Astron. Astrophys. 339, 298–308 (1998)

ASTRONOMY AND ASTROPHYSICS

The electron cyclotron maser in hot thermal plasmas

A.J. Conway[1] **and A.L. MacKinnon**[2]

[1] Department of Physics & Astronomy, University of Glasgow, Glasgow, UK (conway@astro.gla.ac.uk)
[2] Department of Continuing and Adult Education, University of Glasgow, Glasgow, UK (alec@astro.gla.ac.uk)

Received 27 July 1998 / Accepted 18 August 1998

Abstract. We investigate the possibility of an electron cyclotron maser operating in a hot ($T > 10^7$K) thermal plasma with a loss cone. We find that the instability can occur in this scenario because resonance involving electrons in the thermal tail can give rise to significantly large growth rates. We estimate the range of electron velocities that can be in resonance and show how this dictates the frequency and angular distribution of the emitted radiation. In so doing, we show quite generally that it is the range of speeds and not the details of the loss cone edge, as previously thought, that determines the frequency and angular distribution. By considering the energetic constraints on the initial and final (i.e. when the maser is saturated) electron distributions, we estimate bounds on the energy that may be emitted. This energy can be a sizeable fraction (up to 5-10%) of the plasma electrons' total thermal energy, though this is sensitively dependent on the temperature, density and magnetic field. Finally, we discuss the implications of this emission for flaring solar coronal loops and hot stellar atmospheres.

Key words: instabilities – radiation mechanisms: thermal – Sun: corona – Sun: flares – Sun: radio radiation

1. Introduction

The electron cyclotron maser (ECM) is a plasma instability caused by the resonance between electromagnetic waves and plasma electrons in the presence of an externally generated magnetic field. However, the resonance only leads to the instability if the velocity distribution of the electrons has a particular kind of anisotropy. An important milestone in the understanding of the ECM was Wu and Lee (1979)'s description of auroral kilometric radiation, in which they used the semi-relativistic approximation (see Sect. 2), demonstrating that masers could operate in relatively common astrophysical conditions - e.g. in a loss cone distribution. This discovery led to a growth of interest in the ECM, particularly in the context of micro-wave radio emission from flare processes in the solar corona, initiated by the work of Melrose and Dulk (1982). Later, interest in the ECM as an emission mechanism began to wane as it was realised (Sharma and Vlahos 1984) that very contrived conditions were required for the emission to avoid absorption in the solar corona, and thus escape and be directly observed. However, precisely because maser radiation is so readily absorbed, it was also recognised as a potential mechanism for transferring plasma energy across the magnetic field - something which charged particles cannot do directly in strongly magnetised regions, such as the solar corona. In particular, Melrose and Dulk (1984) proposed it as a plausible mechanism for heating solar flare loops from their ambient coronal temperature of a few 10^6K to typical flare temperatures of a few 10^7K. Also, Sprangle and Vlahos (1983) proposed that solar type III bursts could be explained by electrons being accelerated by ECM radiation in the vicinity of flaring loops. In all of these applications, the plasma was modelled as having two (or even three) components. Most usually, a cold ($T \sim 10^6$K) thermal component was assumed with the ECM operating in a separate component of fast (i.e. much greater than thermal speed) electrons that had developed a loss cone. In this scenario, the cold component dictates the details of wave propagation by determining the plasma frequency, and also causes thermal damping of the ECM. Aschwanden (1990) represents the most detailed treatment of this kind, numerically mapping out the vast parameter space by calculating growth rates, saturation times and released energies for wide ranges of plasma properties.

In this paper, we consider a one component thermal plasma and show that, if hot enough, it can develop a loss cone in its tail and release significant amounts of energy by the ECM mechanism. This not only has relevance to the solar flare context discussed above, but also to many other astrophysical contexts where hot thermal plasmas are known to exist. There are two distinct gains from consideration of the ECM mechanism in this way. Firstly, no accelerated non-thermal distribution of particles needs to be explained - hot coronal loop plasmas are an observational fact. Secondly, it reduces the parameter freedom of the model. A two component model essentially requires two densities and two temperatures - the one component model's relative simplicity allows a neater analytic description of the process. Given this simplification, we have also endeavoured to provide useful formulæ and to clarify the many conflicting conditions that have to be satisfied for the ECM to operate. We pay particular attention to the frequencies and angles at which large growth

Send offprint requests to: A.J. Conway

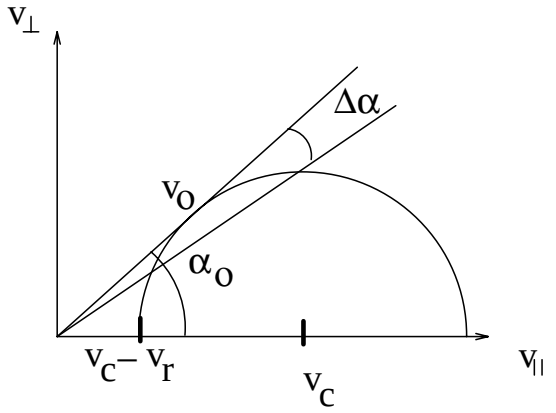

Fig. 1. The resonance circle is placed so it is just touching the edge of the loss cone. Here, the loss cone extends out to the line at angle α_0 to the $v_\parallel$ axis. The $\partial f/\partial p_\perp$ derivative is positive only in a region of angular width $\Delta\alpha$. For $\alpha < \alpha_0 - \Delta\alpha$, the distribution function (and all its derivatives) is zero.

is possible and general limits on the total energy that may be released.

Although our aim in this paper is to provide a general description of the ECM operating in a magnetised thermal plasma, our original motivation for the work was to investigate whether the ECM could operate in solar coronal loops at flare temperatures. For this reason, we use such terms as "hot", "tenuous", "strongly magnetised" etc. with reference to typical coronal loop properties: $n \sim 10^{16}\mathrm{m}^{-3}$, $B \sim 0.1$ Tesla (1000 Gauss) and $T \sim 10^7$K (note we use SI units throughout). A typical coronal flare loop can be thought of as a hot (and therefore collisionless on maser time scales) plasma following an arched magnetic field which is strongest (strength B_1) where it descends into the solar chromosphere and photosphere. It is this strengthening of the field at the "ends" (also referred to as "feet" or "mirror points") of the loop that is responsible for trapping particles with larger pitch angles ($\alpha > \alpha_0 = \sin^{-1}\sqrt{B/B_1}$). (We use the conventional definition of pitch angle as the angle between a particle's velocity vector and the magnetic field direction.) The absence of particles with pitch angle smaller than α_0 gives rise to the notion of the velocity space loss cone. As particles are effectively tied to field lines on the timescales of interest, the actual shape of the "loop" is not important.

In Sect. 2 we outline the relevant ECM theory and analyse the restrictions on the growth rates of the instability in a thermal plasma. In Sect. 3 we place upper and lower limits on the energy that can be released on the grounds of momentum and energy conservation. In Sect. 4 we draw our conclusions from the results, summarising the various conditions required for large growth and the assumptions which restrict our treatment of the problem. Finally we relate our findings to astrophysical situations in which thermal masering can occur.

2. The electron cyclotron maser

2.1. Electron cyclotron maser theory

A wave of frequency ω and wavenumber k, travelling at an angle θ to the magnetic field is in resonance with electrons of speed v and pitch angle α, if

$$\omega - s\Omega(v) - (k\cos\theta)(v\cos\alpha) = 0 \tag{1}$$

where s is the harmonic number (any integer), $\Omega = \Omega_0/\gamