section 5 describes the warm-plasma relativistic-electron cyclotron maser and the long favored loss-cone maser and then turns to the ring-shell maser. We will stress that the latter is the most important one and that radiation generation requires the presence of an electric field component that is aligned with the ambient magnetic field. Hence electron–cyclotron maser radiators are high-probability systems where strong magnetic field-aligned currents have been generated. Section 6 provides arguments for mechanisms which generate such fields, discusses the stability of such fields, and presents the most recent ideas about the generation of maser fine-structures and their informational content about the radiation-source regions. Section 7 will

discuss a number of astrophysical applications, and in Sect. 8 we present some conclusions.

2 History

The possibility of an electron–cyclotron maser emitting radiation at frequencies close to the electron cyclotron frequency and its harmonics was first envisaged about simultaneously but independently by Twiss (1958), Gaponov (1959a), and Schneider (1959). Bekefi et al. (1961) refined the theory. All these authors realized that it should be possible to directly amplify the high-frequency electromagnetic waves in a plasma (cf., Fig. 1) by a resonant interaction between the wave and energetic electrons at the Doppler shifted electron–cyclotron frequency:

$$\omega - k_{\parallel} v_{\parallel} = l \frac{\omega_{ce}}{\gamma}, \quad \gamma = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}. \quad (1)$$

Here $\omega = 2\pi f$ and $\omega_{ce} = 2\pi f_{ce}$ are the angular frequency of the emitted wave and the angular cyclotron frequency, respectively, \mathbf{v} is the resonant electron velocity with components v_{\parallel} , v_{\perp} parallel and perpendicular to the ambient magnetic field \mathbf{B}_0 , respectively, and l is the cyclotron harmonic number. $\mathbf{k} = (k_{\parallel}, \mathbf{k}_{\perp})$ is the wave number of the free-space wave which in magnetized plasmas is related to the frequency via the dispersion relation

$$N^2 \equiv \frac{k^2 c^2}{\omega^2} = \epsilon(\omega, \mathbf{k}). \quad (2)$$

The quantity N is the index of refraction, and $\epsilon = \mathbf{I} + \boldsymbol{\sigma}/i\omega\epsilon_0$ is the dielectric tensor (its projection $\mathbf{k} \cdot \epsilon \cdot \mathbf{k}$ onto the direction of \mathbf{k} is the response function) of the plasma, and \mathbf{I} is the unit tensor. Explicit expressions for the linear plasma conductivity tensor $\boldsymbol{\sigma}$ have been given in (Montgomery and Tidman 1964; Bekefi 1966; Melrose 1980; Baumjohann and Treumann 1996) and by others. A general form is

$$\frac{\boldsymbol{\sigma}}{2\pi i\omega\epsilon_0} = \frac{\omega_{pe}^2}{\omega^2} \int_0^{\infty} p_{\perp} dp_{\perp} \int_{-\infty}^{\infty} dp_{\parallel} \sum_{l=-\infty}^{\infty} \frac{\mathbf{S}/\gamma}{\omega - k_{\parallel} p_{\parallel}/m_e \gamma - l\omega_{ce}/\gamma}, \quad (3)$$

where the tensor \mathbf{S} is represented by the following matrix

$$\mathbf{S} = \begin{pmatrix} v_{\perp} U \left(\frac{U_l}{x}\right)^2 & i v_{\perp} U \frac{U_l J'_l}{x} & v_{\perp} W \frac{U_l^2}{x} \\ -i v_{\perp} U \frac{U_l J'_l}{x} & v_{\perp} U (J'_l)^2 & -i v_{\perp} W J_l J'_l \\ v_{\parallel} U \frac{U_l^2}{x} & i v_{\parallel} U J_l J'_l & v_{\parallel} W J_l^2 \end{pmatrix}. \quad (4)$$

The parameter x is defined as $x = \gamma k_{\perp} v_{\perp} / \omega_{ce}$, $J_l(x)$ are Bessel functions of order l , and the operators U, W are given through

$$U = m_e \gamma \omega \partial_{p_{\perp}} f_0 + k_{\parallel} (p_{\perp} \partial_{p_{\parallel}} - p_{\parallel} \partial_{p_{\perp}}) f_0, \quad (5)$$

$$W = m_e \gamma \omega \partial_{p_{\parallel}} f_0 - m_e l \frac{\omega_{ce}}{p_{\perp}} (p_{\perp} \partial_{p_{\parallel}} - p_{\parallel} \partial_{p_{\perp}}) f_0. \quad (6)$$

In all these expressions $\mathbf{p} = m_e \gamma \mathbf{v}$ is the relativistic momentum of the electrons, and \mathbf{v} the velocity. \mathbf{p} has the components p_{\parallel} and p_{\perp} parallel and perpendicular to the magnetic field. In the magneto-ionic (cold plasma) limit the propagating free-space solutions of this dispersion relation (cf. Fig. 1) are the (fast) RX and LO free-space modes propagating above the respective cutoff frequencies ω_X (8) and ω_{pe} , and the (slow) RZ (or simply Z) mode which cannot leave the plasma, since it is confined below the upper-hybrid frequency $\omega_{uh} \equiv (\omega_{pe}^2 + \omega_{ce}^2)^{1/2} < \omega_X$ (Bekefi 1966; Budden 1988; Melrose 1980; Baumjohann and Treumann 1996). The RX and Z modes are polarized right-hand, while the LO mode has left-hand polarization. Figure 1 gives a graphic representation of the dispersion relations of these modes for the two main directions of propagation: perpendicular and parallel to the ambient magnetic field.

For the electron-cyclotron maser including the relativistic factor γ — even in very weakly relativistic plasmas — is important. This was, however, realized only later.

Hirshfield and Bekefi (1963) pointed out that coherent mechanisms could be at work in planetary magnetospheres such as that of Jupiter when they endeavored to explain the observation of Jovian decametric radiation. Hirshfield and Wachtel (1964) actually demonstrated the emission in a laboratory experiment with relativistic electrons. Non-relativistic versions of electron-cyclotron maser emission had been developed in other papers by Gaponov (1959b), Harris (1959) (who suggested that certain non-thermal electron-velocity distributions favor the emission) and Sagdeev and Shafranov (1960). These theories were further developed by Melrose (1973, 1976), who assumed a bi-Maxwellian electron distribution function and found from quasilinear theory that coherent emission close to the fundamental harmonic $l = 1$ in the X-mode would dominate but can occur only in highly underdense plasma with $\omega_{pe}^2 \ll \omega_{ce}^2$, for temperatures

$$(k_B T_{\perp})^2 |\cos \theta| > \frac{1}{2} m_e c^2 k_B T_{\parallel}, \quad (7)$$

where θ is the angle of propagation with respect to the ambient magnetic field. The second condition (7) is very severe as, in particular for very oblique emission $\theta \approx \pi/2$ close to the perpendicular, it requires very high temperature anisotropies T_{\perp}/T_{\parallel} of the order of $> 10^2$ which are hard to achieve. Emission at harmonics faces weaker restrictions on the plasma density. Melrose (1976) also pointed out that all emissions are subject to the constraint that the radiation should be able to escape from the plasma. Since in magnetoionic theory (Budden 1988) there are only the above-mentioned two free-space modes LO

and RX which can escape, escaping radiation must be above the respective lower cutoff frequencies $\omega > \omega_{pe}$ for the LO-mode, or for the RX-mode

$$\omega > \omega_X = \frac{1}{2} \left[\omega_{ce} + \left(\omega_{ce}^2 + 4\omega_{pe}^2 \right)^{\frac{1}{2}} \right]. \quad (8)$$

The radiation has to choose between these two conditions. It turns out that it favors the RX-mode even though its cutoff is at much higher frequency than that of the LO-mode. Hence escaping radiation should exhibit right-hand polarization in the RX-mode. The (slow) Z-mode — or low-frequency RX-mode — is confined to the plasma and is thus not accessible to remote sensing. These escape restrictions remain valid even under relativistic conditions. All these restrictions temporarily drove the interest away from further consideration of the electron–cyclotron maser in space and astrophysical applications.

Wu and Lee (1979) achieved major progress. They realized that (similar to the importance of the relativistic correction in the bunching maser) including relativistic effects in the electron component — even in the very weakly relativistic case — has a decisive effect on the efficiency of the amplification of the emitted wave. The reason is that the resonance condition (1) taken in the non-relativistic limit $\gamma = 1$ is independent of v_{\perp} and yields a straight line $v_{\parallel} = (\omega - \omega_{ce})/k_{\parallel}$ in velocity space.⁵ Along this line only very few resonant electrons contribute to radiation.

The exact solution of (1) yields an ellipse in the plane $(v_{\perp}, v_{\parallel})$ which can be positioned on the particle distribution in such a way that it visits many of the resonant particles which carry the available free energy. Wu and Lee (1979) found that including the weakly relativistic correction for electrons in the Earth's auroral region, which cover the energy range from 0.3 to 10 keV, substantially enhanced the growth rate of the RX-mode over the rates found earlier in the non-relativistic approach. They proposed that natural emissions of the kind of the auroral radio emissions could be generated by the semi-relativistic electron–cyclotron maser instability. Such terrestrial radio emissions had been detected and described by Benediktov et al. (1965) in 1965, and had been investigated more closely and confirmed by Gurnett (1974). Gurnett (1974) was able to infer the bandwidth of the terrestrial radio emission and its average intensity which he found to be between 1 and 10% of the total energy supplied to the Earth's magnetosphere during a substorm. By 1985 investigations of the direction from which the radiation was emitted (Mellott et al. 1984, 1985) had demonstrated clearly that the emission resulted from the auroral zone just above the auroral ionosphere at altitudes of a few 1,000 km. Later investigations and refinements found that the emission came from the nightside of the magnetosphere and this related it to processes in the Earth's magnetotail, which are now believed to come from the violent reconnection process taking place in the Earth's plasma sheet during substorms.

⁵ See the detailed discussion of the resonance condition in Sect. 5.2.

The paper by Wu and Lee (1979) was immediately followed by a vast activity in the calculation of various models of the electron–cyclotron maser mechanism for different electron-distribution functions and for different objects. The Sun largest number of models concerned the Sun [for an intermediate review of solar observation and theory see McLean and Labrum (1985), Dulk (1985), Bastian et al. (1998)], followed by models for the Earth’s auroral radio emissions and Jupiter’s radio radiation. The idea of Wu and Lee (1979) had been that the electron-distribution function, which is responsible for the electron–cyclotron maser instability, and the excitation of the terrestrial radio emission, which was now called ‘auroral kilometric radiation’ (AKR), is a so-called loss-cone distribution on the precipitating and trapped hot auroral electron component. The loss-cone is dug out from the initially isotropic distribution in pitch angle θ^6 by the collisions of the nearly parallel electrons with the dense and cold upper ionospheric plasma component.

A loss-cone distribution indeed satisfies the condition that electrons virtually reside in a higher energy state than the (isotropic and pitch-angle independent) ‘ground state’ of the equilibrium distribution function. This means that the distribution function in a certain region around the loss cone is anisotropic and lacks the parallel particle component required to isotropize it. This formally causes a positive derivative on the electron-distribution function $f(v_{\parallel}, v_{\perp})$ with respect to the perpendicular velocity of the electrons, viz. $\partial f / \partial v_{\perp} > 0$. Now, since these electrons which are residing along the edge of the loss cone carry the free energy by contributing to the positive derivative, it would be desirable to draw the resonance line (1) along the positive gradient of the distribution function. Indeed, this becomes possible by exploiting the aforementioned relativistic deformation of the resonance line from a straight line into an ellipse, the position of which in velocity space $(v_{\parallel}, v_{\perp})$ can be chosen in such a way that it covers most of $\partial f / \partial v_{\perp} > 0$.

The loss-cone maser theory⁷ was further expanded by Melrose and Dulk (1982), Melrose et al. (1982), Melrose et al. (1984), and Hewitt et al. (1982) and was applied also to solar-system radio emissions, including planetary radiation, solar microwave spikes in impulsive flares, radio emission from binary systems, and narrow-band emissions from dwarf M flare stars. By including the relativistic correction into the electron–cyclotron maser mechanism, the very disturbing earlier requirement of extreme temperature anisotropies Melrose (1973) in the non-relativistic cyclotron maser had now been removed. Realistic small temperature anisotropies became sufficient for letting the maser work. However, the growth rates still remained comparably small and, in addition, the emission turned out to have large bandwidth. This was less disturbing until observation of the fine structure of the AKR revealed that the auroral radio radiation consists of very many narrow emission features (Gurnett and Anderson 1981)

⁶ The pitch angle $\theta = \arccos(\mathbf{v} \cdot \mathbf{B} / |\mathbf{v}| |\mathbf{B}|)$ is the angle between the velocity vector \mathbf{v} of a particle and the magnetic field \mathbf{B} .

⁷ An extended summary of the linear theories of the loss-cone maser can be found in Wu (1985). A short introduction into the maser theory is given in Treumann and Baumjohann (1997).

which move in an irregular fashion on top of a comparably weak background radiation across the dynamical frequency-time spectra. Moreover, these narrow emission features turned out to be very intense. Interpretation of such properties was very difficult when sticking to the loss-cone maser. Models required extremely steep perpendicular velocity gradients on the distribution function at the edge of the loss cone which neither were supported by observation nor could be justified theoretically. The information needed for developing a realistic model mechanism was clearly stored in conditions that could only be explored in situ, i.e., within the radiation source region.

In the early 1980s, then, it turned out that the loss-cone maser theory was not supported by observation of the AKR. Moreover, two-dimensional particle simulations based on measured electron loss-cone distributions (Pritchett 1984a,b; Pritchett and Strangeway 1985) had failed as well. Omidi and Gurnett (1982a,b, 1984), guided by the observations of the polar-orbiting spacecraft S3-3 high above the aurora at 8,000 km altitude, undertook to determine the maser growth rate for the Earth's AKR from experimental data. Investigating the weakly relativistic cyclotron resonance condition as it had been given by Wu and Lee (1979) and had been elaborated in depth in particular by Melrose and Dulk (1982) and Melrose et al. (1982), they tried to determine the resonance ellipse (see the discussion below) in the auroral electron distribution. They concluded that for the weakly relativistic auroral energies between 0.1 and 10 keV the resonance ellipse could be approximated by a resonance circle that was displaced from origin and that, however, this circle covered at least a few electrons in resonance lying mostly inside the loss-cone. So they concluded, sticking to the then general philosophy, that the loss-cone is the main source of AKR emission by the maser mechanism. Under this assumption they numerically determined the growth rate of the maser by integrating the electron distribution along the resonance ellipse and found very small and non-acceptable growth rates, which they tried to justify with the quasi-linear saturation of the emission and the quasi-linear depletion of the loss-cone by energy loss to the radiation.⁸ In fact, the resolution of the electron measurements was not good enough to let them realize that the distribution was closer to a ring or shell distribution with only a small and negligible loss-cone component present. Nevertheless, this attempt

⁸ Such a depletion of the loss-cone was considered by many authors in application to various solar radio bursts and their fine structures (cf., for instance, Aschwanden 1990a,b; Aschwanden and Benz 1988a,b, Melrose 2002 and others) as the dominant saturation mechanism. The idea is that the loss-cone would be very vulnerable to energy losses of the electrons at the boundary of the loss-cone, and that the small energy loss to the radiation would be sufficient to fill in the loss-cone. However, this assumption requires that the radiation does not escape from the source region, an assumption barely satisfied in the source region of the AKR. Moreover, here the loss-cone is very wide, and it is hard to believe that weak pitch-angle diffusion, due to the small energy loss to the radiation, could bridge its entire width. In contrast, it is easy to understand that at equatorial altitudes the loss-cone can be filled due to pitch-angle diffusion in the resonant interaction of electrons with whistler waves, as Kennel and Petschek (1966) argued in their classic paper on the stability of the radiation belts, because here the loss-cone is extremely narrow and the growth rate of the resonant whistlers is very large. However, in the equatorial region the electron-cyclotron maser instability cannot be excited. This argument also applies to other systems like Jupiter, Saturn, and solar coronal loops.

demonstrated that the theory could be tested by exploiting the in situ measurements. Clearly, without such measurements the generation mechanism of the auroral kilometric radiation would still lay in darkness.

Failure of the loss-cone mechanism thus became obvious already early on, even though the overwhelming majority of investigators was still favoring it. It was instead suggested in 1986 that 'ring' or 'shell' distributions produced by field-aligned electric fields may drive the electron-cyclotron maser Winglee and Dulk (1986). A study based on the Swedish Viking satellite⁹ observations suggested that the loss-cone itself is insufficient as a reason for emission of auroral kilometric radiation. Louarn et al. (1990) found in these observations many cases with pronounced loss-cones with no auroral kilometric radiation. Instead, they found that the free energy for the growth of auroral kilometric radiation comes from electron-distribution functions exhibiting a positive gradient which was not necessarily located inside the loss cone. This led to speculation that the emissions, at least as observed in situ in the Earth's environment, do not draw their energy from a loss-cone instability. Rather the energy comes from a deformed distribution which is the result of the action of magnetic-field-aligned electric fields¹⁰ as it had already been proposed by Pritchett and Strangeway (1985) based on the electric-field model investigated by Chiu and Schulz (1978).

The missing fact in the theory of the electron-cyclotron maser of the auroral kilometric radiation was indeed the neglect of the role that is played by a field-aligned electric field in the auroral dynamics. The presence of magnetic-field-aligned electric potential drops to accelerate the auroral electrons in the auroral magnetosphere was a long suspected though mysterious fact. It had already been anticipated by Hannes Alfvén (1958) in the fifties and had since been promoted by his collaborators and others, but did not receive the attention it deserved. In the practically collisionless plasma of the auroral region, parallel potential drops were believed not to occur over the length of time that is sufficient for generating auroral electron beams. The first indication on the real importance of field-aligned potentials in the auroral region had been provided early on in a widely neglected paper published by Calvert (1981a). Calvert (1981a) had observed that the auroral region is practically void of thermal

⁹ Viking (Hultqvist 1990) was a Swedish auroral spacecraft that was launched on February 22, 1986 from Kiruna, Sweden, into a polar orbit at roughly 8,000 km distance from Earth, measuring there until December 1986. It was particularly tailored for the exploration of the properties of substorms and the auroral radio emission.

¹⁰ The first very simple model of such electric fields by Knight (1973) was based on thermal equilibrium and had nothing to do with the real physical process that keeps a field-aligned electric field alive. This model was followed by more recent attempts to construct stationary electric fields along the magnetic field provided by Lyons (1980), Vedin and Rönmark (2005) and others. Measurements of such auroral electric fields have been reviewed Weimer et al. (1985) and more recently by Hull et al. (2003). Earlier, Föppl et al. (1968) and Haerendel et al. (1976) injected a Barium ion cloud into the upper ionosphere and, from observation of emission lines of Barium atoms and Barium ions in the visible, inferred a sudden upward acceleration of the ions along the Earth's auroral magnetic field-lines at high altitude of a few 1,000 km. They concluded that the acceleration had been provided locally by an upward electric field along the Earth's magnetic field.

plasma, and proposed that radiation in this region, where the plasma frequency is much lower than the electron–cyclotron frequency, should be generated by the energetic plasma component and should be bouncing back and forth between the walls of the ‘auroral density trough’ that acts as a wave guide. Thereby the radiation would be amplified before it leaked out from the trough to free space. Calvert (1981a) did not address the question of how the trough would be produced. However, Chiu and Schulz (1978) had already shown 3 years earlier that such a trough would occur, provided a magnetic-field-aligned electric potential drop existed in the auroral region. Mechanisms that generate potential drops and field-aligned electrostatic fields are under consideration even today because finding conditions for maintaining them in a collisionless plasma presents serious difficulties. But Chiu and Schulz (1978) had shown that such drops once they are established deform the original trapped electron-distribution function in a certain characteristic way which much later turned out to be crucial for the maser mechanism. Under the combined action of these electric fields and the mirror effect of the converging magnetic field-lines, the electron-distribution function assumes the shape of a ring in $(v_{\parallel}, v_{\perp})$ -phase space. This was ultimately demonstrated in particle simulations (Winglee and Pritchett 1986).

In order to understand how this situation arises, one may assume that a stationary electric field exists along a magnetic field-line pointing upward. Such a field accelerates electrons down into the converging magnetic field. While the electrons move down into the stronger magnetic field, the first adiabatic invariant $\mu = m_e \gamma c^2 \sin^2 \theta / B$ (the magnetic moment of the particles) is conserved, and as a consequence the electrons assume larger pitch angles θ . This causes a peak in the electron distribution at the larger pitch angles and, by removing electrons from smaller energies due to acceleration in the electric field, produces a hole at small energies, and thus causes a positive perpendicular-velocity gradient, $\partial f_0 / \partial v_{\perp} > 0$, in the distribution function. Thus, in the presence of a field-aligned electric field component, there should exist an inner region in the distribution function at low energies that is void of particles (Chiu and Schulz 1978).

This latter prediction is not in agreement with observations which showed, instead, that the electron distribution exhibits a well-pronounced plateau at smaller velocities. Explaining this plateau involved considerable difficulties. It is clear by now that the plateau is not caused by a quasi-linear relaxation of the distribution through energy loss by electron–cyclotron maser emission. Instead, it was found to be caused by wave-particle interaction between the energetic electrons and low-frequency waves in the Very low-frequency range (LaBelle and Treumann 2002). These low-frequency waves are bound to the plasma and have no way to escape from it. Their excitation is unavoidable; it is intrinsic to the positive perpendicular velocity gradient in the undisturbed distribution function. Their intensity grows high enough to imprint on the electrons which are in resonance with the VLF waves and to deplete the gradient and generate the plateau. In fact, artificially inhibiting the depletion of the gradient by taking out the VLF waves retains the gradient and leads to larger maser emissivity. Escaping free-space modes of the maser radiation are not intense

enough to deplete the gradient, unless the radiation would be kept inside the maser source by some mechanism of confinement. Thus electric fields, particle acceleration, and VLF waves work together in producing auroral kilometric radiation as a most intense emission which can be understood only on the basis of the electron–cyclotron maser mechanism in the presence of a field-aligned electric potential drop. This has been a most important realization of the last decade, and its consequence for astrophysical plasmas has yet to be exploited.

Stimulated by all these investigations, Pritchett (1984a,b), developed the electron–cyclotron maser theory for ring distributions of the kind described above. The important point he realized was that the resonance condition (1) becomes a perfect circle in phase space for strictly perpendicular free-space mode propagation. Positioning the resonance line for perpendicular emission is thus ideally adapted to the ring distribution. It follows from the propagation properties of the RX modes that in this case only the extraordinary wave will become resonantly amplified.¹¹

Interestingly, due to the relativistic mass increase the emitted resonant frequency for perpendicular propagation falls below the (non-relativistic) electron–cyclotron frequency, $\omega < \omega_{ce}$. An estimate of the dependence of the linear growth rate of the radiation brought the result that maximum growth is attained just for strictly perpendicular emission of the fundamental harmonic at such frequencies below ω_{ce} . Of course, higher harmonic emission is possible as well but its growth is much slower and thus radiation in these harmonic modes should be weaker. On the other hand, escaping from the density trough is no problem for them. Under certain conditions, when the fundamental, most intensely excited radiation cannot leave the plasma, one may thus expect to observe harmonic radiation instead.

Calvert (1981a) claim that multiple bouncing of the fundamental radiation inside the density cavity will amplify the radiation before escape will, however, in most cases be barely realized simply because the conditions for such an amplification require that the length of the ray path between the boundaries is a multiple of the radiation wavelength. This does not preclude that under fortunate conditions and for broadband amplification of the RX-mode, the macroscopic density trough may filter out a particular wavelength, amplify and release it to free space. Radiation of this kind will clearly be narrow band. Such a model has been proposed by Louarn and Le Quéau (1996a,b) who considered the auroral density trough as an imperfect wave guide and showed that the resonantly amplified RX-mode is not only amplified but also couples to the external LO-mode outside the walls. The latter leaves the plasma without any problem. A model of this kind could explain the sometimes observed imperfect polarization of the radiation.

Pritchett (1984a,b) theory turned out to provide the relevant tool for interpreting the Earth's auroral kilometric radiation when in the late eighties and early nineties in situ measurements of plasma and radiation above the

¹¹ This is explained in detail in Sect. 5.2.

Earth's auroral zone from the Viking and FAST¹² satellites became available. The measurements demonstrated that the earlier suspected field-aligned potential drops do exist and are accompanied by ringlike particle distributions with only moderate loss-cones. In situ crossings of the generation region of the auroral kilometric radiation by the FAST satellite showed, in addition, that the most intense emissions occurred below the local electron–cyclotron frequency and in the regions of highly diluted plasma density that contained accelerated hot electrons. An enormous fine structuring of the radiation found in the source region suggested that the fine structure observed earlier by Gurnett and Anderson (1981) had not been created at large distance from the source by propagation nor through selection by wavelength but was a property of the source itself. Closer investigation of the plasma properties of the source yielded the result that the source region is filled with small-scale parallel electric field structures, either microscopic solitary waves or so-called electron holes in phase space which actively contribute to the generation of radiation. They possibly even cause the fine structure of the radiation by a mechanism which until the time of writing has not been clarified. The observations indicated that almost all of the radiation consists of narrow-band ‘elementary’ radiators of microscopic size. These act like point-like sources of the radiation moving both in real space and in phase space.

As mentioned before, astrophysical applications of the electron–cyclotron maser mechanism outside the solar system are still sparse. Application to solar radio emission dominated the field from the very beginning just until the discovery of the intense terrestrial auroral radio emission in 1974 (Gurnett 1974). Attempts to explain the solar radiation in the radio band from the longest wavelengths down to microwaves abound in the literature. Solar type IV and type V bursts, solar radio spikes, and various fine structures in the solar radio emission have been proposed to be the result of maser activity. Application to Jupiter and radiation from the other magnetized planets followed, and speculations about radiation emitted from extrasolar planets when they are magnetized strongly enough have been published as well (Bastian et al. 2000; Zarka et al. 2001a,b). Radiation of this kind, because it would be much stronger than any other radio emission, would allow not only to detect ‘radio-loud’ extrasolar planets but also to infer their magnetic field strengths and plasma properties. Its identification is therefore of vital interest. However, given the experience with the Sun which is a strong radio emitter it may be more realistic for the time being to apply the theory to other stellar systems. The proposals of this kind have been around since the earliest application of the electron–cyclotron maser mechanism (Louarn et al. 1986) and have been renewed meanwhile in view of the modern theories (Ergun et al. 2000; Bingham and Cairns 2000; Bingham et al. 2001). Applications of the theory to exotic objects like pulsars (Ma et al.

¹² FAST is the mnemonic for NASA's Fast Auroral Snapshot (FAST) Explorer spacecraft, which has been launched into a 350 km by 4,180 km altitude polar orbit on August 21, 1996, the first of NASA's small spacecraft. FAST was tailored to improve substantially and complement Viking observations with very high-resolution in situ measurements at lower altitudes in the auroral plasma cavity (Pfaff et al. 2001).

1998) and Blazar jets (Begelman et al. 2005) have recently been attempted as well. These more speculative theories have of course to await more detailed observations afforded by the much higher spatial and temporal resolutions of future instrumentation. In this respect, space plasma physics has a tremendous advantage over astrophysics as near-Earth or planetary space offers the possibility to undertake in situ measurements of the very conditions of emission. In the following section, we review the most relevant observations that support the very existence of the natural electron–cyclotron maser in space.

3 Observing the space plasma in situ

Since the radiation generated in the electron–cyclotron maser mechanism is closely related to the plasma configuration in the radiation source, a review of the relevant observations supporting its existence cannot avoid to go into some detailed discussion of the connected plasma observations. This is possible only for the region very close to the Earth (and sporadically as well but with much higher uncertainty for other planets like Jupiter). We therefore start in this section, after briefly reviewing the first observational papers, with the description of the observational properties of the Earth’s auroral kilometric radiation the source of which is the upper auroral ionosphere between, roughly, 1,000 and 8,000 km altitude, a region which has turned out to be so far the only accessible and in any case the closest to Earth electron–cyclotron maser source. Other much less accessible and still quite uncertain sources are the auroral regions of the large planets Jupiter, Saturn, Uranus, and Neptune and, of course, the solar corona of which it has been most frequently speculated that electron-maser emission is one of the dominant mechanisms generating radio emission in the range from meter to micro waves. This concentration on Earth’s environment cannot be avoided in this discussion as only there ideas, theory, and simulations can be compared with the hard facts of observation.

3.1 Large-scale properties of auroral kilometric radiation

Figure 2 shows the first crude measurement (Gurnett 1974) of the full spectrum of the auroral kilometric radiation recorded by the Interplanetary Monitoring Platform Imp 6 spacecraft flying far above the ionosphere and seeing the radiation coming from the entire Earth¹³. Terrestrial radio emission had already been detected earlier. The earliest observation by Benediktov et al. (1965) showed natural radio radiation leaving the Earth’s upper atmosphere but covered only part of the spectrum at high frequencies. Nevertheless, these authors already identified the observed radiation correctly as ‘terrestrial radio noise’. Other

¹³ The Interplanetary Monitoring Platform spacecraft Imp formed consecutively numbered series of early scientific spacecraft. Other similar series were the Explorer spacecraft, the Orbiting Solar Observatory OSO spacecraft, the Orbiting Astronomical OAO spacecraft and others.

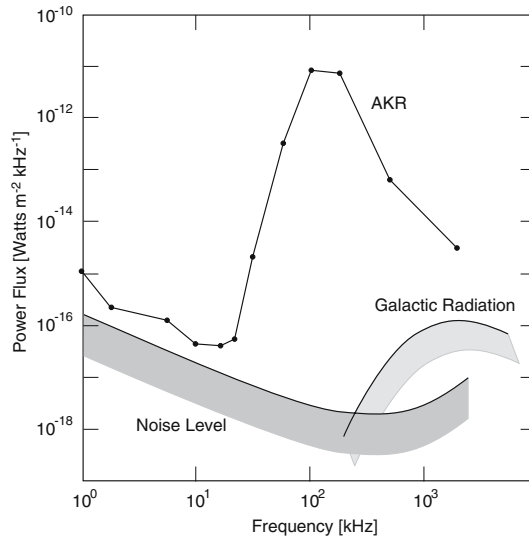


Fig. 2 The first measurement of the auroral kilometric radiation spectrum obtained from the Imp 6 spacecraft. Included into the figure are the instrumental noise level for comparison and the spectrum of the galactic cosmic background. The latter becomes important at frequencies above a few hundred kilohertz. The striking fact of this first still relatively crude observation of the auroral kilometric radiation spectrum is that the spectrum is orders of magnitude stronger than the galactic background radiation. It is a rather broad emission spectrum, covering the frequency range above background (noise level and real near-Earth background) from ~ 1 kHz to a few MHz with a steep maximum around a few hundred kilohertz (after Gurnett (1974), with permission by the American Geophysical Union)

parts of the auroral kilometric spectrum had been observed on the Orbiting Geophysical Observatory OGO 1 (Dunkel et al. 1970) and on Imp 6 (Brown 1973), and were followed by the Hawkeye 1 spacecraft (Kurth et al. 1975) (who coined the name ‘auroral kilometric radiation’), Imp 8 (Gurnett 1975), and the Radiation Astronomy Explorer Rae 2 (Alexander and Kaiser 1976), the International Satellite for Ionospheric Studies ISIS 1 (Benson and Calvert 1979; Benson 1982, 1985), the Swedish spacecraft Viking (Bahnsen et al. 1987), and most recently (Carlson et al. 1998; Ergun et al. 1998a,b,c) by FAST and the Cluster¹⁴ mission (Mutel et al. 2006).

The first rough measurement of the full auroral kilometric spectrum in Fig. 2 exhibits a pronounced maximum at a few hundred kHz, exceeding the near-Earth radio level between about 20 kHz and a few GHz by orders of magnitude. Comparison with the galactic radio noise spectrum shows that the auroral kilometric emission dominates even at the frequencies of the galactic noise. This is clearly only the case if one is close enough to Earth. In addition, the

¹⁴ Cluster is a European Space Agency (ESA) mission comprising four satellites – which were launched in pairs on Soyuz launchers from Baikonur on 16 July and 9 August 2000 – to fly in formation with separation distances ranging from 200 to 10,000 km (Escoubet et al. 1997). The Cluster orbit is 4 by 20 Earth radii with an inclination of 90° . Cluster is currently scheduled to operate until 2009.

galactic noise can be distinguished from the auroral kilometric radiation by its stationarity while the auroral kilometric radiation exhibits fast variations both in time and frequency since it strongly depends on auroral activity.

The typical total power radiated in the auroral kilometric radiation is of the order of 10^7 W during small substorms, while peak power levels can be as high as 10^9 W during strong substorms, caused by the intense disturbance of the Earth's magnetic field resulting from its interaction with the super-alfvénic solar wind. The ultimate energy source of a substorm is the solar wind flow with velocity $v_{sw} > v_{A,sw}$ and the solar wind magnetic field \mathbf{B}_{sw} , the latter acting through the solar wind convection electric field $\mathbf{E} = -\mathbf{v}_{sw} \times \mathbf{B}$, and hence the solar wind is also the ultimate energy source of the auroral kilometric radiation. The mechanism of energy transfer from the solar wind to the geomagnetic field consists of a combination of the mechanical compression of the geomagnetic field by the solar wind ram pressure $p_{sw} = n_{sw}m_p v_{sw,n}^2$, with n_{sw} the solar wind density, $v_{sw,n}$ the solar wind velocity component normal to the boundary between the geomagnetic field and the solar wind (i.e., the magnetopause), and m_p the proton mass — an approximation based on the composition of the solar wind which is $\sim 95\%$ protons and $\sim 5\%$ helium — a possible but so far unidentified friction force acting at and along the magnetopause and, as the main mechanism, reconnection between magnetic fields carried by the solar wind and the geomagnetic field. Substorms are the elementary disturbances through which the large geomagnetic storms are built. The former occur in the wake of the interaction between solar coronal mass ejections and the geomagnetic field, and last over days and play an important role in the effects which are of interest in space weather prediction. During a typical substorm of duration of $\sim 10^3$ s a total energy of between 10^{11} and 10^{14} J corresponds to a power of between 10^8 and 10^{11} W that is transferred to the magnetosphere. Accordingly, the energy emitted in auroral kilometric radiation is typically several percent of the total substorm energy. The resulting brightness temperatures lie between 10^{10} K for the weakest events and 10^{24} K or more for the strongest events, i.e., many orders of magnitude higher than what may be expected from incoherent gyro or synchrotron mechanisms of radiation (Jackson 1962; Rybicky and Lightman 1979). Auroral kilometric radiation — and in the same spirit the highly time-variable radio emissions of the outer large magnetized planets of the solar system — is thus seen to be far from being generated by an incoherent emission process. The measured intensities require a strongly coherent process.

The polarization measurement of the AKR based on Hawkeye spacecraft observations (Gurnett and Green 1978) indicated that the radiation is emitted mostly in the free space RX mode and that OL mode emission or slow Z mode emission may be present but are in general much weaker.

The direction finding measurements of AKR (Jones 1976; Green et al. 1977; de Feraudy and Schreiber 1995) pointed to sources that are located above auroral heights at altitudes between $\sim 2,000$ and $4,000$ km with the intensity of radiation decreasing as the inverse square of distance from Earth. When looked at from far away, the source of the auroral kilometric radiation and thus the location of the electron-cyclotron maser is therefore about point-

like. This region is known as the ‘auroral cavity’ (Calvert 1981b; Hilgers 1992) because of its very low plasma density. This has also been confirmed by the above-mentioned polarization measurements (Gurnett and Green 1978) which found the auroral kilometric radiation being strictly correlated with ratios $\omega_{ce}/\omega_{pe,r} \gg 1$.

The existence of the auroral cavity poses a problem in itself as it is by no means obvious that the plasma density on auroral field-lines must be as low as it is observationally found. For a long time the relevance of this fact for the emission process and the dynamics of the auroral plasma has been grossly underestimated. On the other hand, the auroral cavity is not a permanent, invariable feature. Rather the auroral plasma consists of many wider or narrower cavities with varying latitudinal extension and generally much larger extensions in longitude. Each of them can be looked at as a two-dimensional depletion of the auroral plasma content which, in most cases, is confined latitudinally by dense plasma walls.

3.2 The auroral plasma cavity: plasma and its field properties

Figure 3 shows a representative collection of the plasma and field measurements during a crossing of the auroral plasma cavity by FAST at $\sim 4,000$ km altitude. One of the main properties of this region is that it contains comparably strong field-aligned electric currents. These currents couple the conducting part of the upper Earth’s atmosphere, the ionosphere, to the electrical currents flowing in the outer magnetosphere and being generated by the solar–wind interaction as shown schematically in Fig. 4. The magnetic field B_z of these field-aligned currents is transverse to the strong ambient magnetic field and is shown in panel a. Positive (negative) gradients in B_z indicate downward (upward) currents corresponding to upward (downward) ionospheric electron fluxes.

These relations are shown in panels e and f where the electron energy flux and electron pitch angles (i.e., the angle of individual particle velocity with respect to the ambient magnetic field) are given. The electron flow direction can be taken from panel f, where the broad region of upward current and downward electron fluxes concentrates at angles of 0° and 360° but, in addition, exhibits a broader distribution in pitch angle with practically no upward electrons at 180° . Moreover, the energy flux (panel e) shows the strong downward fluxes concentrated at an energy near 10 keV and absence of low-energy electrons.

Panels g and h show the corresponding ion energy fluxes and pitch angles, respectively. In the upward current region, all ions flow upward at 180° pitch angle covering a narrow energy band with energy varying between a few keV and 10 keV. Obviously these ions have been accelerated in the upward direction. Such behavior in ions and electrons is typical for the presence of an electric field pointing upward along the ambient magnetic field, accelerating the ions upward and electrons downward. In fact, electrons with an energy of ~ 10 keV moving downward along the magnetic field carry this current. This electric field is responsible for the very low cold plasma density in the auroral cavity; it creates the auroral cavity. The temporal width of the auroral cavity in the present case is

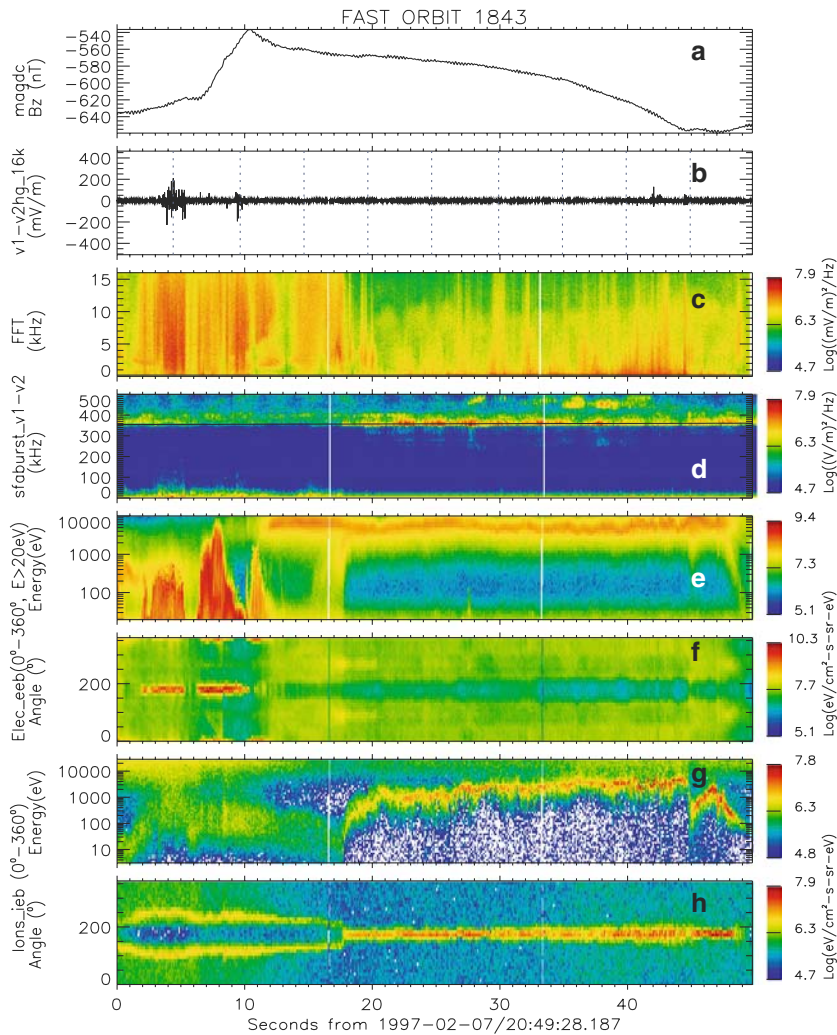


Fig. 3 Overview of plasma characteristics observed by instruments on board of the FAST spacecraft at altitudes in and above the source region of the auroral kilometric-radiation electron-cyclotron maser source. Crossing of the cavity boundary occurs at 18 s. **a** Magnetic field of field-aligned current; **b** electric wave form <16 kHz; **c** Fast Fourier transform (FFT) of wave form showing the electric wave spectrum for frequencies < 16 kHz; **d**: wave spectrum <500 kHz (black line near 360 kHz in panel **d** is the local f_{ce}); **e** parallel electron energy flux; **f** electron pitch angle; **g** parallel ion energy flux; **h** ion pitch angle. For further description see text (after Pottelette and Treumann (2006))

~30 s which at spacecraft velocity of 5 km s^{-1} corresponds to a latitudinal extension of just 150 km. Much narrower cavities have been detected as well (Fig. 5).

In the narrow downward current region upward electron fluxes concentrate at pitch angles 180° while their energies cover the whole range from 0 keV up to ~10 keV. The ions are found here at two intermediate pitch-angles sparing out

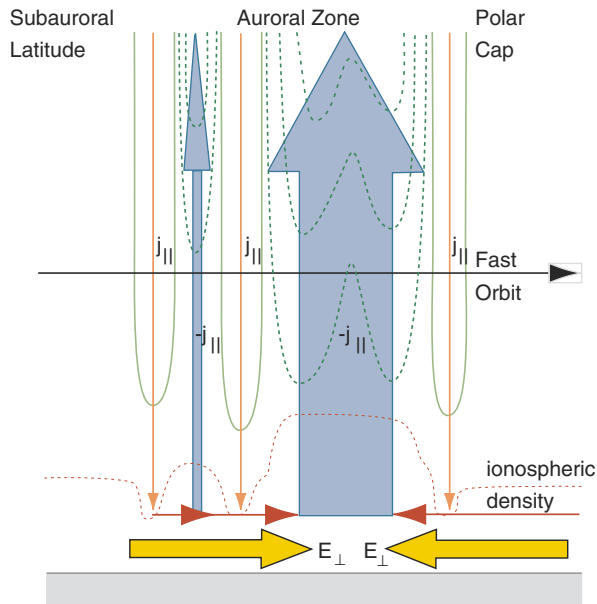


Fig. 4 The auroral cavity-current and electric-field system as explored by the FAST spacecraft. The auroral cavities are the broad regions of upward directed current flow ($-j_{\parallel}$, in blue) parallel to the (vertical) magnetic field-lines (light red). They close the perpendicular ionospheric Pedersen currents (indicated by the converging horizontal red arrows) which flow in the direction of the converging perpendicular electric fields E_{\perp} (yellow vectors). The upward $-j_{\parallel}$ currents correspond to downflow of keV auroral electrons which precipitate into the ionosphere and cause the enhanced ionospheric ionization. The negative potential in the upward current region (shown as dotted green lines) confines the auroral electrons to low altitudes. The enhanced ionospheric electron density is shown in dotted red. In the narrow regions between the two upward currents downward currents j_{\parallel} (light red) flow along the magnetic field from the magnetosphere into the ionosphere. They correspond to positive potentials (shown in solid green) and upward electron fluxes. The ray bar at the bottom is the surface of the Earth (after Elphic et al. 1998 with permission by the American Geophysical Union)

upward pitch angles. This region is of lesser interest for the electron-cyclotron maser.

3.3 Wave fields and radiation

For the AKR, we refer panel d of Fig. 3 which is the dynamic spectrum of wave activity observed during the present pass of the spacecraft. The horizontal black line in this panel at ~ 360 kHz indicates the local (non-relativistic) electron-cyclotron frequency f_{ce} .

Strong auroral kilometric radiation, the broad dark emission band at frequency above f_{ce} has been observed during the entire passage of the auroral zone on this day. This band is modulated at spacecraft spin period as seen in its higher frequency part. The modulation identifies the radiation as being in the RX mode perpendicular to the ambient magnetic field.

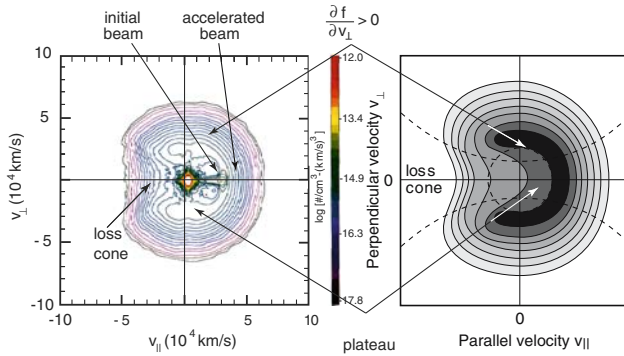


Fig. 5 *Left:* A typical auroral electron-velocity phase-space distribution function as measured by the FAST satellite at about 4,000 km altitude in the auroral region during auroral activity (after Delory et al. (1998), Pritchett et al. (1999) with permission by the American Geophysical Union). This distribution exhibits a loss-cone, a central isotropic cold plasma component around zero velocity which is most likely due to photoelectrons produced by the satellite itself and thus is not considered real, an energetic incomplete ring-shell (horseshoe) distribution with $\partial f / \partial v_{\perp} > 0$, a rudimentary beam component, and a broad nearly flat plateau. Note the nearly circular shape of the horseshoe part. *Right:* A schematic of such a measured distribution as used in calculations and simulations

When the spacecraft enters the auroral cavity at about 10 s the AKR band suddenly shifts below the local electron-cyclotron frequency. At this time the spacecraft has entered the radiation source region where it stays for a short while. Moreover, spin-modulated lower frequency emissions can be seen to occur at this time as well as fading out to low frequencies. These are non-escaping Z-mode (slow mode) emissions of the same polarization as the RX mode.

At frequencies below 30 Hz in this panel, separated by a wide gap in frequency from the radiation, one finds low-frequency waves with sharp upper frequency cutoff. This is the local plasma frequency f_{pe} . These emissions in the VLF band are of immense importance for the dynamics of the plasma. Outside the auroral cavity the emissions extend to higher frequencies, as the plasma is much denser there, and hence the plasma frequency is higher. However, in the auroral cavity with near absence of cold plasma, the VLF is more strongly confined.

Panel c of Fig. 3 shows the lowest 15 Hz of the dynamic VLF spectrum, and panel b its wave form. The latter, in this low resolution, exhibits occasional spiky behavior with peak amplitudes of a few mV m^{-1} in the cavity and larger amplitudes outside it in the downward current region. The VLF spectrum shows that in the cavity the whole VLF consists of spiky broadband bursts of emissions of two distinct types: one modulated in the same way as the radiation being roughly perpendicular to the magnetic field; the other being much more narrowband, strictly bound in frequency by the plasma frequency and about parallel to the magnetic field, such that these emissions are slightly out of phase.

In summary, the AKR source region, i.e. the source region of the electron–cyclotron maser has the following main global properties:

- intense field-aligned currents (in the Earth’s case directed upward),
- very dilute plasma density with $\omega_{pe} \ll \omega_{ce}$; in the present case this ratio is 0.1,
- field-aligned energetic electron fluxes (in the Earth’s case downward and of ~ 10 keV mean energy),
- strong upward parallel stationary electric fields diluting the plasma (in the Earth’s case of field strength $\sim 1 \text{ Vm}^{-1}$),
- ring-shell electron-velocity distributions in addition to the downward electron beam,
- upward collimated ion-beam distribution of same energy as downward electron beam,
- intense electromagnetic radiation in the source region at emission frequency $\omega < \omega_{ce}$,
- 100% circularly polarized radiation in the RX mode (fast electromagnetic mode),
- radiation at ω_{ce} is directed strictly perpendicular to the ambient magnetic field,
- at frequencies $\omega \leq \omega_{ce}$ there is weak radiation in Z-mode polarization (slow electromagnetic mode),
- low-frequency waves are confined to frequencies $\omega < \omega_{pe}$,
- two types of VLF waves are present: intense broadband bursts with highest frequencies exceeding ω_{pe} , and weak modulated VLF emissions confined below ω_{pe} ,
- the electric wave form at VLF modes exhibits spiky pulses in relation to the broadband bursts.

3.4 Fine structure of auroral kilometric radiation

One of the most pronounced properties of the AKR is its fine structure. It has been reported first by Gurnett and Anderson (1981). Figure 6 shows the example of a recent measurement of the auroral kilometric spectrum with the wideband instrumentation on the Cluster spacecraft SC1.

This observation which had been anticipated already much earlier by the Hawkeye (Gurnett and Anderson 1981) and again later at progressively lower distance from Earth by the Viking and FAST spacecraft suggests that the instantaneous AKR covers by no means a broad emission band. Rather it is emitted in narrow banded rays of bandwidth < 1 kHz within sources which move at comparably high speeds along the magnetic field-lines up and down depending on the conditions prevalent in the source region. In the present case the velocity is $\sim 300 \text{ km s}^{-1}$, roughly comparable to the ion acoustic speed. Upward drift in space is reflected by downward drifts in the frequency-time diagram since emission at the electron–cyclotron frequency as suggested by the electron–cyclotron maser mechanism implies that an upward moving source region in

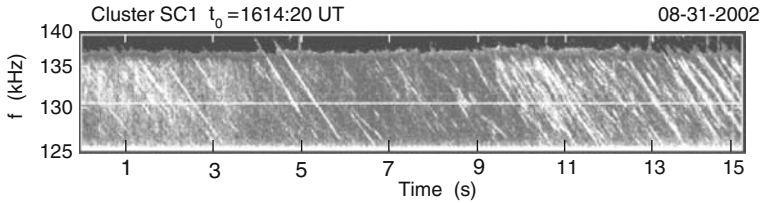


Fig. 6 A 15-s snapshot of wideband observations made by the Cluster mission, depicting auroral kilometric fine structure. This dynamic spectrum representation shows a smooth radiation background with many fast-drifting narrow-band emissions superimposed. The white line around 130 kHz is the local electron-cyclotron frequency at the orbital position of Cluster at the time when these measurements were taken at about 4 Earth radii distance away from Earth. The structures drift from high to low frequencies. This implies an upward drift of the radiation source away from Earth, i.e., towards lower magnetic field and thus lower electron-cyclotron frequency. Note that none of the structures are moving towards higher frequencies, i.e., downward in real space. This indicates that all the structures were generated below the Cluster spacecraft. The elementary radiators detected by Cluster have thus escaped from their generator at low altitudes along the magnetic field up to the altitude of Cluster

the vicinity of Earth passes from high to low magnetic dipole field strengths. Since the emissions cross the local electron-cyclotron frequency, one would also conclude that the radiation sources pass through the altitude of the spacecraft.

Figure 7 shows an observation with very high-time and frequency resolution by the FAST spacecraft of auroral kilometric radiation that was detected at a much lower altitude of only about 4,000 km above Earth's surface. The frequency of the detected emission fine structure is above the local cyclotron frequency, and thus the radiation source region is below spacecraft. This recording extending over only 2 s exhibits a wealth of fine structuring in the emission which is important as it provides information about the very radiation sources.

The first lesson learned from Fig. 7 is that obviously the entire AKR is coming from very fast moving 'elementary radiation sources,' the nature of which has to be identified. Below we will argue that these are all so-called electron holes which are produced under particular conditions in the plasma. The speed of these elementary radiation sources (or electron holes) as inferred from the drift of these elements across the dynamic spectrum of the AKR in the figure is of the order of a few 10^3 km s^{-1} . This value lies close to the electron thermal velocity of $\sim 1 \text{ keV}$ electrons which are known to be present in multitude in the AKR source region. In most cases it is faster than the time resolution.

The second lesson learned is that these elementary radiation sources sometimes arrange themselves in groups which drift as an ensemble through space and thus their emitted radiation drifts as an intense line in frequency space. The instantaneous width of such a line is of the order of 1 kHz which at the altitudes of the source near-Earth corresponds to roughly 1 km along the nearly vertical magnetic-dipole field lines. The arrangement of the elementary radiation sources in a line frequently resembles trapping like it is seen in Fig. 7 where the radiators seem to turn around and stay for some time until being reflected. Moreover, two such lines may interact (see the figure) without merging and

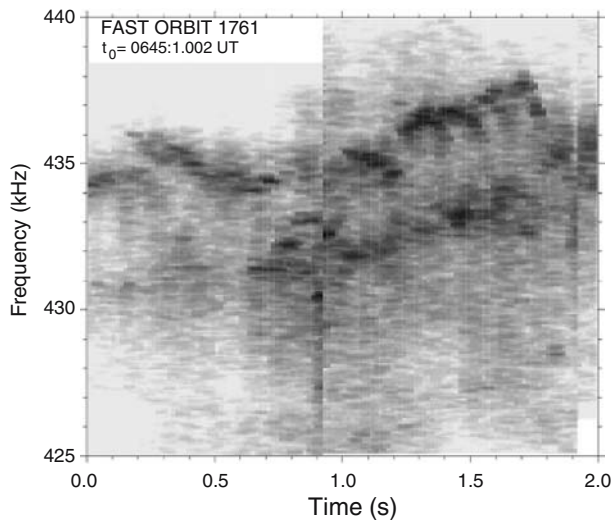


Fig. 7 A two-second snapshot taken with very high time- and frequency-resolution of a dynamic spectrum of auroral kilometric radiation above the radiation source region. The spectrum exhibits two narrow-band structures of radiation that move across the spectrum. The local bandwidth of the two structures is of the order of <1 kHz. Both structures are identified as the superposition of even narrower emission events which reach maximum intensity inside the structure, some of them even changing the direction of motion. Also note that the entire auroral kilometric-radiation emission seems to consist of such ‘elementary radiation events’ of very narrow bandwidth and high spectral velocity, $\partial\omega/\partial t$. The instrumental resolution reaches a bandwidth of ~ 0.1 kHz (after Pottelette et al. 1999 with permission by the American Geophysical Union)

after interaction moving together (in the case of the figure they move downward) as if two regions of similar potential reject each other, the region with the stronger momentum pushing the weaker one ahead. Similar effects have been found in numerical simulations (see below in the section on electron holes).

Finally, the last lesson from Fig. 7 is that the local bandwidth of the modulations of the AKR in one elementary radiation source can be as low as the resolution, viz. $\Delta f \simeq 0.1$ kHz corresponding to a field-aligned extension of the elementary radiation source of the order of $\simeq 100$ m or even less. Such a small bandwidth implies that the elementary radiation sources of the AKR, i.e. the electron–cyclotron maser radiators are very small-scale structures, not larger than of the order of the Debye length.¹⁵ Their observation requires in situ measurements in the source region. However, they can be inferred by a remote observer by their properties and the excitation conditions of the electron–cyclotron maser.

¹⁵ The Debye-length (or Debye radius) $\lambda_D = \sqrt{2m_e k_B T_e / \omega_{pe}}$, with T_e the electron temperature, defines the scale above that the plasma can be treated as an electrically quasi-neutral ‘fluid’, i.e., for lengths $L > \lambda_D$ the densities of electrons n_e and ions n_i can be taken to be approximately the same, $n_i \approx n_e$. On scales $L < \lambda_D$ the electric fields of the single charges of the plasma constituents cannot be neglected anymore. Here single-particle effects come into play and thermal fluctuations become important.

One of the characteristic features of the AKR also seen in the above figures is its pronounced fine structure (Gurnett and Anderson 1981). In high-resolution dynamical power spectra, the auroral kilometric radiation is observed to consist of many components that are closely spaced in frequency and have very narrow-band components that display a high degree of variability with time. The bandwidths span the range from few kHz (Gurnett and Anderson 1981) down to widths of the order of ~ 10 Hz. Such widths have been indirectly inferred first by Baumback and Calvert (1987) but can be recognized also in the high resolution spectra of Fig. 7. It is not yet clear what their nature is as they could be interferences, propagation effects, absorption effects, and also the direct signature of very small-scale structures — even smaller than what we have described as the elementary radiation events which seem to build up the entire spectrum of AKR. In the next section we are going to describe the contemporary state of the art of our knowledge of the electron–cyclotron maser radiation mechanism. Since AKR is the only paradigm on which this mechanism can be, and has been tested, we will always refer to it as the ultimate confirmation of whether or not the mechanism can be taken as a valid explanation for the direct generation of free space radiation.

4 Elementary physics of the cyclotron maser

Keeping in mind that the electron–cyclotron maser is a classical maser with no involvement of quantum correlation effects, we provide in this section simple physical though quantitative arguments for the operation of the cyclotron maser under the two conditions of mono-energetic and hot non-thermal distributions, the former related to electron-phase-space bunching, the latter related to phase-space distributions not in thermal equilibrium. It emphasizes in which way, classically, correlation and coherence can be achieved since this is one of the main properties of the electron–cyclotron maser. In this respect the very emission mechanism is of lesser interest.

4.1 Phase space bunching mechanism

Relativistic phase bunching in the gyrational motion of electron in strong magnetic fields had first been realized by Twiss (1958). The bunching is due to the relativistic mass dependence of the electrons at energies above a few keV transverse energy. Bunching at a particular gyrophase angle replaces quantum coherency in ordinary masers providing sufficiently strong correlation such that the electrons gyrate coherently and thus radiate coherently. This radiation is stimulated radiation because it is stimulated by an electromagnetic wave which the electrons are in phase with satisfying the resonance condition (1).

Such resonant electrons moving along a strong external magnetic field \mathbf{B}_0 at speed v_{\parallel} and together with right-hand polarized electromagnetic wave and gyrate at velocity \mathbf{v}_{\perp} feel the constant transverse electric wave field \mathbf{E}_{\perp} which can either decelerate or accelerate the electron perpendicular to the ambient

magnetic field, depending on whether the respective electrons rotate faster or slower than the wave. Faster electron will thus become retarded and slower electrons will be pushed until both have same phase angle as the wave. Mono-energetic electron beams will in this way bunch at one particular phase angle. The evolution of the kinetic energy of the electrons $W = (\gamma - 1)m_e c^2$ is then given by

$$\frac{dW}{dt} = -e\mathbf{v}_\perp \cdot \mathbf{E}_\perp = -eE_0 v_\perp \cos(\omega t - k_\parallel z - \phi), \quad (9)$$

if ϕ is the angle between \mathbf{v}_\perp and \mathbf{E}_\perp , and E_0 the wave amplitude. The angle ϕ evolves in time as

$$\dot{\phi} \simeq \omega_{ce}/\gamma. \quad (10)$$

On using the resonance condition (which now becomes a synchronism condition) one can easily show that the effective cyclotron angle $\phi_{\text{eff}} \equiv k_\parallel z + \phi$ of the electrons advances as

$$\dot{\phi}_{\text{eff}} \simeq k_\parallel z + \omega_{ce}/\gamma \equiv \omega_{ce,\text{eff}}. \quad (11)$$

Synchronism is thus reached when $\omega \simeq \omega_{ce,\text{eff}}$. This is illustrated in Fig. 8 which shows that the electrons tend to bunch in gyration phase angle along the wave magnetic field perpendicular to the wave electric field in order to minimize the change in energy. Bunching happens simply because electrons gaining energy on the left become heavier and rotate at a reduced cyclotron

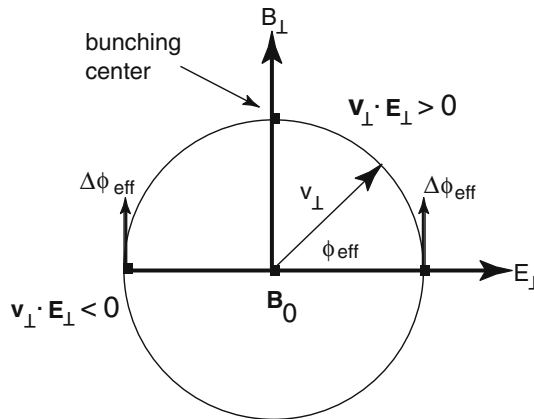


Fig. 8 Qualitative representation of the bunching process in velocity space. The circle indicates the rotation of the perpendicular velocity component with angle ϕ_{eff} against the wave electric field E_\perp in phase space during gyration. Particles in the left (right) half space have $\mathbf{v}_\perp \cdot \mathbf{E}_\perp < 0$ ($\mathbf{v}_\perp \cdot \mathbf{E}_\perp > 0$) and thus, according to Eq. (9), gain (lose) energy with the consequence that their effective phase angle moves toward the direction of the wave magnetic field. On both sides the changes in effective phase angle $\Delta\phi_{\text{eff}}$ point upward, since the relativistic mass of particles gaining energy increases while for particles losing energy it decreases. This causes the bunching of the particles perpendicular to the wave electric field along the wave magnetic field in order to minimize the energy change. Electrons arrange at a phase where the transverse wave electric field in their frame of reference vanishes, i.e. at $\mathbf{v}_\perp \times \mathbf{B}_\perp = 0$, as they always do in collisionless plasmas

frequency ω_{ce}/γ , while electrons which loose energy on the right increase their masses and rotate faster. Since all the particles attain the same phase angle, they bunch and moving together like a single macro-particle. One should note that the bunching at constant magnetic field does not contribute to radiation since the emission and absorption on both sides in Fig. 8 cancel. However, when the magnetic field is detuned the correlated motion of the bunch leads to coherent emission, an effect exploited in the laboratory ‘gyrotron emission’. In addition to this bunching in the cyclotron motion there is also bunching due to the motion parallel to the magnetic field which acts in a similar way through the term $k_{\parallel}v_{\parallel}$ in the resonance condition which as well changes the effective phase angle. Both bunching mechanisms compete, and partially cancel (Chu 2004).

The above arguments for bunching involve only the very simple and basic physics of single particle motion in an electromagnetic wave in a strong magnetic field. It neglects the very complicated (chaotic) phase space dynamics of electrons in a strong electromagnetic wave to which the wave piles up in maser emission. Thus they hold only for the initial linear state. Direct application is thus restricted only $|\delta|t \equiv |\omega - k_{\parallel}v_{0z} - \omega_{ce}/\gamma_0|t \ll 1$, where the index 0 refers to the initial values of particle velocity and energy. In addition, application is restricted to a cold mono-energetic particle distribution only. Extensions have been given to loss-cone distributions by Lau and Chu (1983) and to spiraling magnetic-field aligned electron beams by Kho and Lin (1988).

4.2 Radiation and absorption in bunching

In the bunching process radiation can either be emitted or absorbed depending on whether $\omega_{ce,eff} < \omega$ or $\omega_{ce,eff} > \omega$, respectively. This is easily seen if for instance detuning the ambient magnetic field initially according to the first of these conditions. In this case the electrons in order to escape the $\mathbf{v}_{\perp} \times \mathbf{B}_{\perp}$ -electric field rotate as a whole in the clockwise sense during bunching with the bunch center slipping into the right half-plane in Fig. 8 resulting in a net transfer of energy to the wave resulting in coherent radiation. In the opposite case energy the motion is anti-clockwise and energy is transferred to the particles. Thus, depending on their initial phase ϕ_0 the electrons may radiate or absorb. The average rate of change in their energy $\gamma - 1$ is given by

$$\langle \dot{\gamma} \rangle = \int_0^{2\pi} \frac{\dot{\gamma} d\phi_0}{2\pi}. \quad (12)$$

On the basis of this expression, the kinetic energy transfer efficiency at time t is defined through

$$\eta_{\gamma}(t) \equiv -\frac{1}{\gamma_0 - 1} \int_0^t \langle \dot{\gamma} \rangle dt \quad (13)$$

which when expanded with respect to time gives the linear efficiency at $|\delta|t \ll 1$ in the form

$$\eta_{\gamma,\text{lin}}(t) \approx \frac{e^2 E_0^2 / 2m_e^2 c^2}{\gamma_0(\gamma_0 - 1)} \left[\frac{v_{\perp 0}^2 |\delta|}{12\omega} (\omega^2 - k_z^2 c^2) t^4 - \left(1 - \frac{v_{\perp 0}^2}{2c^2} \right) t^2 \right]. \quad (14)$$

This expression shows that the emission on the short time scale is $\propto (|\delta|t)^4$, while absorption is $\propto (|\delta|t)^2$. Thus initially absorption dominates until emission takes over. Synchronism is maintained until $|\delta|t \sim \pi$. This yields the following ratio for the emitted to absorbed energies:

$$\frac{W_{\text{em}}}{W_{\text{abs}}} \sim \frac{(v_{\perp 0}/c)^2}{1 - (v_{\perp 0}/c)^2}, \quad |\delta|t \sim \pi. \quad (15)$$

So, as expected, emission efficiency decreases with decreasing perpendicular electron energy. These expressions remain valid for a cold particle distribution $f_{v_{\perp}} \propto \delta(v_{\perp} - v_{\perp 0})/v_{\perp}$ where $v_{\perp 0}$ is the bulk perpendicular speed of the particles. For warm distributions with finite velocity spread one would have to integrate all these expressions over the distribution function in order to obtain the emission and absorption coefficients. However, the above expressions for the cold component indicates that emission is nothing else but the generation of a negative absorption coefficient, and this makes up the similarity to a maser.

The above emissivities and absorptivities which are valid for one single electron (or a bunch of non-interacting zero temperature relativistic electrons) could be generalized to other distributions simply by integrating over the distribution weighted expressions. It is, however, more straightforward to determine them in a plasma stability approach. Emission of radiation can then, in a plasma physical picture, be interpreted as a simple (fluid-like) plasma instability pumping a free space electromagnetic mode (Chu and Hirshfield 1978; Winglee 1983). To this end one solves the linearized relativistic Vlasov equation for the electron-distribution function

$$\partial_t f_1 + \mathbf{v} \cdot \nabla f_1 - e \mathbf{v} \times \mathbf{B}_0 \cdot \partial_{\mathbf{p}} f_1 = e (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \partial_{\mathbf{p}} f_0, \quad (16)$$

for the disturbed distribution $f_1(t, \mathbf{x}, \mathbf{p})$ in the presence of the uniform background magnetic field and the undisturbed distribution function $f_0(\mathbf{p})$, where $\mathbf{p} = \gamma \mathbf{v}$ is the relativistic momentum. Taking the undisturbed distribution as one of the cold though relativistic plasma $f_0 = \delta(p_{\parallel}) \delta(p_{\perp} - p_{\perp 0}) / 2\pi p_{\perp}$ and using conventional methods to solve the linearized Vlasov equation (16) and Maxwell's equations for plane waves of frequency ω and wave number \mathbf{k} yields the linear dispersion relation

$$\omega^2 - k_{\parallel}^2 c^2 = \frac{\omega_{\text{pe}}^2}{\gamma_0} \left[\frac{\omega}{\omega - \omega_{\text{ce}}/\gamma_0} + \frac{k_{\perp}^2 v_{\perp 0}^2 (1 - \omega^2/k_{\parallel}^2 c^2)}{2(\omega - \omega_{\text{ce}}/\gamma_0)^2} \right]. \quad (17)$$

The real k , complex $\omega = \omega_r + i\Gamma$ solution to this expression is shown in Fig. 9 (Chu and Hirshfield 1978) (this was the paper which led Wu and Lee (1979) to start considering the relativistic correction in the development of their successful theory of terrestrial kilometric radio emission from the aurora). One realizes that under certain conditions which coincide with the bunching process the two possible free space electromagnetic modes (the ‘fast’ RX and ‘slow’ Z circularly polarized waves) propagating in the plasma can be directly amplified by the plasma. The important conditions are that the plasma must be relativistic even if weak, and that it must be extremely dilute with $\omega_{ce}^2/\omega_{pe}^2 \ll 1$ as already mentioned in the introduction to this paper. Note that it is easy to derive the above phase bunching conditions also in this plasma case by inspecting the evolution of $\Delta\omega_{ce,eff} = k_{\parallel}\Delta v_{\parallel} - \Delta\gamma\omega_{ce}/\gamma^2$ under the evolution of particle energy [see, e.g., Chu (2004), his Eq. (73)].

Instability under these conditions for both modes occurs close to the electron-cyclotron frequency. Whether the amplified slow wave can escape from the plasma is a different one and - is difficult to answer. Analytical and numerical calculations have been presented by Kho and Lin (1988) for the slow wave in a field-aligned beam plasma system and $\gamma_0 = 2, \omega_{ce} = 2\omega_{pe}$ finding a strong instability. It seems that the slow wave can be more easily be made unstable when a beam along the magnetic field is present in the plasma. Another extension of the above theory has been given by Lau and Chu (1983) to a loss-cone plasma which applies to magnetically confined plasmas of the kind present in confinement devices and in many natural systems. This will be discussed briefly in the next subsection.

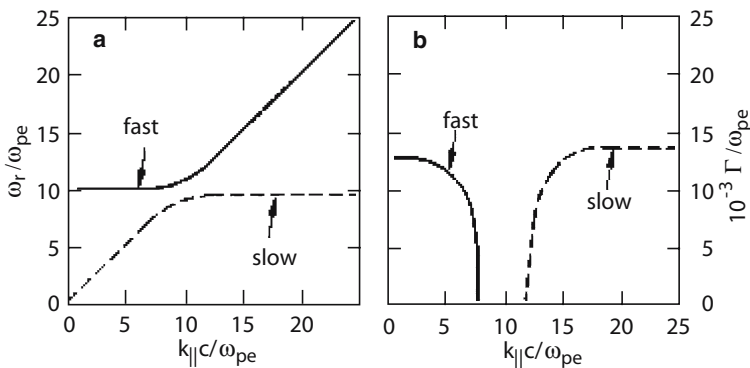


Fig. 9 Dispersion relations $\omega_r(k_{\parallel})$ of the two electromagnetic free-space modes, the ‘fast’ and ‘slow’ waves, in a cold relativistic plasma and their growth rates $\Gamma(k_{\perp})$ (after Chu and Hirshfield 1978). Note that the scale on the growth rate is the same as on the frequency except that the growth rate is multiplied by a factor of 10^{-3} . Hence, the growth rates are only 0.1% of the frequency for the chosen parameter range of a weakly relativistic $\gamma_0 = 1.02$ underdense plasma with $\omega_{ce}/(\gamma_0\omega_{pe}) = 10$. With this normalization the unstable ‘fast’-mode branch is amplified at a frequency very close to the electron-cyclotron frequency ($\omega_r \approx 10\omega_{pe} \approx \omega_{ce}$). The slow wave is amplified at a slightly lower frequency

5 Electron-cyclotron maser theory

In the following we go into greater detail of the electron-cyclotron maser mechanism. We start with a review of the loss-cone maser even though it has become highly improbable that it is of the importance which had been originally attributed to it. We then come to the mechanism of the ring-shell maser and, finally to the question of how the magnetic field-aligned electric potential drops can be generated which will become the subject of the next section.

5.1 Review of the loss-cone maser

Lau and Chu (1983) following Chu and Hirshfield (1978) considered the following problem: Let a right-handed polarized electromagnetic wave with wave frequency close to the electron-cyclotron frequency propagating along a strong ambient magnetic field interact with a warm weakly relativistic loss-cone distribution of electrons

$$f_0(p_\perp, p_\parallel) = A(p_\perp)^{2j} \exp \left[\frac{-p^2}{(\Delta p)^2} \right]. \quad (18)$$

An example of a loss-cone distribution as observed in the Earth's auroral region is shown schematically in Fig. 12. Such distributions imply that there the probability to find a particle at small pitch angles is zero, describing totally emptied loss cones in a mirror magnetic field geometry and had been investigated previously by Dory et al. (1965) in the context of electrostatic loss-cone instabilities in hot non-relativistic plasmas. A is a normalization constant, and $j > 0$ is the loss-cone index. The above distribution is a relativistic distribution if properly normalized as $\int d^3p f_0 = 1$. For $j = 0$ the distribution becomes the famous Boltzmann-Jüttner distribution. Δp is a measure of the electron temperature $k_B T = m_e c^2 [1 + (\Delta p)^2 / m_e^2 c^2]^{\frac{1}{2}} - m_e c^2$. The distribution peaks at $\sqrt{j} \Delta p$. Hence for $j \neq 0$ there is a 'population inversion' in p_\perp indicating 'excitation' of electrons as necessary in masers.

Linear Vlasov-Maxwell theory of the same kind as given in the previous subsection becomes slightly more involved due to the presence of the loss-cone factor. The dispersion relation (2) replacing (17) now becomes

$$\omega^2 - k_\parallel^2 c^2 = 2\pi \omega_{pe}^2 \int_0^\infty \int_{-\infty}^\infty dp_\parallel \frac{f_0}{\gamma} \left[\frac{\omega - k_\parallel p_\parallel / m - e\gamma}{\omega - k_\parallel p_\parallel / m_e \gamma - \omega_{ce} / \gamma} - \frac{p_\perp^2 (\omega^2 - k_\parallel^2 c^2)}{2\gamma^2 m_e^2 c^2 (\omega - k_\parallel p_\parallel / m_e \gamma - \omega_{ce} / \gamma)^2} \right], \quad (19)$$

an expression which has to be solved numerically. Solutions are shown in Fig. 10 and 11.

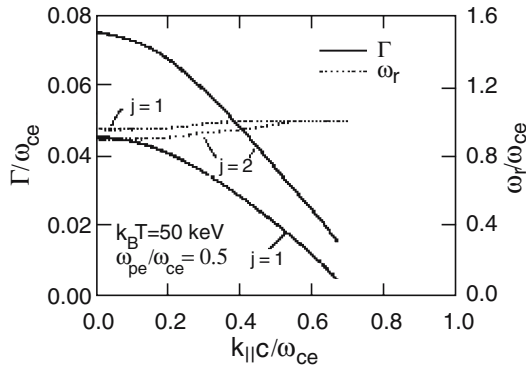


Fig. 10 Growth rate Γ/ω_{ce} and frequency $\omega_r(k_{||})/\omega_{ce}$ of the parallel propagating unstable electromagnetic mode excited in the presence of an energetic electron loss-cone distribution. The loss-cone parameters chosen are $j = 1, 2$. The electron temperature is $k_B T_e = 50$ keV, and the plasma-to-cyclotron frequency ratio is $\omega_{pe}/\omega_{ce} = 0.5$ (after Lau and Chu 1983). The frequency (dotted line) changes very little over the unstable domain in $k_{||}$. The small wave numbers of the fast mode are driven unstable at a substantial growth rate of several percent of the electron–cyclotron frequency

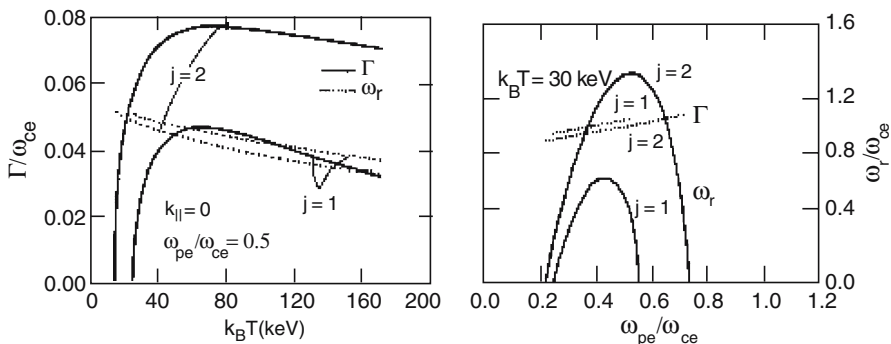


Fig. 11 *Left:* The growth rate and frequency of the loss-cone driven maser emission as function of electron temperature. For instability the electron temperature must exceed a certain threshold value which decreases with increasing loss-cone parameter j . Increased ‘emptiness’ of the loss-cone favors the emission and thus the maser action. *Right:* Dependence of growth rate and frequency of the loss-cone driven maser emission as function of the ratio between plasma frequency and electron–cyclotron frequency at electron energy $k_B T_e = 30$ keV. The important information that can be taken from this figure is that the maser instability occurs only in underdense ($\omega_{pe} < \omega_{ce}$) plasmas and in a narrow range of ratios between plasma frequency and cyclotron frequency (after Lau and Chu 1983)

What is hidden in the above expression is the location where the free energy in the distribution function is located. This free energy is necessarily a source of radiation in the electron–cyclotron maser. It can, however, be identified by inspecting (16) more closely. The above dispersion relation has resulted from an integration over the right-hand side of (16) along the unperturbed particle orbits. This integration yields an expression for the dielectric function ϵ which we write down here for the RX-mode only:

$$\epsilon_{\text{RX}}(\omega, \mathbf{k}) = \begin{pmatrix} 1 + A - N_{\parallel}^2 & -iA & N_{\parallel}N_{\perp} \\ iA & 1 + A - N_{\perp}^2 & 0 \\ N_{\parallel}N_{\perp} & 0 & 1 - N_{\perp}^2 \end{pmatrix}. \quad (20)$$

$N_{\parallel, \perp} = k_{\parallel, \perp} c / \omega$ are the parallel and perpendicular indices of refraction, respectively, of the wave under consideration, in this case the RX-mode, and the quantity A which follows from (4–5) is given by

$$A = \frac{\pi}{2} \frac{\omega_{\text{pe}}^2}{\omega^2} \int_{-\infty}^{\infty} du_{\parallel} \int_0^{\infty} u_{\perp}^2 du_{\perp} (\gamma \omega - k_{\parallel} u_{\parallel} - \omega_{\text{ce}})^{-1} \\ \times \left[\omega \frac{\partial}{\partial u_{\perp}} + \frac{k_{\parallel}}{\gamma} \left(u_{\perp} \frac{\partial}{\partial u_{\perp}} - u_{\parallel} \frac{\partial}{\partial u_{\parallel}} \right) \right] f_0. \quad (21)$$

For simplicity the massless components of the electron momentum $u_{\parallel} = p_{\parallel} / m_e$, $u_{\perp} = p_{\perp} / m_e$ have been introduced here. This expression shows that the parallel momentum component and its derivatives are compensated by the resonant denominator $(\gamma \omega - k_{\parallel} u_{\parallel} - \omega_{\text{ce}})^{-1} \approx i\pi \delta(\gamma \omega - k_{\parallel} u_{\parallel} - \omega_{\text{ce}})$. The only dependence which remains is the partial derivative with respect to the perpendicular momentum $\partial f_0 / \partial u_{\perp}$ which determines the sign of the maser growth rate Γ . The growth rate is positive only for positive perpendicular velocity gradients in f_0 . Population inversion, required for maser emission, thus implies that the particle distribution does not decay monotonically with increasing perpendicular momentum u_{\perp} like it is found in thermal equilibrium distributions. Instead it is required that $\partial f_0 / \partial u_{\perp} > 0$. Figure 12 shows such a behavior in the region of the loss-cone along the right-hand part of the elliptic line. The relevance of this line is discussed below.

Figures 10 and 11 do in principle already contain all the information about the effect of the loss-cone maser. Its efficiency increases the wider the loss-cone is, i.e. the larger j , and the higher the electron energy is. The electron temperature must exceed a certain threshold for the maser to overcome absorption. However, in this calculation there is no assumption of the effect of a possible plasma background which should usually be expected to be present under natural conditions where the loss-cone particle distribution represents, however, a dilute energetic tail distribution of magnetically mirroring particles on a much denser background. Hence the question arises as to how a maser could work in presence of dense background plasmas? It turns out that this is the key question to the whole maser project under natural conditions. Intuitively, one would suggest that a less energetic and isotropic background plasma will inhibit maser action by imposing a high absorbing power. This is basically true as all investigations have shown, leading to the result that it is instrumental for the electron–cyclotron maser to have either no background at all or at least a very dilute one in order to minimize background absorption. Before proceeding we will have to more closely analyze the resonance condition (1) and afterwards return to the most efficient particle distributions in maser action.

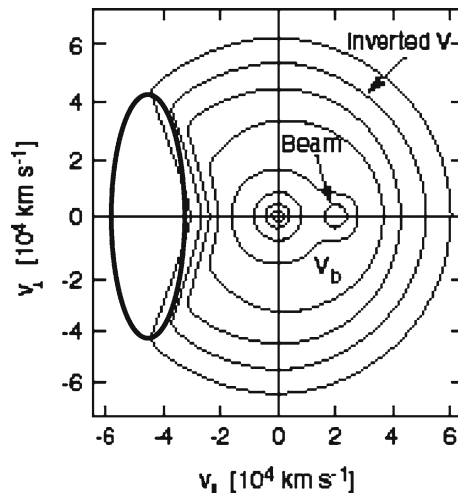


Fig. 12 Sketch of an auroral electron-distribution function in the Earth's auroral region including the loss cone (after Ergun et al. 1993). The loss cone is the triangular shape on the left of the distribution function in velocity space, where one finds a lack of electrons. These electrons are, in the case of the auroral region, lost to the atmosphere since their velocities are mostly parallel to the magnetic field. Particles at velocities with such angles with respect to the magnetic field are not reflected by the magnetic field and thus do not return into the plasma. They are lost. The ellipse drawn inside the loss-cone shows the position of the resonance ellipse for loss-cone maser radiation (see text)

5.2 Analysis of the resonance condition

The main contribution to the cyclotron maser instability comes from particles satisfying the resonance condition (1). This resonance condition is the denominator under the integral sign which appears for instance in (24), (21) or generally in the explicit form of the electromagnetic dispersion relation (2). This dispersion relation requires explicit knowledge of the relativistic dielectric function $\epsilon(\omega, \mathbf{k})$ as has been given in (2). The maximum contribution to the integral in (3) is provided by those particles satisfying the condition of vanishing denominator. In non-relativistic plasmas, these are just those particles which move along the magnetic field with parallel phase velocity of the wave shifted by the cyclotron frequency. Resonant particles experience the stationary electric wave field and are thus either accelerated or retarded. In the relativistic case the cyclotron frequency is modified by the relativistic factor ω_{ce}/γ , and hence the resonant frame depends on particle energy and differs for all particles. Setting the relativistic resonance condition (1) to zero, the resonance condition can be rewritten in terms of the electron velocity components v_{\parallel}, v_{\perp} in the form of the equation of an ellipse

$$\frac{(\beta_{\parallel} - \beta_{\parallel 0})^2}{a^2} + \frac{\beta_{\perp}^2}{b^2} = 1, \quad (22)$$

where $\beta_{\parallel,\perp} = v_{\parallel,\perp}/c$. The center of this ellipse is shifted out of origin along the v_{\parallel} -axis by

$$\beta_{\parallel 0} = \frac{N_{\parallel}}{N_{\parallel}^2 + l^2 \omega_{ce}^2 / \omega^2}. \quad (23)$$

The half axes of the ellipse are given by

$$\begin{aligned} a^2 &\equiv \frac{b^2}{1 + (N_{\parallel} \omega / l \omega_{ce})^2} \\ b^2 &\equiv 1 + (N_{\parallel}^2 - 1)(\omega / l \omega_{ce})^2 \end{aligned} \quad (24)$$

There is a set of resonant ellipses, one for each value of l . Figure 13 shows a schematic of the resonance curves in the velocity space for a non-relativistic case (curve a), when the resonance line is a straight line, and the relativistic case (b) when it is an ellipse, as in Fig. 12 shifted to the position of the loss-cone in a loss-cone plasma including a thermal (cold) electron component. It is easy to see that these ellipses in velocity space are all confined entirely to the interior of the circle with radius $v = c$. Since the ellipses are the geometric locations of the resonances in velocity space, integration in (19 (or any other equivalent equation)) has to be performed along these ellipses. Moreover, one observes that for finite parallel wave numbers $k_{\parallel} \neq 0$, i.e. for oblique propagation to the ambient magnetic field there will always exist a displaced resonance ellipse for given harmonic l which has to be positioned in such a way that for a given harmonic wave $\omega = \omega(k_{\parallel}, k_{\perp})$ it collects all the contributions from the particles which carry the free energy in order to maximize wave excitation. (One such ellipse is shown in Fig. 12 superimposed on a measured though still schematic loss-cone distribution.) Thus for small parallel phase velocities, the ellipse becomes highly elongated with parallel phase velocity small against the velocity of light, while

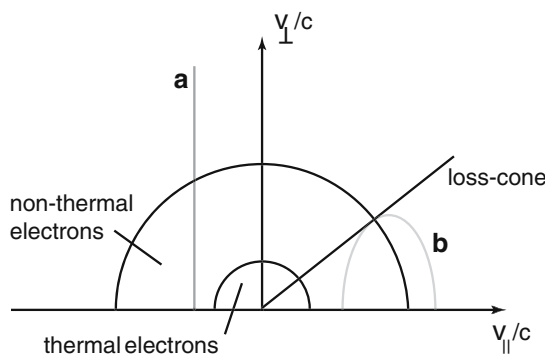


Fig. 13 Resonance lines for the non-relativistic case **a** and the relativistic case **b** in the velocity space of a loss-cone distribution of hot non-thermal particles and a thermal cold electron component. The non-relativistic resonance line is a *straight line*, while the relativistic resonance line is *ellipse* (as already shown on the data displayed in Fig. 12). This ellipse has been shifted into the loss-cone

in the opposite case when the parallel phase velocity is large against the velocity of light, the ellipse becomes nearly circular.

The observation that the unstable frequency in the relativistic electron–cyclotron maser falls below the electron–cyclotron frequency is somewhat disturbing since the RX-mode has a low-frequency cutoff which is given by (8) which is clearly above ω_{ce} . For $\omega_{ce} \gg \omega_{pe}$ this expression simplifies to $\omega_x \approx \omega_{ce}(1 + \omega_{pe}^2/\omega_{ce}^2) > \omega_{ce}$, slightly above the electron–cyclotron frequency. The unstable RX-maser wave to be able to propagate this expression has to be examined for the relativistic case, which has been done by Pritchett (1984a,b), (Le Quéau et al. 1984a,b). Pritchett (1984b) found that the RX cutoff for a weakly relativistic Maxwellian distribution depends on the electron energy through the parameter $\mu = m_e c^2 / k_B T_e$. This is illustrated for the particular case of $k = 0$ in Fig. 14. The frequency ω of the RX mode crosses the electron–cyclotron frequency from above to below at $\mu \leq \frac{3}{2}(\omega_{ce}^2/\omega_{pe}^2)$ and for a certain range below this value stays below ω_{ce} . For $\omega_{ce}/\omega_{pe} = 10$ the transition occurs for electron energies larger than 150 eV. Hence the electron–cyclotron maser unstable RX mode can indeed propagate below ω_{ce} in a limited electron energy range.

Let us briefly discuss the relevance of the resonance ellipse for the loss-cone maser as illustrated schematically in Fig. 12. The resonance ellipse with its right-hand shoulder passes along the loss cone where $\partial f_0 / \partial u_{\perp} > 0$. Here the wave that satisfies the resonance condition along the ellipse becomes amplified. The left-hand shoulder of the ellipse on the other hand encounters a negative gradient in the velocity distribution and thus causes absorption and damping on the same wave. However, the number of absorbing particles is very small in comparison to the number of emitting particles such that along the resonance ellipse emission dominates. Any other resonant ellipse put into the distribution would belong to different resonant waves and would cover only absorbing particles. Hence, the only wave which will become amplified is the one which generates the resonance ellipse along the boundary of the loss cone. Clearly, the number of particles with positive perpendicular velocity space gradient in the case of the loss-cone is small, and the loss-cone maser will therefore be rather weak.

This poses the question for a more efficient particle distribution than the loss-cone distribution. The first observation is that any resonant ellipse will always cover only a small amount of resonant particles unless the distribution function is exotic. For instance, if the resonant ellipse degenerates to a circle and the distribution is ringlike, the number of particles with positive perpendicular velocity gradient will become large and, hence, the emission becomes much stronger. The resonance condition degenerates into a circle for $N_{\parallel} = 0$ which implies strictly perpendicular propagation which is possible only for the RX-mode in magnetoionic theory (Fig. 1). From Eq. (22)–(24) one learns that in this case $\beta_{\parallel 0} = 0$. The resonance line is centered at the origin. Moreover, $a = b$ and the resonance line is a circle of radius $(1 - \omega^2/l^2 \omega_{ce}^2)^{\frac{1}{2}}$. For the fundamental $l = 1$ this implies that $\omega < \omega_{ce}$. Instability and emission will thus be at a frequency which is slightly below the non-relativistic cyclotron frequency. Thus a particularly promising distribution function for intense electron–cyclotron

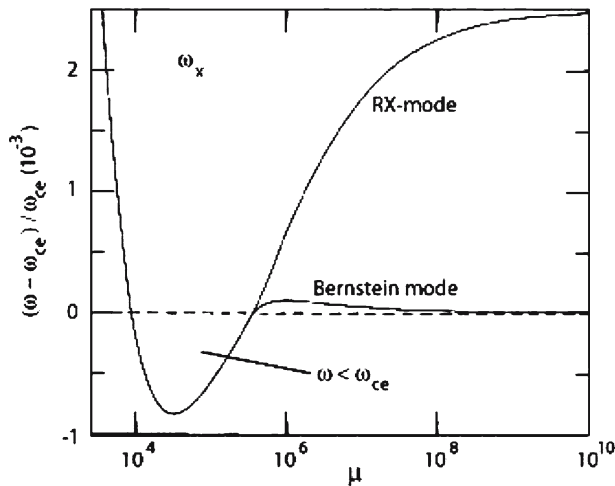


Fig. 14 The relativistic RX-mode cutoff, ω_x , versus the inverse particle temperature μ , and for $\omega_{ce}/\omega_{pe} = 10$ and the marginal case of wave number $k = 0$ at which the wave degenerates and is reflected. For orientation also shown is the $k = 0$ perpendicular electrostatic first-harmonic Bernstein mode. This mode is strictly perpendicular and separates from ω_x at large μ (after Pritchett 1984a)

maser emission is a hot ring distribution, which poses the question of whether or not such ring distributions can exist under natural conditions.

Before investigating this question let us turn to the generation of emission. This is in the RX mode as previously argued. Setting the determinant of (21) to zero, one obtains the dispersion relation for the RX mode Le Quéau et al. (1984a)

$$1 - N^2 + (2 - N_{\perp}^2)A = 0. \quad (25)$$

For $k = 0$, one obtains $A = -\frac{1}{2}$. In this limit the imaginary part of A vanishes: $A_i = 0$. On the other hand, for $k_{\parallel} = 0$ and $k_{\perp}c/\omega_{ce} \rightarrow 1$ and assuming that emission is close to the fundamental, $|\omega - \omega_{ce}| \ll \omega_{ce}$, which is satisfied in view of the above discussion, the real part of A is small. To first order one has that

$$\Gamma = -\frac{1}{2}\omega_{ce}A_i. \quad (26)$$

Then in the semi-relativistic approximation where γ is expanded for small u^2/c^2 and $1/\gamma \approx 1 - u^2/2c^2$ the expression for A becomes

$$A = \frac{\pi}{2} \frac{\omega_{pe}^2}{\omega} \int_{-\infty}^{\infty} du_{\parallel} \int_0^{\infty} du_{\perp} u_{\perp}^2 \frac{\partial f_0(u_{\perp}, u_{\parallel}) / \partial u_{\perp}}{\omega - \omega_{ce} + \omega_{ce}u^2/2c^2}, \quad (27)$$

where $u^2 = u_{\perp}^2 + u_{\parallel}^2$. Since the growth rate will be small such that $\Gamma \ll \omega_{ce}$ can be safely assumed, and defining $u_0^2 = 2c^2(1 - \omega/\omega_{ce})$ the resonant denominator

can be replaced by a delta-function yielding

$$A_i = -\frac{\pi^2 c^2}{2} \frac{\omega_{pe}^2}{\omega \omega_{ce}} \int_{-\infty}^{\infty} du_{\parallel} \int_0^{\infty} du_{\perp}^2 u_{\perp} \frac{\partial f_0}{\partial u_{\perp}} \delta(u^2 - u_0^2). \quad (28)$$

One concludes that for strictly perpendicular radiation $\omega < \omega_{ce}$, permitting perpendicular maser instability and emission only at frequencies below the electron-cyclotron frequency are permitted.

Applying these expressions to the loss-cone distribution (18) one defines $x_0^2 \equiv m_e u_0^2 / (\Delta p)^2$ and evaluates the above integrals to obtain an analytic expression for the growth rate in perpendicular emission

$$\Gamma = \sqrt{\pi} \frac{2^{j+1}}{(2j+1)!!} \frac{\omega_{pe}^2}{\omega \omega_{ce}} \frac{m_e c^2}{\Delta p^2} x_0^{2j+1} e^{x_0^2} \left(j - \frac{2(j+1)}{2j+3} x_0^2 \right) \quad (29)$$

from which it is clear that Γ vanishes for vanishing last bracket in this expression. At larger x_0^2 absorption dominates. Hence the range of applicability of the resonance circle in this case to the loss-cone distribution is rather restricted as discussed above. We will thus have to turn to another distribution function which is more promising for generating intense maser emission, the ring distribution.

Pritchett (1986a) has performed numerical simulations of the weakly relativistic loss-cone maser instability. Figure 15 shows his results for the total emitted

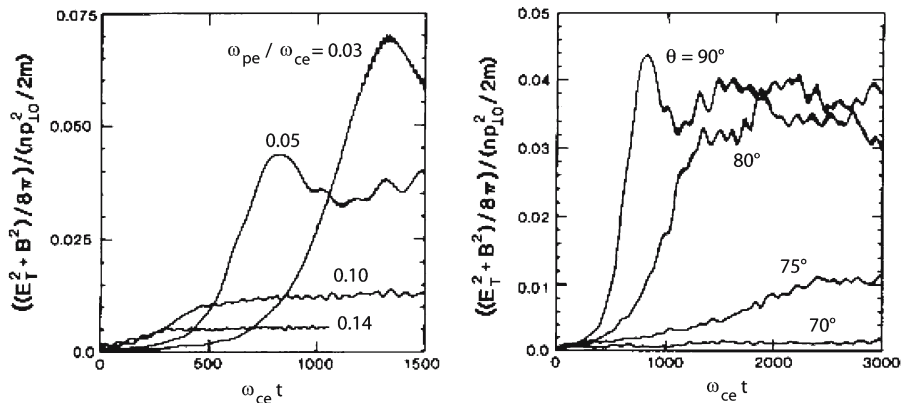


Fig. 15 Numerical simulations of the loss-cone maser instability in strictly perpendicular direction for loss-cone parameter $j = 2$ and various ratios ω_{pe}/ω_{ce} (left), and for several oblique angles at $\omega_{pe}/\omega_{ce} = 0.05$ (right). The normalized total emitted electromagnetic energy in the radiation is shown as a function of time that has been normalized to the electron-cyclotron period. Normalization is to the initial perpendicular kinetic plasma energy. Note that the exponential increase during linear instability which is followed by saturation. The larger the ratio ω_{pe}/ω_{ce} , the weaker is the maser radiation. Similarly, the smaller the emission angle against the magnetic field the weaker is the radiation (after Pritchett 1986a)

electromagnetic energy density. Several remarkable results can be read from this figure. First, the emitted wave energy in all cases passes through an initial exponential rise phase which shows that the maser instability initially can be described by linear theory. After a certain time it reaches saturation. the level of which depends on the plasma to cyclotron frequency ratio. The smaller this ratio the higher is the saturation amplitude. Also, the rise time (growth rate) is a function of this ratio. It is longer for smaller ratio simply because less particles participate in radiation. However, the more particles are involved the stronger is the self-limitation by self-absorption of the radiation. In addition, the more the angle of radiation emission turns away from the perpendicular direction the weaker is the emission, and at an angle of $\sim 70^\circ$ it drops to background noise level, implying that emission takes place in a narrow cone perpendicular to the magnetic field. Finally, the emptier the loss-cone which means the steeper the perpendicular gradient in velocity space, the more intense is the radiation. However, altogether these simulations show that the loss-cone maser is rather inefficient since only very few particles are actively involved into the emission, and other distribution functions have to be investigated.

5.3 The ring-shell maser

The observation that in ring distributions the emission at $k_{\parallel} \approx 0$ is about perpendicular to the magnetic field and that the fundamental emission is beneath the electron–cyclotron frequency restricts the emission at the first place to the RX mode. Figure 16 (top) shows the different strictly perpendicular wave dispersion branches for the first three harmonics $l = 1, 2, 3$ in a plasma dominated by such a cold ring-shell distribution

$$f_0(p_{\perp}, p_{\parallel}) = \frac{1}{2\pi p_{\perp}} \delta(p_{\perp} - p_R) \delta(p_{\parallel}). \quad (30)$$

In such a distribution thermal effects are neglected, and it is assumed that the thermal background is much weaker such that it plays only a minor role. Here p_R is the ring momentum which in Fig. 16 is taken as $p_R = 0.4c$, and $\omega_{ce}/\omega - pe = 7.5$. These branches are indicated by the numbers at each of the curves. The top of the figure shows the coupling between the various branches: RX and purely electrostatic perpendicular Bernstein modes. The coupling regions are responsible for the maser instability. Excitation of the RX-mode is not in a direct way. The perpendicular gradient in velocity space drives relativistic Bernstein modes which at the electron–cyclotron harmonics couple to the RX-mode. In the non-relativistic case such coupling can occur only for extreme temperature anisotropies. In the lower part of the figure the electron–cyclotron maser growth rate $\Gamma(k_{\perp})$ is given as function of the perpendicular wave number for all three harmonics shown. The growth rate is non-zero just below the harmonic wave numbers $k_{\perp}c/\omega_{ce} = l$ and for $\omega < \omega_{ce}$ and decreases rapidly with increasing l , and at small k_{\perp} increases as k_{\perp}^l . Largest Γ s are found at $l = 1$, the fundamen-

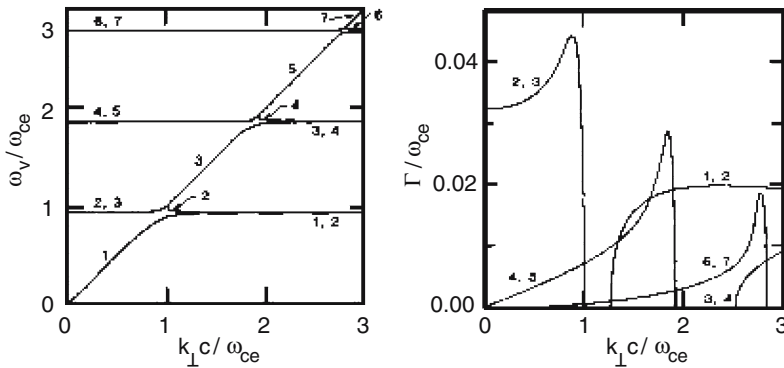


Fig. 16 Weakly relativistic dispersion curves for the RX-modes in the electron ring-shell distribution for perpendicular propagation $k_{\parallel} = 0, k_{\perp} \neq 0$ and the first three lowest harmonics $l = 1, 2, 3$ (after Pritchett 1984b with permission by the American Geophysical Union). *Left:* Real frequency dispersion curves showing the coupling between the RX- and Bernstein modes at the harmonics, numbered from 1 to 6. *Right:* Growth rates Γ for the three harmonics. Instability coincides with coupling ranges below the l th harmonics where they maximize. However, the unstable ranges are quite broad as expected for relativistic effects. Moreover, note that $\Gamma \neq 0$ also below $l = 1$ (labeled 1,2) indicating coupling to the Z-mode, i.e., the slow branch of the RX-mode which cannot escape the plasma

tal. It is expected that in the relativistic or ultra-relativistic regimes the growth rates will overlap over a much broader range leading to a broad synchrotron-like emission spectrum. Possibly, due to propagation effects the fundamental could be reabsorbed in the plasma allowing only the higher harmonics to escape and causing a cut-off on the observed spectra. Note that $\Gamma \neq 0$ also below $l = 1$ (labeled 1,2) indicating coupling to the Z-mode, the slow branch of the RX-mode which cannot escape the plasma.

When adding a small cold background component, the growth rate is suppressed gradually with increasing cold fraction due to the increased wave absorption. This absorption is strongest at the fundamental and is of less importance at higher electron-cyclotron harmonics implying that the fundamental harmonic in the presence of a cold background is suppressed strongest and what remains are higher though weaker harmonics. Moreover, for increasing emission angles (not shown here) the maser-growth rate decreases steeply similar to the case of the loss-cone distribution. The reason for this decrease is that the resonance circle becomes an ellipse which does not cover anymore the positive derivative range on the full ring-shell distribution.

In order to obtain a simple analytical expression for the condition when the maser instability of the ring-shell maser will set on, we consider the dispersion relation (2) with the above undisturbed distribution (30) for $k_{\parallel} = k_{\perp} = 0$ and $\omega_{pe}/\omega_{ce} \ll 1$ and fundamental wave frequency close to the electron-cyclotron frequency, $\omega \approx \omega_{ce}$. Then the dielectric tensor components $\epsilon_{13} = \epsilon_{23} = 0$, and the RX-mode dispersion relation reduces to a cubic equation

$$\omega(\gamma_R \omega - \omega_{ce}^2)^2 = \frac{\omega_{pe}^2}{2\gamma} \left[(\gamma_R^2 + 1)\omega - 2\gamma_R \omega_{ce} \right] \quad (31)$$

which in the non-relativistic case $\gamma_R = (1 + p_R^2/m_e^2 c^2)^{1/2} \rightarrow 1$ just reproduces the Bernstein and RX-modes $\omega = \omega_{ce}$ and $\omega = \omega_x$. In the relativistic case for finite $\gamma_R - 1$, one has $\omega/\omega_{ce} = 1 + \delta$ and

$$\delta = \frac{\omega_{pe}^2}{2\omega_{ce}^2} - (\gamma_R - 1) \pm \frac{1}{2} \frac{\omega_{pe}^2}{\omega_{ce}^2} \left[1 - 4(\gamma_R - 1) \frac{\omega_{ce}^2}{\omega_{pe}^2} \right]^{1/2}. \quad (32)$$

This expression becomes complex and thus unstable for a resonant electron kinetic energy

$$\gamma_R - 1 > \frac{1}{4} \frac{\omega_{pe}^2}{\omega_{ce}^2} \quad (33)$$

exceeding a threshold which is determined by the small ratio of the plasma to cyclotron frequencies. At unstable frequency $\omega_r < \omega_{ce}$ the growth rate of the maser is $\gamma_R \rightarrow (p_R/\sqrt{2}c)\omega_{pe}$.

A simple extension of this approach is to a complete three-dimensional shell which has distribution function

$$f_0(p_\perp, p_\parallel) = \frac{1}{2\pi p_S} \delta(p_\perp^2 + p_\parallel^2 - p_S^2). \quad (34)$$

In the complete shell case the dispersion relation (32) is replaced by

$$\delta = \frac{\omega_{pe}^2}{2\omega_{ce}^2} - (\gamma_S - 1) \pm \frac{1}{2} \frac{\omega_{pe}^2}{\omega_{ce}^2} \left[1 - \frac{8}{3}(\gamma_S - 1) \frac{\omega_{ce}^2}{\omega_{pe}^2} \right]^{1/2}, \quad (35)$$

where $\gamma_S = (1 + p_S^2/m_e^2 c^2)^{1/2}$, and the condition for electron-cyclotron maser instability indicates that instability will still occur but now at higher shell momenta and shell-kinetic energy

$$\gamma_S - 1 > \frac{3}{8} \frac{\omega_{pe}^2}{\omega_{ce}^2} \quad (36)$$

also leading to smaller than ring-shell growth rates. This decrease is simply due to the increased homogeneity of the pure shell distribution in the third dimension. The observation of electron-cyclotron shell-maser emission from a remote object should thus provide an estimate of either the available free shell-energy or the plasma-to-cyclotron frequency ratio in the source region.

The simulations of the evolution of the cyclotron-maser emission by ring-shell and shell distributions are shown in Figs. 17 and 18. In such simulations the ring distribution has an unavoidable finite spread and for small or intermediate relativistic momenta is thus more realistic than the extremely cold shell

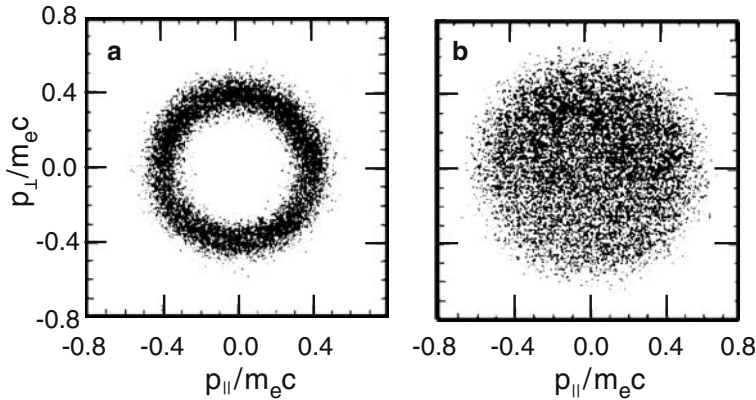


Fig. 17 The ring-shell distribution at times $\omega_{pet} = 0$ (left) and $\omega_{pet} = 30$ (right) used in the one-dimensional simulation of the ring-shell maser for $p_R = 0.4m_e c$ and $\omega_{ce} = 7.5\omega_{pe}$ (after Pritchett 1984b with permission by the American Geophysical Union). Since the radiation cannot escape from the simulation box the initial ring-shell is smeared out non-linearly with time leading to the final, completely filled velocity space inside the initial ring in the right part of the figure

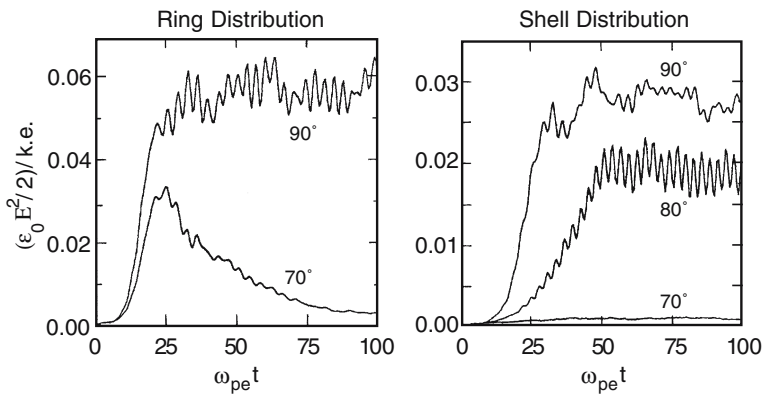


Fig. 18 *Top:* Time evolution of the total transverse electric energy in radiation with $p_S = 0.4m_e c$ and $\omega_{ce}/\omega_{pe} = 7.5$ at the two emission angles 90° and 70° for the ring-shell maser *Bottom* The same evolution for the shell maser at the three emission angles 90° , 80° , and 70° versus time ω_{pet} (after Pritchett 1984b with permission by the American Geophysical Union). The radiation energy has been normalized to the initial kinetic energy. After a short exponential growth phase of the emission the power in the ring maser saturates for perpendicular emission at about 6%. At the smaller angle the radiation is spiky and after having reached a temporary maximum at $\omega_{pet} = 25$, it decays away probably due to the strong non-linear effect imposed by the perpendicular component. The behavior of the energy for the shell maser is similar except that the total intensities reached are lower and the growth is retarded

distribution. These simulations show that ring-shell and complete shell distributions indeed lead to an exponential growth phase of the emitted radiation with the intensity of the emission and growth rate being larger for the ring-shell than for the complete shell distribution. The reason is that the complete shell distribution has a higher symmetry and thus is less unstable than the

ring-shell distribution. The main result in both simulation cases is that for sufficiently large perpendicular electron momenta in the ring or shell distribution maser instability sets on since the distribution provides something like an inverted population in energy space and thus fakes the quantum effect of correlated electrons.

In the simulations the radiation saturates at later time. This is a nonlinear effect, which we will briefly discuss here, as it suggests that in nature the maser would really saturate within a couple of electron plasma periods, for the parameters of the simulations shown here in a few tens of ω_{pe}^{-1} . The electrons loose energy to the radiation as can be seen from the positive growth rate. This energy loss will in principle diminish the transverse momentum space gradient in a purely classical way scattering electrons away out of the positive gradient to lower momenta. This effect though very weak causes the observed saturation. However, this saturation is an artifact of the simulation since the simulation does not allow the radiation to escape. This would be correct for any plasma mode which is trapped in the plasma and thus cannot leave. But radiation above the cut-off can practically freely leave the plasma being lost, and then the small energy loss of the particles is practically not appreciated if only the distribution rebuilds itself by external forces. Here, in the simulations this is inhibited. The radiation stays in the simulation box, being in resonance with the energetic electron component and has plenty of time to react on the distribution until the latter has become depleted. Then the final level of the radiation is reached where it stays until the simulation ends. For natural application to emitted radiation saturation plays a role only if it is confined in the total radiating volume. Otherwise it will immediately leave, encounter other plasmas where it can be absorbed or leave to free space. In this sense a quasi-linear treatment of the maser emission as one can find it in many of the published papers is of little relevance to natural application. These considerations hold for weakly relativistic ring-shell plasmas. If the plasma becomes highly relativistic, two effects will modify the conclusions. One is the beaming effect on the radiation which narrows the emission angle. The other is the gradual overlap of the various cyclotron harmonics which will lead to a broader more synchrotron emission-like but more intense than incoherent radiation spectrum while the low frequency cutoff of the emitted spectrum will remain at a frequency close to though below the electron–cyclotron frequency. Such conditions can be expected in many highly relativistic astrophysical objects like binaries, pulsar magnetospheres and AGNs. Closeness to ω_{ce} , however, depends on the available bulk γ_S . For large γ_S the cutoff could be substantially below ω_{ce} and if incorrectly interpreted could cause an underestimate of the magnetic field in the maser source region. Louarn et al. (1986) has given an attempt to account for such a relativistic effect. The main result is that even though the growth rate of the maser mechanism increases with γ_S , the relativistic overlap of the harmonics and the resulting spreading of the spectrum decreases the emissivity at the single harmonics since the energy is spread over a wide range in frequencies.

5.4 Ring-beam maser

For completeness we note that the ring-shell maser has been extended to include the propagation of the ring-shell distribution along the magnetic field (Freund et al. 1983) though so far only for the case of a dense background plasma with only a thin suprathermal propagating beam. This case is of practical interest for systems containing dense background plasmas traversed by relativistic beams. In this case the distribution becomes a ‘hollow beam’ distribution

$$f_{hb}(u_{\perp}, u_{\parallel}) = \frac{\exp(-\zeta_0^2)}{\pi \Delta u_{\perp}^2 \sqrt{\pi \Delta u_{\parallel}^2 [1 + \zeta_0 Z(\zeta_0)]}} \times \exp\left(-\frac{(u_{\parallel} - u_{\parallel 0})^2}{\Delta u_{\parallel}^2} - \frac{(u_{\perp} - u_{\perp 0})^2}{\Delta u_{\perp}^2}\right), \quad (37)$$

where as before $u_{\parallel} = p_{\parallel}/m_e$, $u_{\perp} = p_{\perp}/m_e$, Δu_{\parallel} , Δu_{\perp} are the thermal spreads of the distribution in parallel and perpendicular direction, respectively, $\zeta_0 = -iu_{\perp 0}/\Delta u_{\perp}$, $Z(\zeta_0)$ the plasma dispersion function, $u_{\parallel 0}$ is the beam velocity along the ambient magnetic field, and $u_{\perp 0}$ is the ‘population inversion’ of the velocity distribution in perpendicular direction. This distribution is to some extent a generalization of the ring-shell distribution to the inclusion of a beam. Such distributions might be of interest in astrophysical application like jets where fast plasma beams are the rule. If such beams, in addition, exhibit a population inversion caused for instance by a field-aligned electric potential drop (parallel electric field), the above distribution offers a model of such a ‘hollow beam’. Maser instability of such a hollow beam has been investigated in the presence of a cold background plasma by Freund et al. (1983). The most important result of the linear treatment of the hollow beam is summarized in Fig. 19. From this figure it is obvious that the RX-mode is excited strongest at the second harmonic $l = 2$, while the LO-mode has maximum excitation at the fundamental $l = 1$.

5.5 Inhomogeneous medium

So far we have dealt with only a homogeneous medium. The electron-cyclotron maser radiation source plasma as paradigmatically represented by the AKR source which is located inside the auroral cavity is known to be violently inhomogeneous on all scales. On the macroscopic scale which we have identified earlier to be of the order of 100 km in latitudinal extension and possibly a few 1,000 km in longitudinal extension it is bounded by dense plasma walls which inhibit propagation out of the source region. This inhomogeneity is even more restrictive for the ring-shell distribution since the electron-cyclotron maser acts most effectively in perpendicular direction. It is thus unclear whether or not the radiation once excited can ever leave the radiation source. Similar arguments will apply to any other system than the Earth’s or planetary auroral

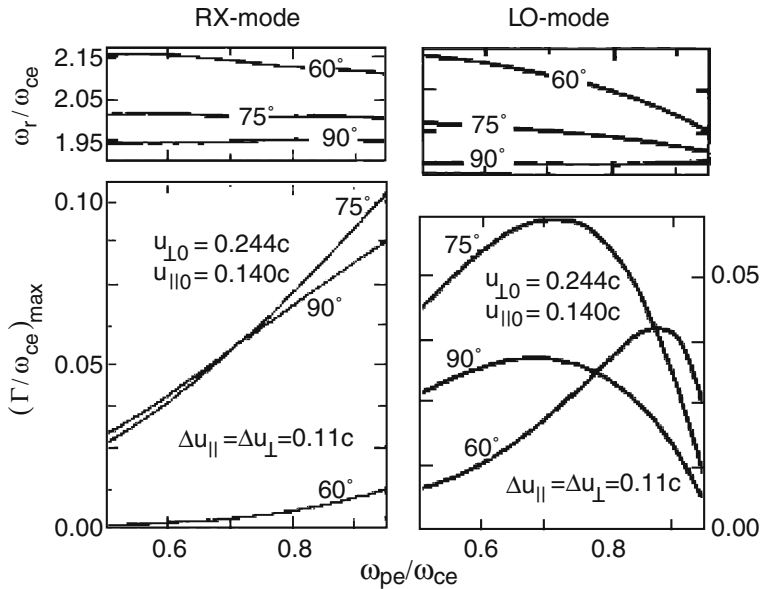


Fig. 19 The maximum growth rates for the RX-mode (*left*) and LO-mode (*right*) in hollow-beam maser emission (after Freund et al. 1983) vs. ω_{pe}/ω_{ce} and for different propagation angles. Note that the RX-mode is excited only at the second harmonic $l = 2$, while the LO-mode has maximum growth rate for the fundamental

regions depending on the relative geometries of the magnetic field and plasma density. It is then tempting to assume that only obliquely emitted radiation can ultimately escape. However, the efficiency of the maser drops steeply with increasing emission angle, and extraordinarily high radiation temperatures cannot be expected for oblique propagation.

Figure 20 shows two cases where radiation from an embedded localized electron–cyclotron maser source emitted about perpendicular might leave the cavity. On the left the RX rays encounter the RX mode cutoff which is most probably along the inclined magnetic field-lines in the mirror geometry. Here the rays become reflected thereby increasing their angle. This increase in the propagation angle will necessarily let them leak out of any further resonant amplification. Further gradual increase of the angle transforms them continuously into the completely circularly polarized R-mode and let them escape to free space parallel to the magnetic field. Other smaller scale inhomogeneities related to the presence of density fluctuations, electron and ion holes will scatter radiation as well. However, the most serious restriction is actually given by the cavity.

The mechanism on the right is different (Le Quéau et al. 1984b; Le Quéau 1988; Louarn and Le Quéau 1996a,b; Louarn 2006). In this case the RX mode cutoff is incomplete, and the RX-mode reaching the boundary of the plasma cavity can couple to the LO mode which propagates in the external dense plasma region and can freely escape to space. This is what happens in a non-ideal wave

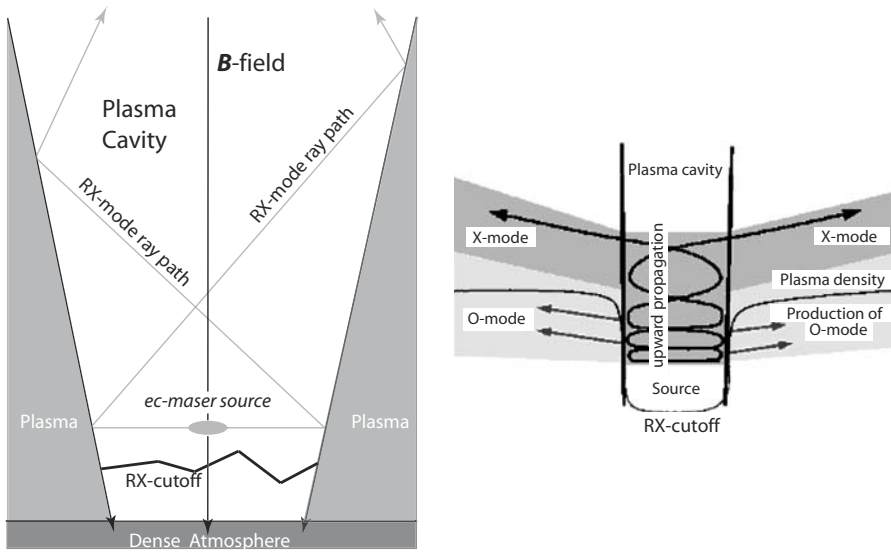


Fig. 20 Schematic of an auroral cavity which traps the electron-cyclotron maser radiation. *Left:* Complete trapping. The RX-mode cutoff does not allow the radiation to pass into the plasma, but the RX-mode rays emitted from the electron-cyclotron-maser source perpendicular to the local field are reflected at the cutoff in the inclined fields and can after several reflections escape to free space. *Right:* Incomplete trapping (after Louarn 2006). Radiation hitting the RX-mode cutoff at the boundaries of the plasma cavity is allowed to couple to the external LO-mode like in an incomplete wave-guide. Rays making it up to the region above cutoff can freely escape in the RX mode, as on the *left* of the figure

guide and may explain the occasional observation (Louarn 2006) of relatively intense LO mode radiation from the auroral plasma cavity. Since direct amplification of the LO mode in the plasma cavity in the electron-cyclotron maser process is by far less efficient than RX mode excitation such, a RX to LO wave transformation mechanism provides a reasonable explanation of the LO mode observation.

Pritchett 1986 (Pritchett and Strangeway 2002) performed full particle simulations of the cyclotron maser radiation in localized sources in an inhomogeneous plasma cavity. In the first of these papers he allowed for the presence of several narrow irregularly distributed density irregularities elongated along the magnetic field in order to investigate their effect on the propagation of the RX mode and to check the waveguide feedback model of amplification of Calvert 1981a,b, 1982, 1987, 1995. These simulation led to the conclusion that the feedback mechanism did not work efficiently. Instead, narrow inhomogeneities were easily crossed by the $\omega < \omega_{ce}$ perpendicular radiation via coupling to the slow RZ mode inside the inhomogeneity and recoupling to the RX mode in the cavity at exit from the inhomogeneity with very little attenuation. Interestingly enough, it was found that the radiations occur in intense bursts which form wave packets traveling at group speeds $(0.1\text{--}0.2)c$, typical for being close to RX mode cutoff. Reflection at the cavity edges turned out to be quite weak,

with only 0.1% of the incident energy being reflected which supports the model of a bad waveguide and coupling to the RZ mode instead of the LO mode. This result to some extent also contradicts the analytical calculations of Louarn and Le Quéau (1996a,b) and Louarn (2006), which favor coupling to or rather excitation of the LO mode at the cavity boundary. Such a contradiction is intelligible as the RX and RZ mode are of same polarization while the LO mode has opposite polarization. One thus intuitively expects that a wave transformation at the boundary should be easier between waves of equal polarization than between waves of opposite sense of polarization. Excitation in the latter case should be much weaker and thus less efficient and in addition should not proceed via transformation from one wave into the other. Rather it should proceed via absorption of one wave of one sense of polarization and re-emission of the wave with the other polarization. The simulations suggest that this process might be possible but is favored less than the direct transformation.

Using a complete shell distribution as suggested by FAST observations, two-dimensional simulations (Pritchett et al. 2002) have basically confirmed these conclusions. In this case the magnetic field was allowed to have a mirror geometry. Again generation of bursts of radiation of 0.5 ms length were observed, and the coupling of the perpendicular RX mode to the Z mode at the boundary was reproduced though in the two-dimensional case stronger attenuation was found than in one dimension, and that the part of the wave propagating in the R mode parallel to the magnetic field was stronger amplified. However, the most important observation in view of trapping of the radiation and amplification was that no standing wave structure was ever observed and thus waveguide amplification is irrelevant for reaching the high emissivities in the shell electron–cyclotron maser. Inhomogeneity determines the propagation properties, attenuation, and coupling to the Z-mode of the originally excited $\omega < \omega_{ce}$ radiation in the RX mode, and the high intensities reached are entirely due to the direct linear excitation properties of the shell maser mechanism with no further macroscopic waveguide amplification necessary.

In addition, it is very interesting to find excessively bursty emission in all these simulations which could not be explained in a simple way. This bursty radiation points on something which is hidden in the radiation mechanism itself and which we believe is related to the generation of microscopic structuring of the plasma in the presence of strong field-aligned electric fields. In the following we therefore review the reasons for the occurrence of magnetic field-aligned potentials and their effect on the structure of the plasma in the plasma cavity region.

6 Electric double layers

The key to the maser mechanism is found in the presence of field-aligned electric potential drops which act to accelerate electrons in one and ions into the opposite direction along the magnetic field. The possibility of such localized potential drops causing parallel electric fields in collisionless plasmas has been

proposed half a century ago (Block 1972, 1977) and has been provisionally though not definitely been confirmed in laboratory (Saeki et al. 1979; Leung et al. 1980) and space plasmas (Temerin et al. 1982; Boström et al. 1998; Ergun et al. 2001a; Pottellette et al. 2003).

Theoretically, the maintenance of electric double layers along the magnetic field in collisionless plasmas encounters severe and still not completely resolved difficulties. Clearly, double layers are related to the presence of sufficiently strong electric currents along the magnetic field. When these currents exceed a critical value they undergo instability and generate electrostatic waves. For strong currents the localized electric fields of these waves cannot be compensated fast enough, and localized potentials occur which survive for times long enough to affect the dynamics of the plasma.

Simple global relations have been derived by several authors between the parallel current strength and the potential drops which can be achieved under different assumptions like Maxwellian and non-Maxwellian distribution functions. These relations are all oversimplified as the real electron-distribution functions deviate strongly from those models in the presence of electric wave fields. They actually develop into incomplete ring-shell ('horseshoe') distributions, the presence of which we have identified as the necessary condition for the electron-cyclotron maser.

6.1 Generation of incomplete ring-shell ('horseshoe') distributions

Chiu and Schulz (1978) investigated the adiabatic motion of electrons along a mirror magnetic field geometry like the one realized in the vicinity of magnetized planets, in many places in the solar atmosphere, the atmospheres of active magnetized stars in general and in other places. They found that in the presence of a field-aligned electric field component the concerted action of the electric acceleration of the electrons and the mirror effect of the magnetic field, which tends to increase the pitch angle of the electrons, should produce a deformation of the initial electron-distribution function.

In addition to the prevalent loss cone in the distribution function, low-energy electrons were excluded from the distribution, and the more energetic beam electrons were found to be diverted to larger pitch angles. Both these effects leading to the formation of an incomplete ring or to ring-shell distributions of the kind discussed in previous sections and now known to be responsible for the electron-cyclotron maser emission in the Earth's auroral region — and probably at many other places as well. Without a parallel electric field the distribution function would simply remain a loss-cone distribution. Thus it is the electric field which — in analogy to molecular masers — excites the particles and lifts them into a non-thermal higher energy level. The particles draw their excess energy from the presence of the electric field. Schematically this is shown in Fig. 21. This requires the generation of electric fields which are directed parallel to the magnetic field and must be directed away from the object (Earth) so that they accelerate electrons towards the object and empty the plasma cavity

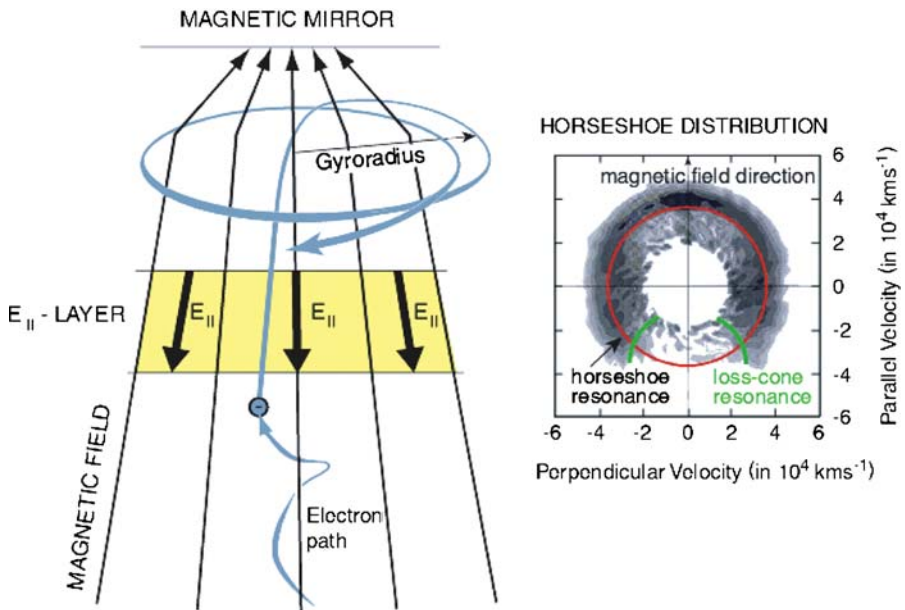


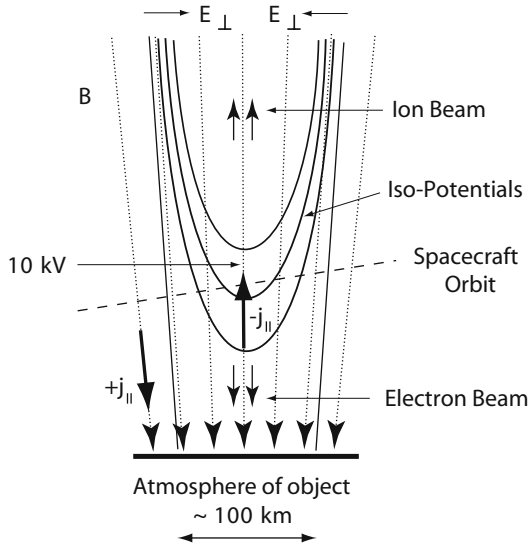
Fig. 21 *Left:* Schematic of the formation of a ‘horseshoe’: a parallel electric field accelerates the electrons towards the mirror, while the mirror effect of the converging magnetic field-lines shuffles the excess energy which the electrons gained in the electric field from parallel into perpendicular energy. The magnetic moment of the electrons is conserved thereby and the pitch-angle of the electrons is increased. This results in the horseshoe-like shape of the electron distribution with respect to pitch-angle and the two velocity components v_{\parallel} and v_{\perp} . On the right part of the figure we have indicated two relativistic resonance lines. *Red* is the horseshoe-resonance circle of the RX-mode for perpendicular radiation, *green* is the loss-cone resonance line. The latter one is necessarily for oblique propagation since it is an ellipse that is shifted out of the origin. *Right:* An incomplete ring-shell (‘horseshoe’) phase-space distribution occurs, given the mirror effect in a converging magnetic field and a parallel electric field. The ‘horseshoe’ is an incomplete ring-shell since the particles inside the loss-cone are missing

by evaporating its low-energy plasma component along the magnetic field into space.

6.2 Generation of double layers

Figure 22 shows the schematic of a double layer in a mirror magnetic field configuration. In ideally conducting collisionless stationary plasma the magnetic field-lines become electric iso-potentials with electric field having only components perpendicular to the magnetic field. The plasma under such conditions performs a current-free drift motion $\mathbf{v}_e = \mathbf{E}_{\perp} \times \mathbf{B}_0 / B_0^2$ perpendicular to both the magnetic and electric fields. However, when by some process two regions of oppositely directed sheared drift motions come into contact, the electric fields either converge or diverge. In this case the transition region between the shear

Fig. 22 Schematic of the conditions inside an electric double layer with the electric potentials deviating from their original direction parallel to the magnetic field-lines. In this case the electric field is directed upward accelerating electrons downward and ions upward. An ion beam is generated in this way while the plasma below the double layer is evaporated. The energetic electrons become deformed into a ring-shell distribution as has been described earlier. The iso-potentials shown are assumed to be 3 kV apart, adding up to a total potential drop of 10 kV



motions acts like a charged layer. Of interest here is only a negative charge layer corresponding to converging electric fields.

Such a layer requires that the electric iso-potentials start deviating from being parallel to the magnetic field-lines as shown in Fig. 22. Numerical simulations of Singh et al. (2006) of such shear flows (Fig. 23) have demonstrated that beneath the shear flows a region evolves where negative potentials accumulate, the initially straight stretched iso-potentials turn into U-shaped and S-shaped potentials, and a field-aligned electric field (which is upward directed under the conditions at Earth) evolves in a limited spatial domain, and thus evaporates all the plasma in the region below the shear flows.

This is nicely shown in the Fig. 23 as a dark U-shaped hole in the plasma density forming at 1,000 plasma times with the dark equipotential line becoming S-shaped. In this way the plasma cavity is generated, and at the same time a region of parallel electric fields appears which lasts as long as the shear flows are maintained atop the layer. Internal non-linearities of this structure let it become unstable for longer simulation times with the bottom of the cavity becoming erased, developing filamentary density structures and moving upward as an ion-acoustic disturbance in order to erase the shear flows. One thus concludes that the process of double-layer formation is non-stationary: it lasts for limited times only and causes pulsed parallel electric fields to appear and to disappear. Since, however, $\omega_{pe} \ll \omega_{ce}$ these times are much longer than those required for the generation of the ring-shell or horseshoe distributions and emitting electron-cyclotron maser radiation.

A rare direct in situ measurement by a FAST-spacecraft crossing an electric double-layer is shown in the lower part of Fig. 24 where the two components parallel and perpendicular to the magnetic field have been plotted as a function of time. The double layer appears as a short unipolar excursion of the paral-

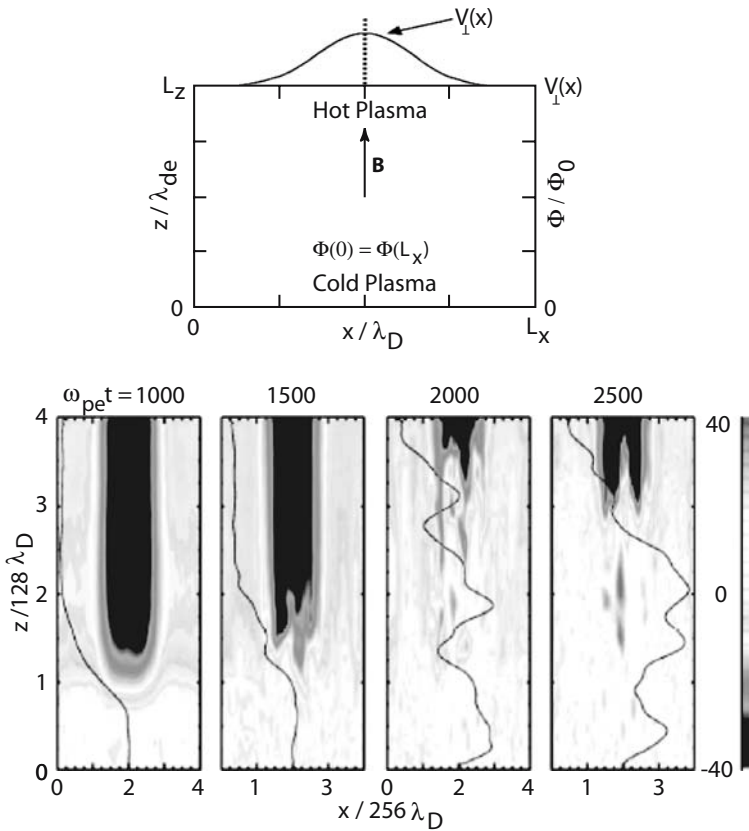


Fig. 23 *Top:* A shear-flow model, as it is used to simulate the evolution of a double layer in a magnetic-mirror geometry. Shown is the simulation box with the smooth shear flow profile on top. The simulation is periodic in x , which requires that the electric potentials at $x = 0$ and $x = L_x$ are the same. *Bottom:* Sequence of the two-dimensional evolution of the plasma cavity (gray-scale coded plasma density in *box*) and of one individual vertical double-layer potential-line at four different times (measured in inverse plasma times ω_{pe}^{-1}). The distances in the vertical (i.e., parallel to the magnetic field) and horizontal (i.e., in the perpendicular direction) are measured in multiples of Debye lengths λ_D . The deviation of the density from its initial value is given on the *gray scale bar* in arbitrary units (after Singh et al. 2006, with permission by the European Geophysical Society). A deep plasma cavity evolves at time $t\omega_{pe}^{-1} = 10^3$. After another 10^3 time steps it has moved up the field lines and has developed a distinct substructure. The smooth shape of the originally S-shaped iso-potential line is distorted, and short-scale alternating parallel electric fields have evolved below the cavity along the magnetic field. Such fields correspond to local charges and thus represent phase-space holes

lel electric field. Interestingly, the generation region of the electron–cyclotron maser radiation in this crossing, which is located on the left-hand side, corresponds to the *low potential side* of the layer in agreement with the presence of a negative space charge (electron beam) in the center of the cavity and converging perpendicular electric fields as expected in a shear flow layer. The behavior of the perpendicular electric field also shows that the double layer is

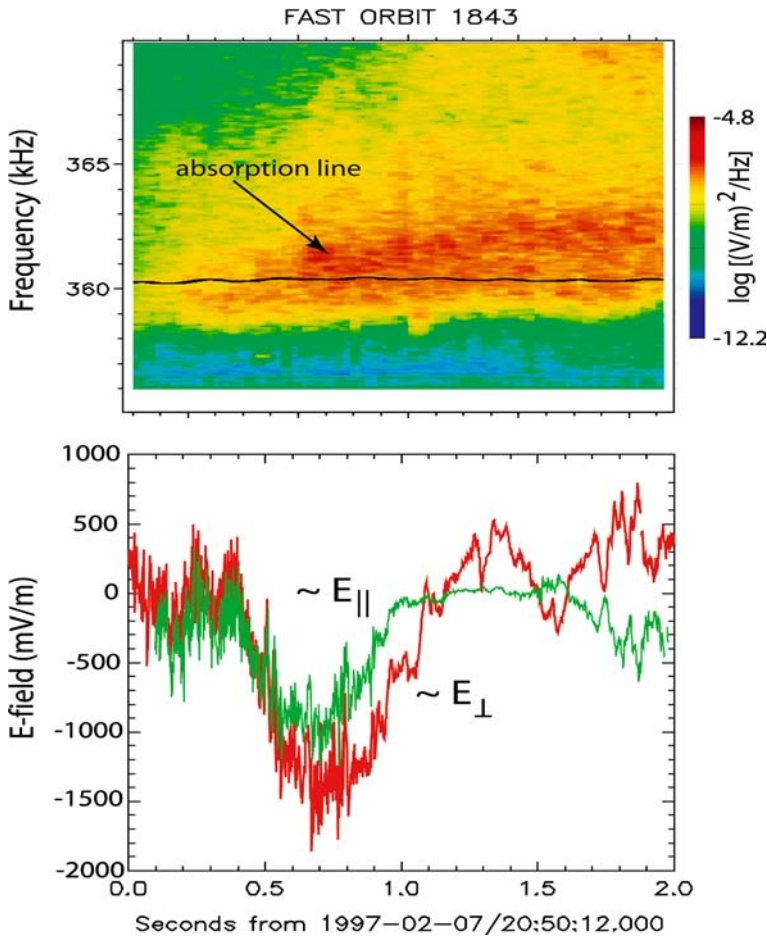


Fig. 24 *Top:* The dynamic spectrum in the source region of auroral kilometric radiation obtained with very high time resolution. The emission is partially local, namely at frequencies below the local electron-cyclotron frequency (*black line*). The very narrow-band fine structure suggests that it consists of many small-scale elementary radiators which, however, here in the source region are less well separated than in Fig. 7 where the spacecraft is at larger distance from the radiation source. Therefore, their emissions overlap and produce a nearly continuous spectrum. Nevertheless, one can distinguish some fast drifting vertical structures, narrow-band emissions and some narrow-band absorptions. These indicate radiation from electron holes as explained below (see Fig. 30). *Bottom:* DC electric field measurement during the same time interval in parallel (*green*) and perpendicular (*red*) directions. The parallel field exhibits a unipolar anti-Earthward signature typical for a double-layer potential ramp (after Pottelette and Treumann 2005 with permission by the American Geophysical Union)

crossed at a position above its dip (cf. the detached line indicating the spacecraft orbit in Fig. 22). At dip position the perpendicular electric field should vanish whereas here it does not. Possibly, the double layer is a three-dimensional structure with the two-dimensional assumption which is being a crude approximation only. Observations of this kind identify electric double-layers

as a reality in collisionless high-temperature plasmas when imposed on strong converging magnetic fields in mirror geometries.

We should also note that the position of the double layer is the site of direct electron and ion acceleration. In fact, Fig. 3 shows that, at the boundary crossings of the plasma cavity, ions become steeply accelerated into a fast cold upward beam. Closer inspection of the high time- and energy-resolution data of the electron distribution (Pottelette et al. 2004) reveals a very similar effect on the electrons. The electrons become effectively accelerated when entering the double layer just by the amount of the measured potential difference which adds to their already quite high initial energy which they received when leaving as a warm weakly relativistic beam from the distant source. In the case of Earth's auroral plasma cavity this is the magnetic reconnection site which is found a couple of Earth radii away from the Earth in the tail of the magnetosphere.

Figure 24 in its upper part gives the highest time and frequency resolution that is currently available in the radiation spectrum during a double-layer crossing. We should remember that, while crossing the boundary, the spacecraft is not yet deep enough inside the plasma cavity, i.e., such that the optimum conditions for the electron–cyclotron maser are not yet reached. The plasma here is still relatively dense and the maximum emission efficiency should therefore not be attained yet. Moreover, it is expected that the emissivity will not be highest below the local electron–cyclotron frequency in this case and that the emission will be slightly oblique. This is reflected in the spectrum shown. Highest intensities are reached just at or just above the local cyclotron frequency. Nevertheless, strong emission is already observed also from beneath ω_{ce} (which is the black line in the figure). Finally, the outstanding phenomenon revealed by this figure is the enormously fine structuring of the emission in frequency. This suggests that almost all the radiation in the source region is generated by very small ‘elementary radiation sources’.

6.3 Phase space holes

These in situ observations raise the interesting but tantalizing question of the nature of the ‘elementary radiators’ identified in the radiation spectra of Figs. 7 and 24. This question is not an academic one: only when one is able to answer it, can one expect to understand the nature of the radiation and, on the other hand, extract information from the radiation received from a more remote object. We note that this becomes possible even though one will probably never be in a position to resolve any distant source as well as is the case for the near-Earth radiation sources of the AKR. In order to proceed, we consider two questions:

- What is the physical nature of the elementary radiators, i.e. what is the physical nature of the elementary sources of the electron–cyclotron maser emission?
- What is the mechanism that turns a small-scale entity like a phase-space hole into an ‘elementary radiation source’, i.e. how does the electron–cyclotron maser mechanism work within a small-scale source?

In the present subsection we attempt to deal with the first question only remembering that the interior of the cavity (and therefore the macroscopic electron–cyclotron radiation source region) contains a strong field-aligned current that is passing through a very dilute plasma. Such currents excite kinetic plasma instabilities of the family of the two-stream instability, which is excited in the interaction of the counter-streaming electron and ion beams. When the current drift speed $\mathbf{v} = \mathbf{v}_{\parallel i} - \mathbf{v}_{\parallel e}$ (which is the difference of the average ion and electron drift velocities) exceeds the electron thermal velocity $v_{e,th}$, this instability growth very fast. It readily traps electrons in the wave-electric potential and thus generates localized structures on the electron-distribution function, which are known as ‘electron phase space holes’. These holes correspond to a splitting of the electron-distribution function into trapped (consisting of electrons with energies less than the potential of the wave on the wave frame of reference) and untrapped (consisting of electrons with energies larger than the wave potential) distributions (Schamel 1979, 1986). At lower current drift velocities, the instability is on the ion-acoustic branch and may excite similar localized structures on the ion distribution function. Such entities are known as BGK (Bernstein–Green–Kruskal) modes and contain localized electric fields that can be detected with high-resolution instrumentation.

Figure 25 gives a recent example of the wave form and spectrum of a sequence of such localized structures as detected by the FAST spacecraft in situ in the auroral plasma cavity Pottelette and Treumann (2005). The representation is only 200 ms long showing seven electric field signatures of localized structures. It takes the spacecraft less than < 10 ms to pass through a structure of typical

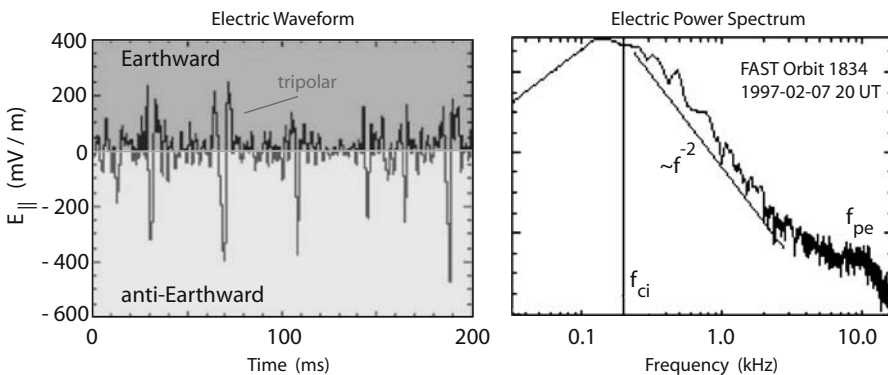


Fig. 25 *Left:* A 200-ms excerpt of parallel electric VLF wave form that shows a chain of seven nearly equally spaced tripolar electric-field structures that are embedded into a fluctuating background of electric waves. One distinguishes the large-amplitude anti-Earthward field signatures which typically last for a few milliseconds. They are flanked by two weaker Earthward field excursions. Such anti-Earthward fields accelerate electrons Earthward along the magnetic field. *Right:* The VLF spectrum for the same time period maximizes below the ion cyclotron frequency and exhibits the power law shape towards higher frequencies that is typical for broadband noise. A second weak maximum occurs at the plasma frequency, indicating the presence of electron-plasma waves that are excited by the accelerated electron beam in the ‘horseshoe’ distribution (after Pottelette and Treumann 2005 with permission by the American Geophysical Union)

length, and this probably corresponds to only a few electron Debye lengths. The electric amplitudes reach values of 0.5 Vm^{-1} inside a structure, a very strong electric field, indeed. Such fields can be maintained in a collisionless plasma only if particles with an energy less than a few 100 eV are trapped inside the localized electric wave fields. The electric power spectrum of the phase-space hole is broad and featureless as is typical for a localized structure of this kind. Such a spectrum is shown in the figure.

In Fig. 26 we present a simple model of an electron phase-space hole which has been produced by Muschietti et al. (1999, 2002). This model is a BGK structure based on the theory of Schamel (1986). From left to right the figure shows the electric potential, the electric field, and the phase-space density. The dip on the ambient distribution function is obvious from this figure. The exactly symmetric potential and electric fields are however, an artifact of the numerical model and need not actually occur in nature. Rather a finite potential difference across the hole may be retained there. In the presence of many holes, these potential differences add up to a large-scale potential drop that can make

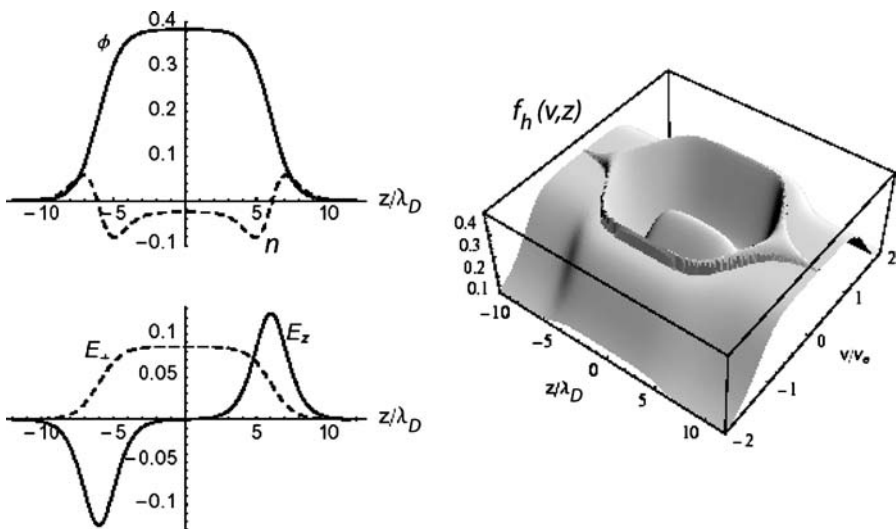


Fig. 26 A numerical model of a Bernstein-Green-Kruskal (BGK) electron-hole in a collisionless plasma (after Muschietti et al. 2002). *Left:* Potential ϕ , normalized density n_e (in arbitrary units), and electric field components $E_{||}$, E_{\perp} at the hole position. The density depletion in the interior of the hole is well recognized. Lengths are given in units of the Debye-length λ_D , potentials are measured in units of $k_B T_e/e$; the unit of the electric field is $k_B T_e/e\lambda_D$. The *solid curves* represent values along the ambient magnetic field, the *dashed curves* hold for the direction perpendicular to the magnetic field. The parallel and perpendicular cuts through, respectively, the potential and the perpendicular field have the form of hats, while the parallel electric field has a clear bipolar structure, as is measured in many cases. The perpendicular electric field is unipolar instead. The symmetry of the hole is intrinsic to the model and will not necessarily be reproduced in nature when a residual potential difference and, hence, an electric field in the parallel direction may be measured across the hole. *Right:* Pseudo-three-dimensional representation of the electron-distribution function showing the hole as a dip in the total (external) distribution with a small number of electrons trapped and sharp boundaries of externally attached electrons

up for the entire potential drop along the magnetic field in the plasma cavity. For instance, in Fig. 25 electron-holes occur at a rate of one hole every 20 ms yielding 500 holes per 10 s. The small potential difference of 10 mV m^{-1} in the sequence shown then corresponds to a 5 Vm^{-1} electric field along the field-line. Stretched over a length of $\sim 1 \text{ km}$ effective length (depending on the velocity of the electric structures), this yields a $\sim 5 \text{ kV}$ potential difference or, applied to particles, an electron acceleration up to 5 keV which is in pretty good agreement with observation.

On the one hand, these phase-space holes result from the existing parallel-potential drop, on the other hand they contribute themselves to the existing parallel electric field. This ambivalence has not yet been resolved.

The physics of formation of phase-space holes is highly non-linear and thus is not accessible to a perturbational approach. The only way of investigating their evolution is by numerical simulation. Such simulations have been performed by several authors (Newman et al. 2001, 2004; Goldman et al. 1999, 2003). They have either been based on the assumption of approximate equations or performing numerical particle-in-cell simulations, in one and more dimensions and with large numbers of particles involved. Recently such simulations have produced important results as given in Figs. 27–28.

Figure 27 shows the space-time evolution in one dimension of the energy of the parallel electric-field (normalized to thermal energy) in a plasma carrying an electric double-layer. The double-layer is modeled by a dip in the plasma density (indicated here as quasi-stable ramp). As suggested by the discussion in the former sections, a field-aligned electric potential-drop (parallel electric field) evaporates the plasma from the spatially limited region (interval 300–330 on x -axis in the figure) where the electric field is present. The double-layer in these simulations is the dark-shaded region near $x = 300 \lambda_D$. It is artificially kept alive by regenerating it permanently. (In nature it may be regenerated by the external maintenance of the shear flow.) The important results to be read from this simulation are the following:

- Immediately after initiating the simulation, the double-layer starts releasing narrow electron-holes of small amplitude which move at high velocity to the right away from the ramp. Their velocity is comparable to the electron-thermal velocity;
- These electron holes are of bipolar electric-field signature like the ones modeled in Fig. 26;
- After roughly 1,000 plasma times, the double-layer releases ion-holes which move comparably slowly to the left. These ion-holes are dips on the ion phase-space distribution and represent local ion-density decreases in real space;
- Being of larger spatial extension, these holes are themselves ramps on the electron scale. They reflect electrons and generate secondary electron-holes;
- The electron holes on the left try to escape to the right, but remain captured between the ion holes and the double-layer ramp for a relatively long time;

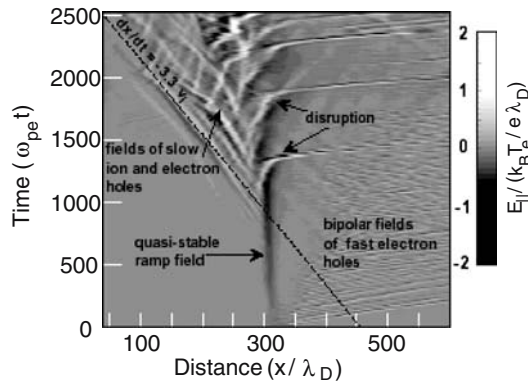


Fig. 27 One-dimensional numerical full particle simulation of the evolution of a localized double layer in a high-temperature plasma (after Goldman et al. 2003 with permission by the European Geophysical Society). The shading, as given on the right, signifies the strength of the parallel electric field $E_{||}$, which is normalized to $k_B T_e / e \lambda_D$. The double layer (*shadow*) is represented as a local, steep density-decrease causing an electric potential ramp in space. It continuously releases small-amplitude electron holes which escape at high velocity to the right. At a later time large-amplitude ion holes form which move slowly to the left thereby attracting electron holes which grow until they are large enough to break through the ramp and escape at high speed to the right. During breaking they briefly erode the ramp

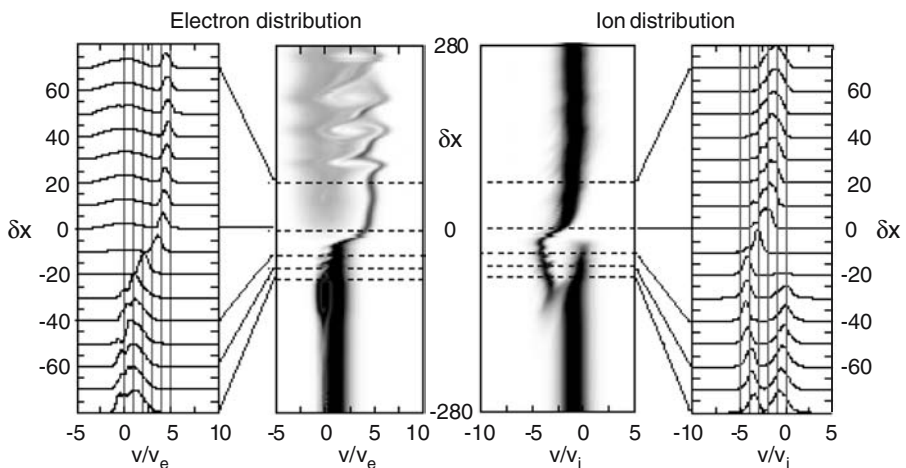


Fig. 28 Phase-space history of the (one-dimensional) electron (*center left*) and ion (*center right*) distribution functions during the evolution of a double-layer with phase-space hole formation (after Goldman et al. 2003 with permission by the European Geophysical Society). The vertical axis is the space coordinate along the magnetic field, and the horizontal axis is the velocity along the magnetic field. The double-layer ramp is in the center near $\delta x = 0$. A chain of electron-holes evolves on the electron distribution function for positive $\delta x > 0$, while for $\delta x < 0$ one single ion-hole develops on the ion distribution function. This ion-hole generates a reverse ion beam. In the velocity the electron-holes are characterized by a cold-electron beam, a hole gap, and by a broad hot-electron distribution. The figures at far left and far right show the small central sections of the electron and ion phase spaces, as indicated, in high resolution

- During this time, while the electron holes move slowly to the left together with the ion-holes, they grow and reach large amplitudes;
- The electron-holes which are trapped between the ion-hole and the ramp, interfere with other newly created electron- and ion-holes;
- When having reached large enough amplitudes, the electron holes attain enough momentum to break through the initial double-layer ramp to the right from where they escape at high speeds;
- In this process the double layer ramp becomes periodically eroded but rebuilds permanently;
- After long simulation times $> 3,000$ plasma times the production of ion-holes slows down, and the system settles into a quasi-steady state (not shown here) when strong ion-electron hole pairs are generated on the left and the break-through occurs on a longer timescale.

The sequence of electron- and ion-hole generation shown above indicates the dynamical coupling between electron- and ion-holes in the presence of a strong double-layer. This coupling is important for the application to the electron-cyclotron maser mechanism. It explains a number of observations which so far have been left open, even though the simulations have been performed on a Maxwellian background plasma. They neglect the formation of the ring-shell or horseshoe distribution which in the presence of a magnetic field-aligned electric field is unavoidable. Therefore the simulations do not contain any signature of the radiation. Inclusion of a ring-shell distribution would considerably change the dynamics of the holes in phase space and also in real space, as we will discuss below. It also leads to energy loss and thus attenuation of the electron-holes which on the basis of this simulation we can identify as the elementary radiation sources.

The large number of small-amplitude electron-holes escaping at high speed from the double layer to the right forms the elementary background-radiation sources (Fig. 7) of the emitted radiation. On the other hand, the smaller number of large-amplitude holes that are trapped on the left side of the ramp between the ion-holes and the double-layer ramp and move along with the double layer produce the intense narrow-band signatures of the fine-structure of the electron-cyclotron maser emission referred to in Figs. 6 and 7. One should note, of course, that the ion-holes themselves do not contribute to radiation. The generation of radiation can only be done by the electrons. However, the simulations show that the occurrence of the intense fine-structure emission-bands can be explained only when ion-holes and secondary electron-holes are produced at the double-layer ramp and the electron holes are trapped between the ion-holes and the ramp. This is in fact possible since ion-holes are the result of a parallel current instability of the kind of the ion-acoustic instability. Ion-holes are nothing else but localized depletions of the ion-plasma density. They represent negative charges which are localized on the plasma background and reflect electrons of less energy than the ion-hole potential. These reflected electrons form cold electron-beamlets that are attached to the ion-hole and are themselves

capable of producing the electron-holes which, in the simulation, occur on the left of the potential ramp.

Figure 28 shows the evolution of the electron and ion distribution functions in the presence of the double-layer ramp at different locations in phase space. The large ion-hole is visible in the ion distribution function on the right, where the ion-hole appears as a kink with missing low-energy ions in the ion distribution. At the location of the ion-hole the ion distribution splits into the two parts: the ordinary ion distribution and a return-ion beam, which both are needed in the formation of the ion-hole.

Instead, the electron distribution on the left of Fig. 28 shows the formation of many electron holes which form a long and wavy chain on the electron distribution. These holes cause the electron distribution to split into a thin cold beam and a broad hot distribution function of considerably larger width than the original distribution. The position of the electron holes in velocity space is identified as the lack of electrons between the narrow beam and the hot broad distribution. It is very interesting to note that the formation of electron-holes is accompanied by two effects on the electron distribution: these are the strong heating of the electron plasma and at the same time the cooling of the initial beam.

6.4 Radiation from electron holes

In regard to the second of the above question, one asks as to the presence of an electron hole can contribute to radiation. This is a delicate question, since as electron holes, even those trapped at the ramp of the double-layer, constitute small-scale structures with a size of only a few electron Debye-lengths. Nevertheless, radiation is generated in the presence of these small-scale entities because of two reasons. First, as we will show below, each of the holes produces a very steep phase-space gradient at its boundary, much steeper than any other gradient in the distribution function. Such gradients are required for intense maser-action and generation of the high observed radiation temperatures found in the auroral kilometric radiation (and the radiation from other magnetized planets like Jupiter and Saturn). Second, under the conditions described above electron-holes appear in very large numbers and thus their individual contributions to the radiation add up to produce high radiation intensities.

In order to contribute to the generation of radiation via the electron-cyclotron maser mechanism, an electron-hole must be able to modify the original distribution function by imposing a steep perpendicular velocity-gradient on the electron-distribution function. How this can be achieved is schematically demonstrated in Fig. 29. On the right in this figure an electron hole is imposed on the horseshoe distribution which is known to be present in the auroral cavity-radiation region of the electron-cyclotron maser. The hole is essentially an entity in the direction of the parallel velocity. The hole forms a narrow strip in velocity space that initially, for all perpendicular velocities v_{\perp} , is located at initial velocity $v_{\parallel,h}$. Since it lacks electrons, this hole is a positive charge on

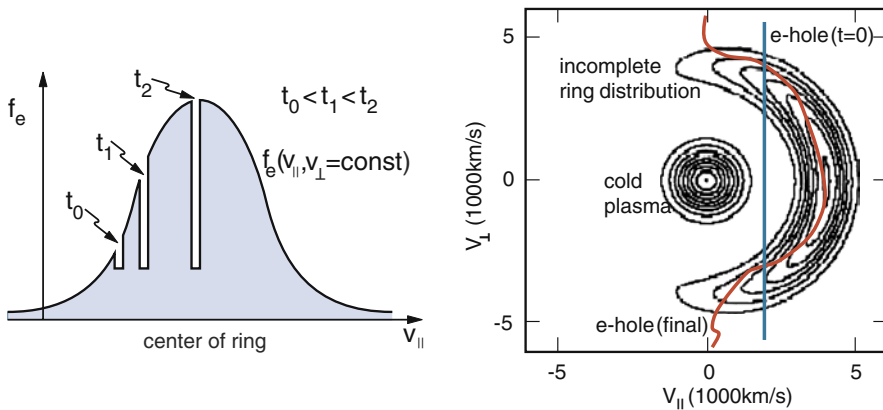


Fig. 29 Schematic evolution of an electron-hole when interacting with an incomplete ring-shell distribution. *Left:* Growth of the electron-hole due to parallel momentum exchange with the electron distribution. The hole, being a positive charge on the electron background, is attracted by the main distribution. When it moves into the bulk of the distribution, its relative depth increases. This corresponds to growth of the hole amplitude. *Right:* In phase space the hole can ‘move’ only in parallel velocity v_{\parallel} at fixed v_{\perp} . Moving (at each fixed v_{\perp}) into the bulk of the distribution, the initial hole at $t = 0$ (shown in blue) is deformed. Thereby the hole becomes nearly circular (shown in red). The final state is reached when momentum balance is achieved. The deformation of the hole in momentum space creates a gradient in perpendicular velocity along the hole. The bulk velocity of the hole along the ambient magnetic field varies during this process. The bulk velocity is obtained by performing the integral $2\pi \int_{\text{hole}} v_{\perp} dv_{\perp} dv_{\parallel} f_{\text{hole}}(v_{\parallel}, v_{\perp})$ along the deformed shape of the hole in phase-space. Here f_{hole} is the velocity distribution of the hole

the electron background and therefore experiences an attraction by the bulk of the negative electron-distribution function. This attraction acts only in the direction parallel to the magnetic field.

Consider a small part of the electron-hole located at constant perpendicular velocity. The attractive force that is exerted on the electron-hole by the bulk of the horseshoe-distribution results in momentum exchange between the hole and the horseshoe distribution and causes this particular part of the hole to move into the direction of the maximum of the bulk electron-distribution. This continues until this part of the electron-hole settles at a parallel velocity v_{\parallel} , where momentum balance is achieved. Note that for a completely symmetric ring the attraction to the left and right would partially cancel and the hole would move slowly to that part of the ring-shell distribution that is closer to the electron-hole. An electron-hole that is initially situated at $v_{\parallel,h} = 0$ experiences symmetric drags to the left and to the right and will thus remain in position. However, for a non-symmetric distribution like the horseshoe-distribution, the different parts of the hole at different $v_{\perp} = \text{const}$ will move different distances in parallel-velocity direction, and the hole will become deformed and will assume a bent structure. This is shown schematically on the right in the figure. There the hole has become deformed into a nearly circular section located close to the maximum of the horseshoe-distribution function.

The left-hand side of Fig. 29 shows what happens to the hole when it moves into the bulk of the distribution. As long as it can resist inflow of electrons from the outside into the hole, i.e., as long as there are no particles of low energy such that they become trapped inside the electron-hole, the number of particles in the interior of the hole remains constant. The local environment of the hole in velocity space attains higher and higher density as the hole moves into the bulk of the distribution, and the hole effectively grows to large amplitude. Since this potential increase when the hole moves into the bulk distribution, electrons are effectively expelled from the hole. This is quite realistic as the whole traps only those particles which have smaller energy than the local potential in the hole frame. Thus, the number of trapped electrons inside the electron-hole remains constant as long as the walls of the hole do not become erased by instabilities. Ultimately, this will of course happen and will destroy the hole. The hole will become oblate in velocity space when this occurs. However, before it happens, the velocity gradients on the distribution function at the boundaries of the electron-hole increase steeply in both directions — parallel and perpendicular to the magnetic field — when the hole deforms and enters the horseshoe distribution.

In this way deformed electron-holes produce steep perpendicular gradients $\partial f / \partial v_{\perp} \perp$ in the velocity distribution at the hole boundaries. This effect turns the distorted electron-hole into an efficient emitter of radiation, an ‘elementary radiation source’. Emission occurs at the upward edge of the electron-hole where $\partial f / \partial v_{\perp} > 0$, while at its downward edge, where $\partial f / \partial v_{\perp} < 0$ and which is at smaller v_{\perp} , the hole absorbs radiation. Radiation and absorption are due to the different signs of the perpendicular velocity gradients at the two boundaries of the hole, as shown in Fig. 30. Fortunately, due to the different positions of the

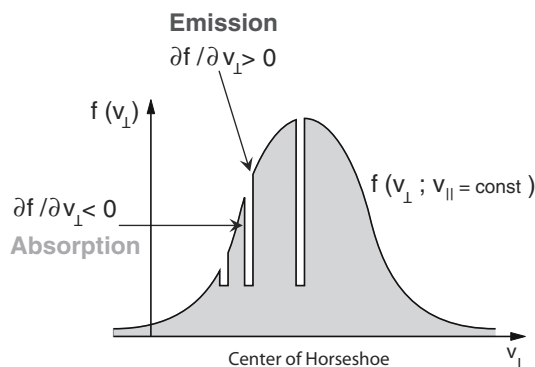


Fig. 30 The mechanism of electron–cyclotron maser radiation emitted from an electron hole. Shown is the cross section of the horseshoe distribution for a given constant parallel velocity v_{\parallel} showing the signatures of deformed electron holes. These have positive and a negative gradient in the perpendicular velocity. The positive gradient generates maser radiation, the negative gradient absorbs radiation. Since both are separated by the hole-width Δ , they belong to different resonance circles. Hence, emission and absorption occur at different frequencies separated by Δf . This implies that the hole emission consists of a combination of an emission and an absorption line. From the resonance condition (1), one concludes that for perpendicular direction of radiation the absorption is at the high-frequency side of the emission

perpendicular velocity gradients in the hole and the finite width Δ of the hole in velocity space, the emissions and absorptions take place at slightly different frequencies. The resonant circles passing through the upward and downward edges of the hole have radii which differ by the velocity spacing Δ . From the resonance condition (1) for strictly perpendicular radiation, one concludes that for constant v_{\parallel} the absorption at the location of smaller v_{\perp} implies absorption at a slightly higher frequency than emission. And a hole which has passed through the maximum of the distribution would have this order reversed. This latter case should, however, be less probable. The frequency spacing is small. From the resonance condition for perpendicular radiation and for a velocity spread Δ of the hole one estimates the frequency gap between emission and absorption to be

$$\frac{\Delta f}{f_{ce}} \sim \frac{\Delta v_h}{c}, \quad (38)$$

where v_h is the resonance velocity at the location of the hole in velocity space where it forms a dip in the horseshoe-distribution. In deriving the above expression, it is assumed that the shape of the hole in the velocity space can be approximated by a resonance circle. For the AKR the spacing between emission and absorption is of the order of $10^{-5} < \Delta f/f < 10^{-4}$. This corresponds to a frequency gap which is of the order of ~ 100 Hz. We note that the relativistic broadening of the emission emerging from a highly relativistic plasma will readily smear out the gap.

The occurrence of a sequence of emissions and absorptions in weakly relativistic plasmas is a very interesting effect. It should permit the identification of hole radiation from the sequence of closely spaced absorption and emission bands. It also should allow us to then determine the widths of the electron holes. Altogether one finds that the phase-space dynamics of electron holes provides the key to an understanding of the intense emission obtained from horseshoe-distributions when electron-holes are generated in the current-carrying plasma. Some signatures of narrow emission and absorption lines during a crossing of the auroral kilometric radiation-source region can actually be recognized in the dynamic spectrum in the upper part of Fig. 24.

6.5 Stability of phase-space holes

Phase-space holes are unstable they lose energy by emitting radiation. However, this loss is minuscule and can be neglected. Moreover, since radiation will not become trapped by the hole, non-linear wave-particle interactions and quasilinear reaction of the radiation on the electron-hole distribution function can be neglected as well. The instability of an electron-hole is nevertheless an important characteristic property which limits the emission from them.

The instability of electron-holes results mainly from waves that are excited inside the hole by the bouncing trapped-electron distribution, and from waves

that are excited at the steep boundaries of the hole. Various types of plasma waves, namely electrostatic oscillations and very low frequency electromagnetic waves, can be excited in these ways. Both these types of waves can exit from the hole but remain to be trapped inside the plasma. They deplete the electrostatic energy that is stored in the holes and erode the electron-distribution function as reviewed by LaBelle and Treumann (2002).

These secondary effects caused by wave excitation in the electron-holes have been investigated in numerical simulations by various authors (Oppenheim et al. 1999, 2001; Muschietti et al. 2000). Their most important prediction was that electron-holes should generate whistler waves. This has been confirmed dramatically by in situ observations of FAST (Ergun et al. 2001b). Ergun et al. (2001b) found very intense emissions of whistler waves when the FAST spacecraft was crossing the environment of electron–cyclotron maser-source regions. These whistlers propagated upward along the magnetic field at the whistler resonance-cone boundary. Emissions of this kind were called ‘saucer emissions’ because of their saucer-like images in the dynamic wave spectrum.

When following the path of these saucers, Ergun et al. (2001b) could identify almost every saucer source with a very localized region in space which itself turned out to be an electric-field structure having the properties of an electron- or ion-hole. These saucers are the result of the instability of the trapped hole-electron population. Propagating away at the local Alfvén speed, the saucer-whistlers transport away a substantial part of the hole energy into the ambient plasma and thereby heat it. This heating by secondary instabilities and saucers is an important effect. In the wave-field of the saucers the particles in the gradient of the horseshoe-distribution experience quasilinear diffusion in pitch-angle and energy; this depletes the general perpendicular gradient of the horseshoe-distribution and produces the plateau at lower electron energies that has been continuously found in the measurements of the electron distribution function. Thereby, it weakens the effect of the contribution of the general gradient in the perpendicular velocity to the maser-emission from the horseshoe-distribution. The main contribution to radiation thus comes from the deformed electron-holes. Finally, we note that the heating of the plasma is most efficient at the boundaries of the plasma cavity. There a substantial amount of plasma is still present in order to absorb the whistler energy.

6.6 Reconnection

Jaroschek et al. (2004b) suggested that a large volume filled with many reconnection sites, which evolve when the plasma is turbulent, will generate detectable incoherent synchrotron-emission. An interesting question is whether such reconnection, one of the most important processes in collisionless plasmas, can map itself into electron–cyclotron maser emission that would be of much higher intensity than incoherent synchrotron radiation.

Reconnection is the process where anti-parallel magnetic fields, which are driven by plasma flow towards each other until they come into contact, merge

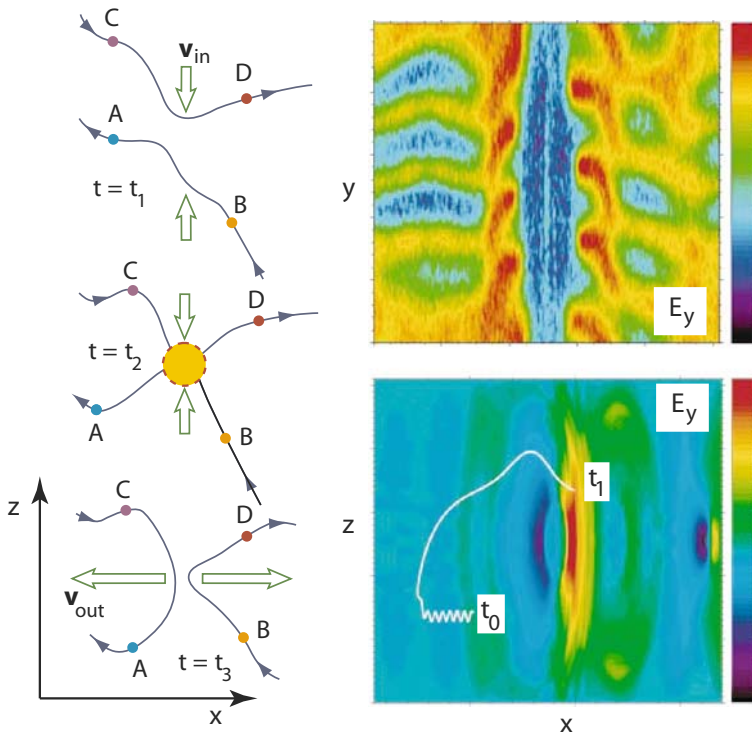


Fig. 31 The physics of reconnection of magnetic field-lines in a schematic representation. *Left, from top to bottom:* Three phases of the reconnection process between two field-lines of opposite direction approaching each other with velocity v_{in} , merging and rearranging at contact, and separating with larger velocity v_{out} determined by the relaxation of magnetic curvature tensions. *Right, from bottom to top:* The (xz) -plane in the bottom part of the figure is the same plane as the planes of the three figures on the left. The (xy) -plane in the top part of the figure is the plane perpendicular to the figure on the left. Both parts show the color-coded electric field E_y in the central reconnection region during merging and rearranging in two sections. *Bottom:* Same plane section as in *left* part of figure. Shown is just the X-point region. The magnetic geometry is about the same as in the bottom graph on the left. Dark blue-to-green and yellow-to-red colors indicate opposite polarities of the electric field E_y . E_y points either in or out of the simulation plane. The white line shows a number of one electron's gyrations in projection and then the path taken between start and end time of the simulation and shows the violent acceleration of the electron in the electric field before it escapes along the magnetic field (in the x -direction) becoming a field-aligned electron beam-particle. *Top:* Same in the perpendicular plane showing the finite extension of the electric field and reconnection region and the wavy structure of the field (*right part* after Jaroschek et al. 2004a). In this section E_y is in the plane. Blue-to-green fields point downward, yellow-to-red fields point upward. All lengths are in terms of the electron inertial length c/ω_{pe} . Times are measured in inverse electron plasma frequencies ω_{pe}^{-1} . Electric fields are normalized to cB_0 , the product of the velocity of light and the magnetic field outside the current sheet

and annihilate. In this process the initially anti-parallel magnetic fields restructure, and the initially strictly separated plasmas mix and become accelerated to high bulk velocities (cf. Fig. 31).

Reconnection in vacuum can proceed without any problem since magnetic fields can reorder there as they like.¹⁶ In a plasma, however, and in particular in a collisionless plasma, simple reordering of the magnetic fields is inhibited by the frozen-in character of the magnetic field. Locally the particles gyrate around the magnetic field on circular orbits at cyclotron frequency ω_{ce} with gyroradius $r_{ce} = v_{\perp}/\omega_{ce}$. Conservation of the magnetic flux, $\Phi = \pi r_{ce}^2 B = \text{const}$, that cuts through the surface of the cyclotron orbit implies that the plasma particles cannot get away from the magnetic field-lines to which they are tied. This is expressed by the Lorentz transformation law for the electric field in moving plasmas which requires that the electric field transforms according to $\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B} = 0$. Hence, in the laboratory system $\mathbf{E} = -\mathbf{v} \times \mathbf{B}$. Reconnection is thus possible only if some diffusion process works which breaks the frozen-in condition.

Numerical simulations of the reconnection process have shown that the diffusion process results in the generation of localized electric fields transverse and parallel to the magnetic field. When plasma is not supplied on a very fast scale to the reconnection site — which happens only in a strongly driven reconnection scenario — reconnection depletes the plasma density around the reconnection site. The plasma leaves from there in two forms: First, the tension forces exerted on the plasma by the strongly bent magnetic field lines accelerate the bulk of the plasma into low-velocity jets perpendicular to the magnetic field, as seen on the left in Fig. 31. The velocity of these jets is a fraction of the Alfvén velocity. Second, the reconnection electric field $\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t$ near the X-point accelerates a substantial fraction of the more energetic tail electrons to high energies. These electrons attain an increase in their gyroradii. When their gyroradius exceeds the width of the reconnection site, they escape and form a high-energy beam along the stretched magnetic field lines.

Because of the fast plasma outflow, the reconnection site itself retains very low plasma densities such that locally $\omega_{pe} < \omega_{ce}$ becomes possible. In the presence of strong guide-magnetic fields that point out of the reconnection plane this condition can even more easily be satisfied. In this case reconnection sites might act as electron–cyclotron maser sources. Similarly, the electron beams ejected from the reconnection site leave along the magnetic field lines which pass through the X-point in reconnection, the two separatrices. Along the separatrices the plasma density is lower than in the surroundings. Here again the above condition can be satisfied. Together, with the passing electron beam that transports the Hall-current, and the increasing magnetic field strength with distance from the X-point — which is similar to a mirror geometry — the separatrix region are also candidates for a working electron–cyclotron maser. The physics of this process has, however, not yet been explored. The geometries are considerably more complicated than in the case of the auroral regions where the magnetic fields possess clear mirror symmetries. Near a reconnection site, the mirror

¹⁶ In fact, in terms of the reconnection picture, the vacuum velocity of light, c , can be interpreted as having the property of a velocity of diffusion of the electromagnetic field across empty space. Such an interpretation is based on the vacuum resistance $(\mu_0/\epsilon_0)^{1/2}$.

symmetry is different from the auroral region and thus the maser-mechanism will also be different. Hollow-beam masers will probably be more important here than ring-shell masers, and the emission will possibly be at higher harmonics of the local electron–cyclotron frequency, in which case it will be weak.

In this context it should be noted that chains of phase space holes have recently been observed (Cattell et al. 2005) in a region where magnetic reconnection was going on in the collisionless plasma sheet of the Earth's magnetotail at roughly 15 Earth radii anti-Sunward distance behind the Earth. A two-dimensional simulation supporting these observations is shown in Fig. 32. The simulation assumes the presence of a strong magnetic guide-field which

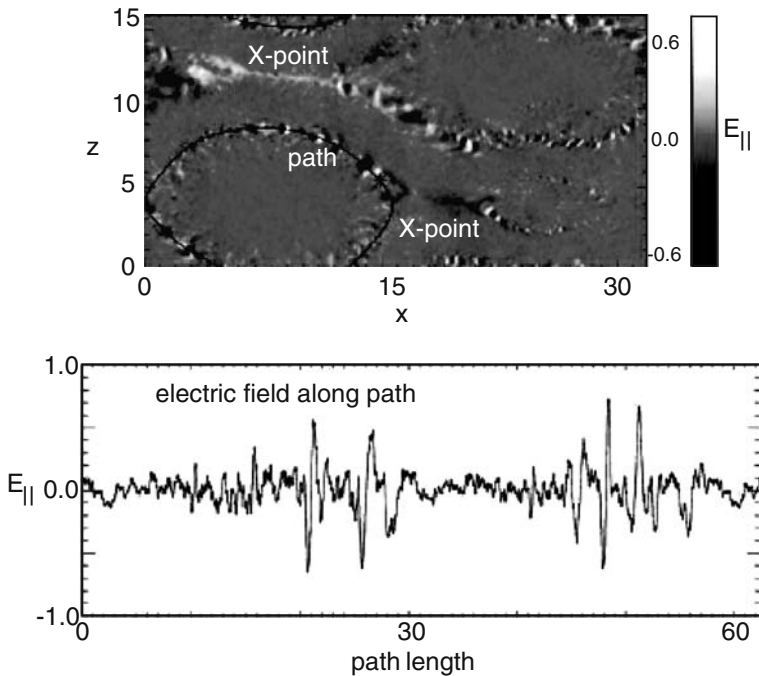


Fig. 32 Chains of electron holes in a simulation of magnetic reconnection including a strong guide field which points out of the plane (after Cattell et al. 2005, with permission by the American Geophysical Union). Shown is the (xz) -plane of reconnection as defined in Fig. 31. Lengths are in terms of the ion-inertial length c/ω_{pi} (where c is the velocity of light, and $\omega_{pi} = \omega_{pe}\sqrt{m_e/m_i}$ is the ion plasma frequency). The gray-scale code of the electric field E is given on the right vertical bar in non-physical computer units. E points out of the simulation plane. Because of technical reasons the simulation is performed in a box containing two vertically (along z) well-separated anti-parallel current sheets (the current in the lower sheet flows out of the plane in $+y$ -direction, the current in the upper sheet flows into the plane in $-y$ -direction). This causes the development of X-points in both current sheets. *Top:* The *upper part* shows the simulation plane with the final magnetic-X-point configuration. Superimposed is the electric field. Along the black about closed magnetic field line (labeled 'path') that connects to the main X-point, the electric field exhibits a chain of localized, small-scale structures. *Below:* The lower part is the amplitude (wave form) of the parallel electric field taken along the 'path'-black magnetic field line. The formation of bipolar and tripolar electron holes is clearly seen

points out of the plane. The upper part of the figure shows the formation of the X-line which is typical for reconnection, and the egg-like structure of the final reconnected magnetic fields. Along the magnetic-field boundary the electric field is highly structured. Its representation in the lower part of the figure exhibits the characteristic structure of bipolar and tripolar electron holes which are generated by the instability of the current flowing along the magnetic field.

Inspection of the density (not shown here) reveals moreover that the plasma along this path is less dense than in the surroundings by up to a factor of ten. This is the result of the presence of a field-aligned component of the electric field which evaporates the plasma locally. Hence, the conditions for the electron–cyclotron maser to work seem to be satisfied. It is, however, not clear whether or not enough energy is available under these conditions in the deformation of the electron-distribution function to feed an electron–cyclotron maser efficiently enough.

Thus, one may conclude that electron–cyclotron maser emission could, in principle, occur at or in the surroundings of a reconnection site. One expects, however, only weak intensities. The main reason for only a weak-intensity emission is that one does not expect horseshoe distributions to be generated in the modest mirror geometry of reconnection. Therefore, reconnection will probably manifest itself in direct electron–cyclotron-maser emission only in an average way when a very large number of reconnection sites — X-points — are embedded into the volume and all the contributions of the single X-points and separatrices add up to the emission.

7 Outlook toward astrophysical applications

The weakly relativistic electron–cyclotron maser emits narrow-band radiation just below the local non-relativistic electron–cyclotron frequency f_{ce} . The highly relativistic electron–cyclotron maser emits relativistically broadened radiation with maximum intensity around f_{ce} . Hence, their observation is a measure of the local magnetic field, B . Its strength is approximately given by

$$B \text{ (gauss)} = 3.57 \times 10^2 f_{\text{GHz}}. \quad (39)$$

From the working condition $f_{pe}^2/f_{ce}^2 \ll 1$ of the electron–cyclotron maser one is then in the position to set an upper limit on the local plasma density in the radiation source region as

$$n_e \text{ (cm}^{-3}\text{)} < 1.24 \times 10^{10} f_{\text{GHz}}^2. \quad (40)$$

Moreover, as we have argued, the source regions of the radiation are small. The radiation, on the other hand, moves at high velocity across the dilute plasma such that the non-linear response of the plasma is negligible. The linear growth of the radiation is limited simply by convection out of the plasma-cavity source region. With L the extension of the source, the transition time of the radiation

is approximately $t_{\text{trans}} \approx L/c$. The wave intensity W_{wave} grows exponentially like $W_{\text{wave}} = W_S \exp(2\langle\Gamma\rangle t_{\text{trans}})$, where $\langle\Gamma\rangle$ is the average electron–cyclotron maser growth rate in the cavity, which is approximately

$$\langle\Gamma\rangle \approx \frac{\pi^3}{\sqrt{2}} \left(\frac{m_e c^2}{k_B T_e} \right)^{\frac{3}{2}} \frac{f_{\text{pe}}^2}{f} Q. \quad (41)$$

Here Q is a (dimensionless) measure of the steepness of the phase-space gradient on the distribution function. It can be taken to be at most in the range of one to two orders of magnitude, $Q \sim 10$ to 100 . $f \approx f_{\text{ce}}$ is the emitted frequency. The wave intensity is proportional to the wave power, $W_{\text{wave}} \sim P$, such that the above expression for the growth of intensity can be written in terms of the power. $W_S \sim P_S$ is the initial wave intensity respectively the initial wave power. Resolving for the transition time provides an estimate for the spatial extension L of the plasma cavity region as

$$L \approx \frac{c}{2\langle\Gamma\rangle} \ln \frac{P}{P_S} \sim \frac{c}{\sqrt{2}\pi^3} \frac{f_{\text{ce}}}{f_{\text{pe}}^2} \left(\frac{k_B T_e}{m_e c^2} \right)^{\frac{3}{2}} Q^{-1} \ln \frac{T_B}{T_S}. \quad (42)$$

The radiation grows out of the available background radiation. This background can be taken as the (initial) incoherent gyro-synchrotron-emission power P_S level at the estimated magnetic-field strength. Then, the logarithm is simply the enhancement factor $\ln(P/P_S) = \ln(T_B/T_S)$ of the coherently emitted power over the incoherent gyro-synchrotron radiation power, which is also the enhancement factor of the coherent brightness temperature over the incoherent gyro-synchrotron brightness temperature, T_S . This factor is typically of the order of a few times 10. Since the density, n_e , and magnetic field strength, B , are already known, the above expression (42) allows to estimate the spatial extension, L , of the plasma cavity, i.e. the electron–cyclotron maser-radiation source. In the most fortunate case when intensity variations are observed, the extension L is also known from light-travel arguments. Then, the expression (42) provides an estimate of either the plasma temperature in the source or the average steepness parameter, Q , which is a characteristic average property of the elementary radiators.

Having reviewed the relevant physics of the electron–cyclotron maser, we now focus on the principal goal of this investigation: the possible relevance of the new physics involved in the electron–cyclotron maser for astrophysics. As usual, such an investigation will have the character of an outlook only, as it is impossible to provide in situ observations of any remote astrophysically interesting systems, exceptions being the nearest planets in our own solar system and some occasional measurements at the outskirts of the heliosphere.

Even the Sun, the nearest star, is not accessible to in situ observations. The Sun hides its secrets on most of the interesting physics which underlies the emission of radio waves under the skirt of its high optical and X-ray radiation power. Nevertheless, astrophysics in this case, like in many others, may use our

knowledge about the relevant physical processes and mechanisms that can be accessed directly on Earth and in the Earth's environment, and look toward applying this knowledge in the investigation of exotic objects in distant space. It is one of the beauties of scientific progress that we can expand the bounds of our knowledge by ever more precise measurements in our accessible environment, thereby providing ever better evidence for the reality of such astrophysical processes.

7.1 Other planets: Jupiter, Saturn

The paradigm of the electron–cyclotron maser emission is the auroral kilometric radiation of Earth. It is tempting to extrapolate from Earth to the other strongly magnetized planets of the solar system and to search for auroral radiation from these planets. Radio emissions from Jupiter and Saturn, the nearest outer strongly magnetized planets in the solar system, have been known (Zarka et al. 1986, 2004; Zarka 1992a,b, 1998, 2004) for a long time from ground-based observations. Jovian radio emission in S bursts (see Fig. 33) reaches brightness temperatures of 10^{18} K at 30 m wavelength strongly suggesting a nonthermal emission mechanism.

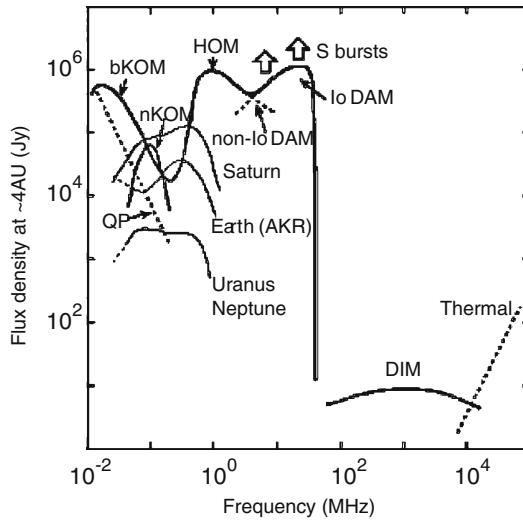
The five families of Jovian radio emissions, listed in Table 1, are believed to be generated by the electron–cyclotron maser mechanism. Some of the emissions are clearly related to the presence of the active Jupiter satellite Io which carries with it a magnetic flux tube around Jupiter and is thus strongly coupled to Jupiter. Io exchanges plasma and currents with the planetary atmosphere of Jupiter, and these currents give rise to intense radio emissions whose strength and spectrum vary strongly with time. However, it is not clear whether or not these radio emissions are actually generated by the electron–cyclotron maser mechanism as the plasma configuration and the actual shape of the electron-distribution function are not really known.

Table 1 Electron-cyclotron maser radio emission of Jupiter

	Type	Frequency	Source	Power (W)	Mechanism
I	Non-Io DAM	<40 MHz	Auroral field-lines	10^{10} – 10^{11}	Ring-shell Horseshoe
II	Io DAM (L & S bursts)	Io field-line	Io torus	$< 10^{11}$	Ring-shell Ring-beam
III	HOM → few MHz	200 kHz ~ 70°	Auroral	10^8 – 10^9	Hollow beam
IV	bKOM	10–300 kHz	Auroral	10^8 – 10^9	Horseshoe?
V	QP bursts, <i>nth</i> -cont	< 700 kHz	Auroral MP	$\sim 10^8$? ?

DAM decametric, *HOM* hectometric, *bKOM* broadband kilometric, *nKOM* narrow band kilometric, *L* long (minutes), *S* S-shaped short (few tens of ms) bursts, *nth-cont* non-thermal continuum, *MP* magnetopause

Fig. 33 Comparison of the spectral intensities of the various radio emissions from the outer planets and Earth (after Zarka 1998 with permission by the American Geophysical Union). Jovian spectra, including those originating from the Io torus, are shown as *bold lines*. (The meaning of the abbreviations is HOM = hectometric, DAM = decametric, bKOM = broadband kilometric, nKOM = narrow-band kilometric, QP = quasi-periodic pulses, S bursts = S-shaped short radio bursts, DIM = decimetric)



Provided that intense electric fields evolve in the Io flux tube, the plasma may become sufficiently diluted so that favorable conditions are created for the electron-cyclotron maser close to the Jupiter surface. The magnetic field would indeed be strong enough, and one believes that the magnetic-field-aligned currents that flow along the Io flux tube and couple the satellite to the planet are so strong that they undergo current instability. In this case they excite waves, generate Bernstein-Green-Kruskal (BGK) modes, and support large-scale field-aligned electric-potential drops — or double layers — in the Io flux tube. Moreover, strong shear flows could generate a double layer as well as is the case in the Earth's auroral plasma cavity, where external reconnection is the driver of the shear flow. Such shear flows may, for example, arise when Io's flux tube moves across the ambient plasma together with the satellite around Jupiter and causes vortices in the plasma flow.

Spectacular aurorae have been observed on Jupiter in the ultraviolet spectral domains by the Hubble Space Telescope, and similar emissions have been seen on Saturn. Figure 34 shows such an observation centered around Jupiter's north pole. The bright spot on the left indicates the magnetic foot point of the Io flux tube in the Jupiter atmosphere. The bright emission coming from there results from the Io plasma which collides with the constituents of the Jovian atmospheric constituents. The extended bright emissions forming the Jovian auroral oval are caused by the same processes as aurorae on Earth, resulting from reconnection at Jupiter's magnetopause and from reconnection in the Jovian magnetotail. Similar auroral rings are observed around the poles of Saturn. They are shown in the right-hand panel of Figure 34.

The deformation of the auroral oval seen particularly on the night side of Jupiter — and to some extent also on Saturn — are caused by shear flows in Jupiter's tailward plasma content in the wake of ongoing reconnection. From this one can conclude that reconnection-generated shear motions will be strong enough to

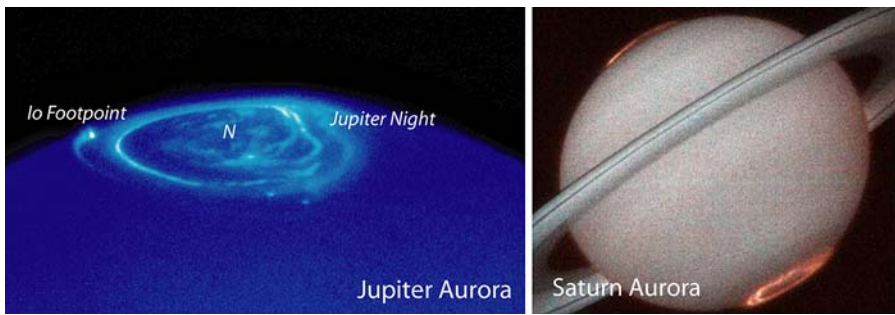


Fig. 34 Hubble Space Telescope views of planetary aurorae. *Left:* Aurora around the pole in the northern hemisphere of Jupiter. Similar to aurora on Earth, auroral activity occurs in a circle that is approximately centered at the magnetic pole, the so-called auroral oval. Seen from space the aurora occurs in several irregular bands which indicates the dynamics of the auroral oval, ongoing Jupiter substorm activity and the immense dynamics of Jupiter's plasma-magnetic-field environment. Also seen on the dayside is the intense emission from the footpoint of Io's magnetic flux tube on the surface of Jupiter; *Right:* Ongoing aurora in both polar caps of Saturn. On both giant planets of our solar system, the aurora appears bright and irregularly structured. In all cases its location encircles the magnetic poles of the planets which it shares with Earth's aurora. This circle is formed by the footpoints of the magnetic field lines which connect the planet to the regions of reconnection at the magnetopauses and in the magnetic tails of the planets. Substorm activity on these planets is related to the ongoing magnetic reconnection in their nightside magnetotails, generation of magnetic-field-aligned currents and the existence of an energetic plasma component in the planetary tails. The latter stems from solar-wind plasma entering the magnetosphere by reconnection at the magnetopauses, and from plasma of the planetary atmospheres, from their rings, and from their satellites, like Io in the case of Jupiter

generate the required auroral currents and cause field-aligned potential drops to develop, even though in situ measurements confirming such a conjecture have not yet been made. But the similarity to the Earth's auroral motions and flux-tube deformations is so tempting that the analogy must be invoked! With magnetic-field-aligned electric fields present, all the constraints on the electron–cyclotron maser would be fulfilled in Jupiter's — and as well in Saturn's — auroral zones. Therefore, the assumption that auroral radio emissions in the hectometric (HOM) and kilometric (KOM) bands and the short and fast drifting S-burst radio emissions, which are believed to be related to the Io flux tube, are caused by the electron–cyclotron maser in the ring-shell mechanism are reasonable.

The loss-cone maser might also contribute to emission, but the experience with the dominance of the ring-shell maser and emission in the nearly perpendicular RX-mode are strong enough to support the assumption that the fast moving striations in Jupiter's radio emissions, like those of the S-bursts, are caused by chains of microscopic electron-holes and ion-holes that move across the plasma in the auroral cavities on Jupiter in connection with the double layers.

Recently, Zarka et al. (2005) argued that S-bursts are the result of loss-cone maser activity inside the Io flux tube. They find consistency with a loss-cone maser model in frequency drift and source-electron energy of ~ 4 keV inside the

flux tube. In addition, however, they find evidence for a magnetic-field-aligned electric potential drop of 1–3 keV which coexists with the emissions. They attribute this electric field to Alfvén waves along the Io flux tube. Moreover, they advocate the validity of adiabatic theory for the motion of energetic electrons.

In the light of the arguments in the present review, such large potential drops together with adiabatic motion of electrons along the converging Io flux-tube would rather indicate a deformation of the electron-distribution function into a ring-shell distribution and the action of the much stronger ring-shell maser. One could easily argue that such parallel potential drops inside the Io flux-tube could be generated by the motion of the flux-tube across Jupiter's magnetosphere. This necessarily causes circulation around the flux tube and thus shear motion of the plasma at its boundary. And, following our previous arguments, this will lead to field-aligned potentials drops and plasma evaporation on a smaller scale. Both these consequences argue in favor of the ring-shell maser. No doubt, Alfvén waves and in particular kinetic or shear Alfvén waves will be involved in the transport of information along the magnetic field since they are the main mode for exchanging information along the magnetic field in a plasma. But in order to account for high potential drops, these modes must become highly non-linear so that it is not reasonable anymore to speak about single Alfvén modes as the ultimate energy source of the radiation.

Recently, the Cassini mission has provided first high time and frequency resolution results of auroral radio emissions from Saturn (Kurth et al. 2005a,b). These observations indicate that Saturnian auroral radio emissions are strikingly similar to those in the Earth's auroral plasma cavity, exhibiting the same kind of drifting narrow band structure (Pottetlet et al. 2001) observed here. It is therefore highly probable that the mechanism generating the Saturnian auroral radio emissions is the same as that on Jupiter and Earth, of which we today believe that they are caused by the ring-shell maser mechanism. All these radio emission similarities among the magnetized planets in the solar system suggest that the electron-cyclotron ring-shell maser mechanism is the most efficient radiation mechanism even though loss-cone maser contributions cannot be excluded.

No generally agreed upon mechanism of generation of the electron-cyclotron maser fine structure emission does yet exist (Pritchett et al. 2002). However the arguments for the involvement of double layer electric fields, electron and ion holes have now become overwhelming, even though the very fine structure of the Saturnian or Jovian auroral radio emissions cannot yet be resolved as well as that of the Earth's kilometric radiation. The more, however, these structures can be resolved the more will a model develop which confirms the electron-cyclotron maser emission mechanisms as resulting from the generation of steep perpendicular gradients in the electron velocity distribution by deformed electron holes on the smallest scales. And the radio emission will be seen as the superposition of the radiation from many such microscopic elementary radiation sources.

This conclusion is far reaching. It provides a universal mechanism for application to other remote, strongly magnetized systems, planets, stars, brown dwarfs,

etc., allowing from observation of highly non-thermal radio emission from such objects to draw conclusions about the structure and parameters of their magnetized plasma environments and their dynamics. It has been concluded (Mitchell et al. 2005) that Saturn exhibits similar magnetic substorms as the Earth. Substorms are generated by reconnection as noted above and provide the ultimate energy source of the auroral electron–cyclotron maser emission. The observation of substorms in remote objects together with coherent radio emission confirms the identification of reconnection as a widely realized universal mechanism of redistribution of energy in collisionless magnetized plasmas.

7.2 Exoplanets

The discovery of extrasolar planets has immediately stimulated the idea that in analogy to the strongly magnetized planets in the solar system, exoplanets could also radiate in the radio band, possibly emitting maser radiation (Farrell et al. 1999; Bastian et al. 2000; Zarka et al. 2001b; Winterhalter et al. 2005). Farrell et al. (1999) suggested that detection of exoplanets in such an emission would be most favorable in the meter and decameter bands. From the knowledge of the emissivities of Earth and Jupiter under solar-wind conditions, one may scale emissivities up to exoplanets of the family of ‘hot Jupiters’ close to their mother stars (Zarka et al. 2001b). One expects intensities that are higher by factors 10^3 to 10^5 than those of Jupiter’s decameter radiation. Radiation of such high intensities may become detectable with the largest radio-telescopes that already exist or are under construction. Searches have been going on at the Very Large Array (VLA) in New Mexico (Bastian et al. 2000) and a comparable array UTR-2 in Kharkov (Ukraine) (Zarka et al. 2001b).

High-resolution observations have already been performed with the VLA by Bastian et al. (2000) who looked at a number of extrasolar planets at 333 and 1,400 MHz. In addition they observed 47 Ursa Major at 74 MHz, but did not find any signature on any of these objects, so far. The Kharkov observations were also not able to detect radio emission from their exoplanet candidates. This negative result might be a consequence of either observing in the wrong frequency band, insufficient sensitivity and insufficient frequency resolution, incorrect assumptions on the expected magnetic field strengths of exoplanets, or incorrect assumptions about the magnetic interaction between the mother star and the exoplanet. The latter might not be of the same kind as the interaction between Io and Jupiter.

In fact, the exoplanets which have been taken as candidates are those which belong to the class of planets which had not been expected to exist. They are very large and very close to their mother stars. Their interaction with the mother star is probably very intense. If the planets are magnetized it is possible that their magnetic field is tied to the star or that it at least interacts very strongly with the stellar magnetic field. This has stimulated the idea that reconnection might dominate the interaction. One would not, however, be able to observe the reconnection site in radiation. Nevertheless, one then expects that strong

currents flow along the field-lines up and down to the planet and generate violent phenomena which should be similar to the phenomena in the auroral plasma cavity. If this held, then the planet should emit electron–cyclotron maser radiation. It is, however, not yet clear in which frequency band this emission would occur.

Another serious argument why radio emission from the electron–cyclotron maser might not be detectable from exoplanets is strong self-absorption of the radiation at the fundamental. This argument has been put forward by Melrose (1999) in the context of solar radio spikes. In fact, in the case of solar radio spikes the argument might be less strict than it sounds (see below). The known exoplanets have, however, been found very close to their mother stars and are thus probably in very dense plasma environments. Even though in such an environment plasma cavities will exist close to the planetary surface, it is not unreasonable to assume that the optical depth for the radiation becomes too large outside the source region, so that the radiation will be fully absorbed over the pathlength it has to travel in order to escape from the planet. This situation is different from that of the planets in the solar system where the radiation can either tunnel through the walls of the narrow plasma cavity, or it can escape by performing wave transformation; it is, however, not unreasonable to assume that strong self-absorption occurs in the exo-planetary environment. Any escaping radiation must then be at higher harmonics of the electron–cyclotron frequency where the excitation is much weaker.

Far higher instrumental sensitivities in the frequency range from 10 to 240 MHz have been announced (Farrell et al. 2004) to be expected from the new Low Frequency Array (LOFAR) which is currently under construction. Its more precise observations may confirm that exoplanets are, in fact, radio emitters.

7.3 The quest for solar and stellar radio-maser emission

Solar radio emissions have not only been the first candidates of electron–cyclotron maser action, with Earth and planets they continue to be a potential host of electron–cyclotron masers. In spite of its proximity, the Sun — like other more remote astrophysical objects — does not allow the investigation in situ of the various potential source regions of the expected maser emissions. This impossibility has not inhibited and should not inhibit attempts to apply maser theory to solar phenomena. Actually, there is a long tradition: the first ideas of the possibility of a direct amplification of the free-space modes appeared in solar physics. Coronal electron-temperature anisotropies were identified as the energy source of the maser emission. This assumption was based on analogies with radiation-belt electron trapping at Earth, but turned out to be ineffective. Subsequently, following the work of Wu and Lee (1979) on the weakly relativistic electron–cyclotron maser, the assumption of loss-cone distributions survived until very recently as the basis for almost all solar application of the maser mechanism.

Since the structure of the solar magnetic field close to the photosphere, i.e., in the regions where the electron–cyclotron maser may be assumed to work, is extremely complicated there are very many such source regions that can be distributed all over the solar surface at various distances, and resolution of them will still for a long time be inhibited by their small scales and short survival times. Experience with the Earth's AKR indicates that the size of a terrestrial plasma cavity, which is determined by the presence of shear flows above the source region, is restricted to merely several 10 km to 100 km. This may be larger on the Sun because of the larger dimensions of the solar active regions, but shear flows at the base of the corona will probably not exceed scales of say 10^3 km to 10^5 km, and shear flows caused by turbulence might occur on even smaller scales. Since large-scale magnetic-field-aligned electric fields cannot be maintained over very long times, externally driven shear flows will be the most important in generating magnetic field-aligned electric-potential drops and parallel electric fields.

Numerical simulations have suggested that plasma cavities are highly time-variable. Such cavities will therefore decay into even smaller structures. Their overall size is determined by the size of the reconnection and acceleration region of which we know today that they are not larger than a few ion-inertial lengths in the transverse dimension. Located at coronal altitudes — say at plasma densities of 10^9 cm^{-3} — their diameter will not be larger than a few times ~ 100 m at best. In mirror-magnetic-field geometries, this distance maps down at the foot-points of the magnetic field to even narrower regions of which a multitude fit into the environment of one solar active region. One hence expects that plasma cavities of the size of some 10 m will be created, and will be elongated both in the azimuthal direction and in the vertical direction along the magnetic field. Such regions are the generator regions of electron–cyclotron maser emission. Field-aligned electric fields will evolve in them and generate ring-shell and horseshoe distribution functions on the electrons and cause the ions to escape along the magnetic field in the form of cold fast ion beams. The emitters themselves, being identified as BGK modes, are structures of the size of a few electron Debye lengths only, i.e. of the size of a few centimeters, in which case myriads of them will make up the radiation, and none of them will ever be resolved by observation.

From the observational side, extremely high brightness temperatures have been reported from the Sun for the solar radio spikes (Melrose and Dulk 1982; White et al. 1983; Willson 1985; Benz 1986). Similarly high brightness temperatures in the radio emission have been found for flare stars (Kuijpers 1985; Benz et al. 1998; Stepanov et al. 2001), dwarf M stars (Lang et al. 1983; Lang 1994), and recently in the bursty radiation of T Tauri stars (Smith et al. 2003). In addition, the fine structure of solar type IV radio bursts has often been attributed to maser action [cf., for instance, the reviews by (Aschwanden 1990a,b; Bastian et al. 1998; Fleishman et al. 2003; Fleishman 2006)]. This might or might not be the case. The models developed so far require confinement of the radiation and are thus not entirely convincing. We will, therefore, not enter into a discussion of them.

Solar microwave spikes are short radio bursts of brightness temperature $T_B \sim 10^{12}$ K. They are observed at frequencies from ~ 100 MHz to 5 GHz and have a duration of < 100 ms during solar flares. Their circular polarization is $\sim 100\%$ and their bandwidth is narrow, namely only of a few percent of the bandwidth of the background emission. These radio emissions are clearly expected to be generated by a coherent process like the one which is found in the electron–cyclotron maser mechanism. This conclusion is backed by the very high degree of circular polarization of these emissions, which identifies them as being radiated in the RX-mode close to the local electron–cyclotron frequency ω_{ce} . If this is true, the emission maps the strength of the magnetic field in the emission region and identifies the source as a plasma cavity with $\omega_{pe}/\omega_{ce} \ll 1$. Some of these emissions do practically not drift across the spectrum, and therefore other mechanisms like electron-beam excited coherent radiation (which occurs in the case of solar type-III bursts) are inappropriate as a generation mechanism.

Melrose (1999) has criticized the application of the electron–cyclotron maser to all these emissions by pointing out that the emission suffers from high self-absorption of the radiation near the fundamental when trying to escape from the plasma. He shows the inefficiency of several mechanisms to help overcoming self-absorption, among them tunneling. However, while tunneling might be no way for the radiation to escape, it has been shown above that radiation can escape from inhomogeneous media in several ways, namely by being channeled along the inhomogeneity, or by being transformed to the Z-mode and then, at the outer boundary of the plasma, being re-transformed into the RX-mode. In rare cases the RX-mode radiation may even be transformed into the LO-mode as it happens in incomplete wave-guides. The LO-mode will then freely escape since its frequency is above the plasma frequency — its low-frequency cutoff. In all these cases the radiation would be able to escape, at least in part. Hence, self-absorption might not provide a really serious argument against the electron–cyclotron maser mechanism, in particular not when the radiation is very intense, so that it is sufficient to convert only a percentage of it into escaping radiation.

As the loss-cone maser is highly improbable as source of solar radio spikes, the shell maser provides an effective alternative. Here the emission is below the electron–cyclotron frequency and the electrons are weakly relativistic, so that the effective temperature of the electrons satisfies the approximate condition $k_B T_e / 2m_e c^2 > \gamma^2 f_{pe}^2 / f_{ce}^2$. The density in the source region is very small such as $f_{pe}/f \sim 0.01$. Hence, self-absorption in such a dilute plasma and for perpendicular propagation is negligible. What is left unclear is only the problem of radiation escape.

Flare stars are similar candidates. The radio image of one of them, UV Ceti – a nearby flare star –, has recently been obtained with the VLBI at 3.6 cm wavelength by Benz et al. (1998) exhibiting a brightness temperature of $T_B > 1.2 \times 10^{12}$ K. The image showed two emission components separated by 4.4 stellar radii. The alignment of the components is similar to the most likely axis of rotation and magnetic field for the star — that is, the two components are apparently located above the polar regions. The authors interpreted the two

components as being caused by huge magnetic flux tubes, whereby the magnetic field was estimated, based on the gyro-synchrotron mechanism, to be between 15 and 130 gauss and the electron density was between 10^6 cm^{-3} and 10^8 cm^{-3} , respectively. All the radio flares were entirely right-hand polarized. Stepanov et al. (2001) have presented similar observations of the dMe-star AD Leo at 4.85 GHz with a bandwidth of 480 MHz and found brightness temperatures $T_B > 3 \times 10^{13} \text{ K}$ in strictly right-hand polarization. Assuming electron–cyclotron maser emission they estimate magnetic field strengths $B \sim 800$ gauss and electron densities $n_e \sim 2 \times 10^{17} \text{ m}^{-3}$.

Such observations are in clear agreement with the electron–cyclotron maser emission mechanism of coherently generated radiation. Bingham et al. (2001) have successfully applied the shell-maser mechanism to the case reported in (Benz et al. 1998). Given a plasma frequency for the above estimated densities of only ~ 100 MHz, and taking the observed frequency of ~ 8 GHz as the approximate cyclotron frequency, we are clearly in the range of a dilute plasma embedded in a strong magnetic field of order of 300 gauss. The dilution on the other hand points to the presence of a substantial field-aligned electric field that is present in the source region. Such a field, in the converging magnetic-field geometry of UV Ceti must generate a shell-horseshoe distribution and will therefore satisfy all the conditions of a RX-maser in perpendicular propagation. This is also in agreement with the completely circular polarization of the radiation in the RX-mode.

Recent VLBI observations of T Tauri South (T Tauri Sb) (Smith et al. 2003) also indicate very high brightness temperatures of $\sim 10^6 \text{ K}$ during the burst phases of the emission, which comes from very compact sources, and exhibits 100% circular polarization. These observations are again typical for electron–cyclotron maser action, and have been used to estimate the magnetic field of the star to be of the order of 1.5–3 kilogauss.

Coming, finally, to the dwarf M stars, radio spikes at 20 cm wavelength (i.e., 1.5 GHz) with a duration of less than 100 ms and with $\sim 100\%$ circular polarization have been reported. The emission amplitude is in excess of 100 mJy (Lang et al. 1983). Since the bursts are very impulsive, the source region must be small — based on light travel-time arguments, much smaller than the stellar diameter. This implies brightness temperatures of the order of $T_B > 10^{15} \text{ K}$, clearly indicating a coherent process (Lang 1994). It is reasonable to assume that the electron–cyclotron maser would be responsible for the observed emissions. However, from the available information it cannot be decided whether or not the loss-cone or the shell maser is at work in these objects. This would require higher time- and frequency-resolution that could reveal fast moving structures. It also would require an indication of the plasma density in these objects. Assuming that the emission is at the fundamental, the magnetic field in these objects should be of the order of 60 gauss. In order to satisfy the cavity condition the plasma density in this case must be not larger than a few times 10^9 cm^{-3} . Such densities correspond to the plasma density in the solar corona.

7.4 Emission from pulsars

The generation of radio emission from pulsars is much less and in fact insufficiently understood. The emission is intrinsically very bright, with brightness temperature reaching values up to $T_B > 10^{25}$ K which suggest a coherent mechanism probably involving some cyclotron instability.

The plasma in the pulsar magnetosphere close to — and probably also inside — the source region of the radiation is highly relativistic. The mechanisms that have been proposed to explain the generation of intense radiation belong to the direct and indirect emission mechanisms and range from curvature radiation, to linear acceleration of electrons and positrons to the free-electron-maser emission, or to relativistic plasma emissions, which are indirect. Magnetic fields in pulsars are superstrong with estimated strengths of $B \sim 10^{11}$ gauss to 10^{13} gauss. Millisecond pulsars have magnetic fields four to five orders of magnitude weaker. The plasma in pulsar magnetospheres can freely escape from the polar region, but is trapped in the closed-field regions around the equator.

In current models it is assumed that very large electric potential drops form along the field-lines in the polar cap and lead to the spontaneous creation of electron-positron pairs which form an electron-positron plasma. In the lowest Landau levels the radiation emitted is gyromagnetic emission. Estimates of the charge density can be given, but are quite uncertain.

In any case, for the objects with strong magnetic fields one expects that $f_{ce} \gg f_{pe}$ which, together with the strong electric fields and the mirror geometry of the magnetic field, seems to be in favor of the electron-cyclotron maser emission mechanism.

To provide an estimate we consider the typical parameters of a pulsar magnetosphere. The surface magnetic-field strength in the pulsar magnetosphere can be expressed through the pulsar rotation period P and its time derivative \dot{P} (Manchester and Taylor 1977) as

$$B_0 \text{ (gauss)} = 3.2 \times 10^{19} (P\dot{P})^{\frac{1}{2}}. \quad (43)$$

Similarly, the plasma frequency is given by (Manchester and Taylor 1977)

$$\omega_{pe} = \left(\frac{8e\gamma\Omega B_0}{m_e c} \right)^{\frac{1}{2}} \left(\frac{R}{r} \right)^{\frac{3}{2}}, \quad (44)$$

where r is the radial distance, R the pulsar radius, Ω its rotation frequency, and γ the relativistic factor of the accelerated electron-positron plasma inside the pulsar magnetosphere. The ratio of plasma to cyclotron frequency of streaming — with gamma factor γ_s — particles in a dipolar is then given by

$$\frac{\omega_{pe}(r)}{\omega_{ce}(r)} = 1.7 \times 10^{-9} \left(\frac{\gamma\gamma_s^2}{PB_{0,12}} \right)^{\frac{1}{2}} r^{\frac{3}{2}} \quad (45)$$

where $B_{0,12}$ is the surface magnetic field measured in 10^{12} gauss. Therefore, closer to the pulsar surface the ratio of plasma-to-cyclotron frequency is very small, and very large gamma factors would be required to reduce this ratio to unity. However, under the special conditions in the pulsar magnetospheres direct electron–cyclotron maser emission seems to work only for the non-escaping modes via the anomalous Doppler resonance (Machabeli and Usov 1979). Emission near the light cylinder has also been proposed by Machabeli and Usov (1989) and Kaabegi et al. (1991). Thus, the theoretical emission height is much farther out than has been inferred from observation (Hoensbroech and Xilouris 1997).

Thus, all mechanisms including the electron–cyclotron maser process encounter serious difficulties in an electron-positron plasma. Recently a hollow-beam cyclotron-maser mechanism has been proposed by Ma et al. (1998) that should work, however, for streaming positrons and for the LO-mode only. Whether this result can be maintained is uncertain. It is also not clear whether an analogy can be drawn between ordinary magnetized planets with their auroral emissions and pulsar emissions. This leaves open the problem of how pulsars can so intensely radiate, and what role could be played by the highly relativistic (ring-shell) electron- and positron-cyclotron maser under the conditions close to the pulsar.

7.5 Coherent radiation from Blazar jets

The most recent and most speculative attempt of application of the electron–cyclotron maser by Begelman et al. (2005) concerns the time-varying emission from Blazar jets.

Blazar jets are strongly magnetized, relativistic low-density plasma jets ejected from a central machine which is believed to be an accreting massive rotating black hole. Blazar jets are visible at radio frequencies, where they exhibit rapid intra-day variability at GHz frequencies. By using light-travel arguments one can infer angular sizes of these sources and thus some of these sources arrives at brightness temperatures $T_B \sim 10^{21}$ K, exceeding the synchrotron self-absorption upper limit of $\sim 10^{11}$ K by a very large factor. This can be reduced only by invoking a relativistic boost of the temperature by a factor of γ and an increase in solid angle by an additional factor of γ^2 . Relativistic factors of $\gamma \sim 10^3$ would thus be able to reduce the observed brightness temperature to the above limit. However, as Begelman et al. (1994) argued, synchrotron efficiencies in high- γ jets are rather low. This would require unreasonably large energy fluxes in the jet to explain the observed radiation intensities. Thus a mechanism that creates very high brightness temperatures in a natural way offers a practical way out of the dilemma.

Such a mechanism can, under several, not entirely unreasonable assumptions, be provided by the electron–cyclotron maser and thus replace the incoherent synchrotron mechanism. It should be noted that other coherent radiation mechanisms based on the non-linear evolution of various plasma waves in the jet

plasma have been proposed as well. These mechanisms do, however, require large plasma densities and most important, given plasma frequencies $f_{pe} > f_{ce}$ increasing the emission frequency. This is the opposite limit to the more reasonable one of the electron–cyclotron maser which is expected to exist in the jet in the strong magnetic fields that are generally assumed. Such high plasma densities must still be justified.

The plasma in the jets moves at a velocity that is practically the velocity of light. Since the magnetic guide-fields of the jet are believed to be very strong, the maser condition $f_{ce} \gg f_{pe}$ is readily satisfied. The second condition, the inversion of the electron (and possibly also the positron) population, requires the presence of a mirror magnetic field. Its existence is crucial for either the ring-shell cyclotron-maser or the loss-cone maser mechanisms to work. There are clearly no direct observations of such populations. Since the magnetic fields are very strong and may be too stiff for a susceptible convergence, one must assume that mirror geometries are generated locally by some mechanism like magneto-hydrodynamic instabilities, turbulence, and shock waves. Either of these mechanisms might locally force the magnetic field on a large enough scale into a mirror geometry along the current-carrying jet flux-tube. Proposals of this kind are now known in the literature. For instance, Bingham et al. (2003) proposed that turbulence in strong shocks may locally produce sufficient mirror-magnetic field-structure.

Another possibility is the Weibel instability (Jaroschek et al. 2004c, 2005) which occurs when the jet is pulsed irregularly by its central machine. The Weibel instability generates transverse magnetic fields that may locally bend the jet magnetic field thereby causing mirror geometries. Once this happens the developing field-aligned electric field will accelerate the particles in order to produce ring-shell or horseshoe distributions. This then sets the framework for the cyclotron maser in the jet.

Assume that local mirror magnetic-field geometries are generated by turbulence, shocks, or Weibel instabilities inside a jet. Also assume that parallel electric potential drops arise as a consequence of the turbulent shear motions that apply to the magnetic field — or otherwise are generated by current instabilities which are driven by the field-aligned current inside the jet. We then have arrived at the scenario in which a ring-shell distribution function can locally be created on the scale of the turbulence. Each of the small mirrors contributes to the emission of radiation, and the radiation from all these turbulent mirrors adds up to intensities that are high enough to match the high observed brightness temperatures of $T_B \sim 10^{15}$ K which have been estimated for Blazars.

Such a scenario has recently been presented by Begelman et al. (2005). In this theory the distance r of electron–cyclotron maser emission at frequency f from the central engine of the jet depends on the (average) mirror ratio $\rho = B(r)/B_m$ of the small turbulent mirrors and the jet power L . Here B_m is the mirror-magnetic field-strength emission for $\rho \sim 5$ occurs at a typical distance of 10^3 gravitational radii for a central-engine mass of $10^8 M_\odot$, where M_\odot is the solar mass. In order to reproduce the observed luminosity (i.e., the brightness temperature) of the Blazar radiation, Begelman et al. (2005) need only a very

small fraction of surface coverage in the sky by cyclotron-maser sources — of the order of 10^{-5} to 10^{-4} . The volume filling factor is then only $\sim 10^{-15}$, and the total number of sites required in one jet amount to $N \sim 3 \times 10^{15}$ which is very small considering their small size. The total broadband maser power measured at the distance $r = r_{\text{obs}}$ of observation can then be estimated as

$$P(r = r_{\text{obs}}) \sim 9 \times 10^{25} \rho^2 T_{\text{B15}} L_{38} f_{\text{GHz}} \gamma^{-4} \quad \text{W}, \quad (46)$$

where T_{B15} is the brightness temperature measured in units of 10^{15} K, L_{38} is the total jet power measured in 10^{38} W, and f_{GHz} is the frequency measured in gigahertz. Thus the broadband isotropic power is boosted by the Lorentz factor to fourth order. The volume-averaged emissivity of the maser radiation in this case at distance r_{obs} is only a fraction $\sim 2.5 \times 10^{-10} \rho^2 T_{\text{B15}} \gamma^{-2}$ of the comoving energy density passing through the observation point r per unit of time. These low volume efficiencies are both surprising and encouraging as they show that only very little kinetic or magnetic energy is needed to be converted into radiation in order to produced extraordinarily high brightness temperatures.

Whether the maser-radiation can escape from the jet is a different problem that has been discussed extensively by Begelman et al. (2005). They came to the conclusion that for all but the most extreme parameters the jet plasma would be optically thick to the radiation. Radiation will probably be observed only from a narrow, optically thin boundary layer of the jet and will produce high brightness temperatures without excessive demands on the model parameters. The required electron energy densities remain far below equipartition energies. Hence, only little magnetic energy is dissipated in the maser process. In conclusion, then, the electron–cyclotron maser model seems to provide a very interesting and very successful model of emission from Blazar jets.

8 Concluding remarks

The electron–cyclotron maser is capable of generating intense and coherent radio emission under some very peculiar – but not unlikely — conditions. It has, therefore, already become a powerful diagnostic tool for such conditions. Its theory has been developed up to quite a mature level. Once the electron–cyclotron maser can be identified as generator mechanism in some remote object, it permits us to infer the magnetic field strength and the density in the radio source. It also allows us to infer some qualitative properties of the source: the presence of an electric field along the magnetic field-lines, the presence of large-scale plasma cavities and the degree of density depletion and, finally, the energy of the resonant electrons. Moreover, the measured degree of polarization tells us whether the radiation can escape directly from the source in the RX-mode or is partly transformed into the LO-mode when passing across regions of higher density.

We have presented two versions of the cyclotron maser: the bunching maser and the genuine electron–cyclotron maser. Theoretically both are two sides

of the same coin. However, under natural conditions the electron–cyclotron version of the maser is probably more frequently realized since the plasma distributions found in the source are in the overwhelming majority of cases hot, and the mono-energetic approximation ceases to be valid. In ultra-relativistic plasmas, however, the situation might be reversed. Such plasmas probably radiate via the mechanism of particle bunching; their spectrum should become relativistically broadened resembling a synchrotron spectrum of very high brightness temperature. Applications of the bunching maser have not yet been attempted to ultra-relativistic objects.

The canonical paradigm of the electron–cyclotron maser emission is the Earth’s AKR. Here, its properties have been extensively reviewed. The AKR has been identified as a narrow-band radiation emitted beneath the local (non-relativistic) electron–cyclotron frequency $f < f_{ce}$. It possesses a very narrow bandwidth, sometimes of the order of $\Delta f/f_{ce} \sim 10^{-4}$ to 10^{-3} . The drifts of these narrow emission bands across a frequency-time diagram reflect the enormous dynamics of the radiation process and its elementary sources of emission. We have identified these sources as electron-holes in phase space which are of the spatial size of just a few Debye-lengths. They can thus be resolved only by investigating the sources in situ. This is possible near the Earth but is impossible for any remote object. Such elementary radiators occur in the presence of magnetic-field-aligned currents and strong magnetic-field-aligned electric fields in a magnetic-mirror geometry.

The first candidates for conditions in favor of the electron–cyclotron maser emission mechanism are the magnetized planets, extrasolar planets, magnetic stars, flare stars, pulsars, and active galactic nuclei or Blazars. Radio emission from some of those objects has been interpreted more or less successfully in terms of the electron–cyclotron maser mechanism. Probably one of the most interesting results of these applications has been that the loss-cone maser mechanism plays an insignificant role only. In the objects discussed in this review the ring-shell maser is by far the more efficient and more interesting emission process. In addition it is the only one that allows for quantitative estimates. Such estimates for the different objects that have been considered here yield volume emission efficiencies — in terms of the plasma kinetic energy density — ranging from several percent in the magnetized planets of our solar system to 10^{-10} in Blazar jets.

Somewhat surprisingly, no radio emission has so far been identified from the newly discovered hot Jupiters. This may suggest that they are very different in their magnetic behavior from the planets known in the solar system.

The range spanned by the emissivities is very wide, indeed. Even though the emitted radiation is coherent and very intense in absolute terms, it represents only a fraction of the energy losses in the medium. The radiation itself does, therefore, not play an important role in the dynamics of most systems under consideration — with the exception of the auroral processes in magnetized planets where it provides a direct and not insignificant loss of energy of the aurora to free space. In the other objects this loss is much smaller. For instance, in Blazar jets the main loss process is related to the jet itself. The radiation

emitted in radio waves appears as a by-product of the dynamic processes inside the jet which can be used with success to diagnose the state of the plasma in the emitting region. Any back-reaction on the plasma can be safely ignored. This applies also to the more intense emitters like the Earth's aurora. Here the electron–cyclotron maser is not responsible for the depletion of the distribution function nor is it responsible for degrading the electric potential.

The potential is degraded, because electrons are accelerated in the electric potential drop and the depletion of the distribution function is the work of very low frequency waves that are excited in the dilute plasma background by the same positive gradient in velocity space that is the source of the electron–cyclotron maser radiation. That the radiation is ineffective in distorting the distribution function is a result of its nature which causes it to disappear at signal speed from the source region. In contrast, the low-frequency plasma waves propagate at much slower velocity and stay in resonance with the energetic electrons for long enough time to be capable of affecting the electron distribution by quasi-linear diffusion. The proof of this conjecture can be found in the observations which show that a broad plateau is generated by the low-frequency waves on the horseshoe distribution function which maintains only a modest phase-space gradient that is just sufficient to keep the electron–cyclotron maser working.

We therefore conclude that the electron–cyclotron maser mechanism provides modest energy losses only for the magnetized planets and negligible energy losses for more violent objects. On the other hand, it not only explains the coherent, highly polarized, highly time- and frequency-variable radio emissions from many sources, but also serves as a valuable diagnostic tool for the physical properties in those sources.

The most interesting of these properties is that in the presence of the electron–cyclotron maser mechanism we are dealing with plasmas that are capable of generating strong magnetic field-aligned electric-potential drops. Such electric potential drops cause strong magnetic-field-aligned electric fields which accelerate electrically charged particles — electrons, positrons, and ions — to the energies corresponding to the full electric potential drop along the magnetic field-line. Such objects are of considerable interest as sources of energetic charged particles released to their environment. They can be identified through the coherent radio-emission generated by the electron–cyclotron maser mechanism.

Acknowledgments The author thanks M.C.E. Huber for suggesting this research, for many discussions, and in particular for his indispensable help in editing the paper. He also thanks M. André (IRF Uppsala, Sweden), R. Ergun (Boulder, CO, USA), C. Jaroschek (Tokyo University, Tokyo, Japan), J. LaBelle and K. Lynch (Dartmouth College, Hanover, NH, USA), H. Lesch (Munich University, Munich, Germany), S. Matsukiyo (Tohoku University, Kyushu, USA), and R. Pottellette (CETP/CNRS St. Maur des Fossés, France) for discussions on the present subject. Preparation of this review has been part of a Senior Visiting Scientist programme at the International Space Science Institute, Bern. The support of the ISSI staff and the ISSI directors, R.-M. Bonnet, R. von Steiger, and A. Balogh, is thankfully acknowledged. This work has benefitted from a Gay-Lussac-Humboldt Prize awarded by the French Government, Direction des Relations Internationales et de la Coopération, Paris. Awarding this prize is gratefully acknowledged.

References

- Alexander JK and Kaiser ML (1976) *J Geophys Res* 81:5948
- Alfvén H (1958) *Tellus* 10:104
- Aschwanden MJ (1990a) *Astron Astrophys* 237:512
- Aschwanden MJ (1990b) *Astron Astrophys Suppl Ser* 85:1141
- Aschwanden MJ and Benz AO (1988a) *Astrophys J* 332:447
- Aschwanden MJ and Benz AO (1988b) *Astrophys J* 332:466
- Bahnse AB, Jespersen M, Ungstrup E, Iversen IB (1987) *Geophys Res Lett* 14:471
- Bastian TS, Benz AO, Gary DE (1998) *Ann Rev Astron Astrophys* 36:131
- Bastian TS, Dulk GA, Leblanc Y (2000) *Astrophys J* 545:1058
- Baumback MM, Calvert W (1987) *Geophys Res Lett* 14:119
- Baumjohann W, Treumann RA (1996) *Basic space plasma physics*. Imperial College Press, London, p 225
- Begelman MC, Rees MJ, Sikora M (1994) *Astrophys J* 429:L57
- Begelman MC, Ergun RE, Rees MJ (2005) *Astrophys J* 625:51
- Bekefi G (1966) *Radiation processes in plasmas*. Wiley, New York, pp 202
- Bekefi G, Hirshfield JL, Brown SC (1961) *Phys Rev* 122:1037
- Benediktov EA, Getmantsev GG, Sazonov YA, Tarasov AF (1965) *Sov Phys Cosmic Res* 3:492 (Kosm. Issled. 3:614)
- Benson RF (1982) *Geophys Res Lett* 9:1120
- Benson RF (1985) *J Geophys Res* 90:2753
- Benson RF, Calvert W (1979) *Geophys Res Lett* 6:479
- Benz AO (1986) *Sol Phys* 104:99
- Benz AO, Conway J, Güdel M (1998) *Astron Astrophys* 331:596
- Bingham R, Cairns RA (2000) *Phys Plasmas* 7:3089
- Bingham R, Cairns RA, Kellett BJ (2001), *Astron Astrophys* 370:1000
- Bingham R et al (2003) *Astrophys J* 595:279
- Block LP (1972) *Cosmic Electrodyn* 3:349
- Block LP (1977) *Astrophys Space Sci* 55:59
- Boström R et al (1998) *Phys Rev Lett* 61:82
- Brown LW (1973) *Astrophys J* 180:359
- Budden KG (1988) *Propagation of radio waves: the theory of radio waves of low power in the ionosphere and magnetosphere*. Cambridge University Press. Cambridge
- Calvert W (1981a) *Geophys Res Lett* 8:919
- Calvert W (1981b) *J Geophys Res* 86:76
- Calvert W (1982) *J Geophys Res* 87:8199
- Calvert W (1987) *J Geophys Res* 92:8792
- Calvert W (1995) *J Geophys Res* 100:14887
- Carlson CW et al (1998) *Geophys Res Lett* 25:2017
- Cattell C et al (2005) *J Geophys Res* 110:A01211. DOI 10.1029/2004JA 010519
- Chiu YT and Schulz M (1978) *J Geophys Res* 83:629
- Chu KR (2004) *Rev Mod Phys* 76:489
- Chu KR and Hirshfield JL (1978) *Phys Fluids* 21:461
- Delory GT et al. (1998) *Geophys Res Lett* 25:2069
- Dory RA, Guest GE, Harris EG (1965) *Phys Rev Lett* 14:131
- Dulk GA (1985) *Ann Rev Astron Astrophys* 23:169
- Dunckel N, Ficklin B, Rorden L, Helliwell RA (1970) *J Geophys Res* 75:1854
- Elphic RL et al (1998) *Geophys Res Lett* 25:2033
- Ergun RE et al (1993) *J Geophys Res* 98:3777
- Ergun RE et al (1998a) *Geophys Res Lett* 25:2061
- Ergun RE et al (1998b) *Geophys Res Lett* 25:2025
- Ergun RE et al (1998c) *Astrophys J* 503:435
- Ergun RE et al (2000) *Astrophys J* 538:456
- Ergun RE et al (2001a) *Phys Rev Lett* 87:045003
- Ergun RE et al (2001b) *Geophys Res Lett* 28:3805
- Escoubet P, Schmidt R, Goldstein ML (1997) *Space Sci Rev* 79:11. DOI 10.1023/A:1004923124586

- Farrell WM, Desch MD, Zarka P (1999) *J Geophys Res* 104:14025
- Farrell WM et al (2004) *Planet Space Sci* 52:1469
- de Feraudy H and Schreiber R (1995) *Geophys Res Lett* 22:2973
- Fleishman GD (2006) Generation of emissions by fast particles in stochastic media, In: LaBelle JW and Treumann RA (eds) *Geospace electromagnetic waves and radiation*. vol 687. Springer, Berlin Heidelberg New York, p 85
- Fleishman GD, Gary DE, Nita GM (2003) *Astrophys J* 593:571
- Föppl et al (1968) *J Geophys Res* 73:21
- Freund HP, Wong HK, Wu CS, Xu MJ (1983) *Phys Fluids* 26:2263
- Gaponov AV (1959a) *Izv VUZ Radiofizika* 2:450
- Gaponov AV (1959b) *Izv VUZ Radiofizika* 2:836
- Goldman MV, Oppenheim MM, Newman DL (1999) *Geophys Res Lett* 26:1821
- Goldman MV, Newman DL, Ergun RE (2003) *Nonlinear Process Geophys* 10:37
- Green JL, Gurnett DA, Shawhan SD (1977) *J Geophys Res* 82:1825
- Gurnett DA (1974) *J Geophys Res* 79:4227
- Gurnett DA (1975) *J Geophys Res* 80:2751
- Gurnett DA and Anderson RR (1981) The kilometric radio emission spectrum: relationship to auroral acceleration processes, In: Akasofu SI and Kan JR (eds) *The physics of auroral arc formation*. American Geophysical Union, Washington, p 341
- Gurnett DA and Green JL (1978) *J Geophys Res* 83:689
- Haerendel G et al (1976) First observation of electrostatic acceleration of barium ions into the magnetosphere, In: ESA European programmes on sounding-rocket and balloon research in the auroral zone, ESA, Paris, p 203
- Harris EG (1959) *J Nucl Energy C2*:138
- Hewitt RG, Melrose DB, Rönnmark KG (1982) *Aust J Phys* 35:447
- Hilgers A (1992) *Geophys Res Lett* 19:237
- Hirshfield JL and Bekefi G (1963) *Nature* 198:20
- Hirshfield JL and Wachtel JM (1964) *Phys Rev Lett* 12:533
- Hoensbroech A and Xilouris KM (1997) *Astron Astrophys* 324:981
- Hull AJ et al (2003) *J Geophys Res* 108: 1007. DOI 10.1029/2001JA007540
- Hultqvist B (1990) *J Geophys Res* 95:5749
- Jackson JD (1962) *Classical electrodynamics*. Academic New York
- Jaroschek CH, Treumann RA, Lesch H, Scholer M (2004a) *Phys Plasmas* 11: 1151. DOI 10.1063/1.1644814
- Jaroschek CH, Lesch H, Treumann RA (2004b) *Astrophys J* 605:L9
- Jaroschek CH, Lesch H, Treumann RA (2004c) *Astrophys J* 616:1065
- Jaroschek CH, Lesch H, Treumann RA (2005) *Astrophys J* 618:822
- Jones D (1976) *Nature* 260:686
- Kazbegi AZ, Machabeli GZ, Melikidze GI (1991) *Monthly Not Royal Astron Soc* 253:377
- Kennel CF and Petschek HE (1966) *J Geophys Res* 61:1
- Kho TH and Lin AT (1988) *Phys Rev A* 38:2883
- Knight S (1973) *Planet Space Sci* (1973) 21:741
- Kuijpers J (1985) In: *Radio Stars*, Hjellming RM and Gibson DM (eds.), Reidel, Dordrecht, p 3
- Kurth WS, Baumback MM, Gurnett DA (1975) *J Geophys Res* 80:2764
- Kurth WS et al (2005a) *Nature* 433:722
- Kurth WS et al (2005b) *Geophys Res Lett* 32:L20S07. DOI: 10.1029/2005GL 022648
- LaBelle J and Treumann RA (2002) *Space Sci Rev* 101:295
- Lang KR (1994) *Astrophys J Suppl* 90:753
- Lang KR, Bookbinder J, Golub L, Davis MM (1983) *Astrophys J* 272:L15
- Lau YY and Chu KR (1983) *Phys Rev Lett* 50:243
- Le Quéau D (1988) *Comput Phys Commun* 49:85
- Le Quéau D, Pellat R, Roux A (1984b) *J Geophys Res* 89:2831
- Le Quéau D, Pellat R, Roux A (1984a) *Phys Fluids* 27:247
- Leung P, Wong AY, Quon BH (1980) *Phys Fluids* 23:992
- Louarn P (2006) Generation of auroral kilometric radiation in bounded source regions. In: LaBelle JW and Treumann RA (eds) *Geospace electromagnetic waves and radiation*, LNP 687. Springer, Berlin Heidelberg New York, p 53

- Louarn P et al (1990) *J Geophys Res* 95:5983
- Louarn P and Le Quéau D (1996a) *Planet Space Sci* 44:199
- Louarn P and Le Quéau D (1996b) *Planet Space Sci* 44:211
- Louarn P, Le Quéau D, Roux A (1986) *Astron Astrophys* 165:211
- Lyons RL (1980) *J Geophys Res* 85:17
- Ma CY, Mao CY, Wang DY, Wu XJ (1998) *Astrophys Space Sci* 257:201
- Machabeli GZ and Usov VV (1979a) *Astron Lett* 5:238
- Machabeli GZ and Usov VV (1979b) *Astron Lett* 15:393
- Manchester RN and Taylor JH (1977) *Pulsars*. Freeman, San Francisco
- McLean DJ and Labrum NR (eds) (1985) *Solar Radiophysics*, Cambridge University Press, Cambridge, pp 211
- Mellott MM et al (1984) *Geophys Res Lett* 11:1188
- Mellott MM, Huff RL, Gurnett DA (1985) *Geophys Res Lett* 12:479
- Melrose DB (1973) *Aust J Phys* 26:229
- Melrose DB (1976) *Astrophys J* 207:651
- Melrose DB (1980) *Plasma astrophysics: nonthermal processes in diffuse magnetized plasmas*, vol 1. Gordon and Breach, New York
- Melrose DB (1999) *Astrophys Space Sci* 264:391
- Melrose DB (2002) *Publ Astron Soc Aust* 19:34
- Melrose DB, Rönmark KG, Hewitt RG (1982) *J Geophys Res* 87:5140
- Melrose DB, Dulk GA, Hewitt RG (1984) *J Geophys Res* 89:897
- Melrose DB and Dulk GA (1982) *Astrophys J* 259:844
- Migenes V and Reid MJ (eds) (2002) *Cosmic masers: from protostars to blackholes*, IAU Symposium no. 206, IAU, Sheridan Books Chelsea
- Mitchell DG et al (2005) *Geophys Res Lett* 32: L20S01. DOI 10.1029/2005GL 022647
- Montgomery DC and Tidman DA (1964) *Plasma kinetic theory*. McGraw-Hill, New York
- Muschietti L, Ergun RE, Roth I, Carlson CW (1999) *Geophys Res Lett* 26:1093
- Muschietti L, Roth I, Carlson CW, Ergun RE (2000) *Phys Rev Lett* 85:94
- Muschietti L, Roth I, Carlson CW, Berthomier M (2002) *Nonlinear Process Geophys* 9:101
- Mutel RL et al (2006) *J Geophys Res* 111 (in press)
- Newman DL, Goldman MV, Ergun RE, Mangeney A (2001) *Phys Rev Lett* 87:255001
- Newman DL et al (2004) *Comput Phys Commun* 164:122
- Omidi N and Gurnett DA (1982a) *J Geophys Res* 87:2241
- Omidi N and Gurnett DA (1982b) *J Geophys Res* 87:2377
- Omidi N and Gurnett DA (1984) *J Geophys Res* 89:10801
- Oppenheim M, Newman DL, Goldman MV (1999) *Phys Rev Lett* 83:2344
- Oppenheim M et al (2001) *Geophys Res Lett* 28:1891
- Pfaff R et al (2001) *Space Sci Rev* 98:1
- Pottelette R and Treumann RA (2005) *Geophys Res Lett* 32:L12104. DOI 10.1029/2005GL022547
- Pottelette R and Treumann RA (2006) *Auroral acceleration and radiation*, In: LaBelle JW and Treumann RA (eds) *Geospace electromagnetic waves and radiation LNP 687*. Springer, Berlin Heidelberg New York, p 103
- Pottelette R, Treumann RA, Dubouloz N (1992) *J Geophys Res* 97:12029
- Pottelette R et al (1999) *Geophys Res Lett* 26:2629
- Pottelette R, Treumann RA, Berthomier M (2001) *J Geophys Res* 106:8465
- Pottelette R, Treumann RA, Berthomier, M, Jasperse J (2003) *Nonlinear Process Geophys* 10:87
- Pottelette R, Treumann RA, Georgescu E (2004) *Nonlinear Process Geophys* 11:197
- Pritchett PL (1984a) *Geophys Res Lett* 11:143
- Pritchett PL (1984b) *J Geophys Res* 89:8957
- Pritchett PL (1986a) *Phys Fluids* 29:2919
- Pritchett PL (1986b) *J Geophys Res* 91:13569
- Pritchett PL and Strangeway RJ (1985) *J Geophys Res* 90:9650
- Pritchett PL et al (1999) *J Geophys Res* 104:10317
- Pritchett PL, Strangeway RJ, Ergun RE, Carlson CW (2002) *J Geophys Res* 107:1437. DOI 10.1029/2002JA009403
- Rybicki GB and Lightman AP (1979) *Radiative processes in astrophysics*. Wiley, New York, p 167
- Saeki K, Michelsen P, Pécseli H, Rasmussen JJ (1979) *Phys Rev Lett* 42:501

- Sagdeev RZ and Shafranov VC (1960) *Sov Phys JETP* 12:130
- Schamel H (1979) *Phys Scr* 20:336
- Schamel H (1986) *Phys Rep* 140:161
- Schneider J (1959) *Phys Rev Lett* 7:959
- Singh N et al (2006) *Nonlinear Process Geophys* 12: (in press)
- Smith K et al (2003) *Astron Astrophys* 406:957
- Stepanov AV et al (2001) *Astron Astrophys* 374:1072
- Temerin M, Cerny K, Lotko W, Mozer FS (1982) *Phys Rev Lett* 48:1175
- Treumann RA and Baumjohann W (1997) *Advanced space plasma physics*. Imperial College Press, London, p 120
- Twiss RO (1958) *Aust J Phys* 11:564
- Vedin J and Rönmark K (2005) *J Geophys Res* 110:A08207. DOI 10.1029/2005JA011083
- Weimer DR et al (1985) *J Geophys Res* 90:7479
- White SM, Melrose DB, and Dulk GA (1983) *Proc Astron Soc Aust* 5:188
- Willson RF (1985) *Sol Phys* 96:199
- Winglee RM (1983) *Plasma Phys* 25:217
- Winglee RM and Dulk GA (1986) *Astrophys J* 310:432
- Winglee RW and Pritchett PL (1986) *J Geophys Res* 91:13531
- Winterhalter D et al (2005) Search for radio emissions from extrasolar planets at 150 MHz, American Geophysical Union, Fall Meeting 2005, abstract no. SA53B-1181
- Wu CS (1985) *Space Sci Rev* 41:215
- Wu CS and Lee LC (1979) *Astrophys J* 230:621
- Zarka P (1992a) *Adv Space Res* 12:99
- Zarka P (1992b) Remote probing of auroral plasmas, In: Rucker HO et al (eds), *Planetary radio Emissions III*. Australian Academy of Science Vienna, p 351
- Zarka P (1998) *J Geophys Res* 103:20159
- Zarka P (2004) *Adv Space Res* 33:2045. DOI 10.1016/j.asr.2003.07.055
- Zarka P, Le Quéau D, Genova F (1986) *J Geophys Res* 91:13542
- Zarka P, Queinnee J, Cray FJ (2001a) *Planet Space Sci* 49:1137
- Zarka P, Treumann RA, Ryabov BP, Ryabov VB (2001b) *Astrophys Space Sci* 277:293
- Zarka P, Cecconi B, Kurth WS (2004) *J Geophys Res* 109:A09S15. DOI 10.1029/2003JA010260
- Zarka P, Hess S, Mottez F (2005) Io–Jupiter interaction, millisecond bursts and field aligned potentials, American Geophysical Union, Fall Meeting 2005, abstract no. SM51A-1275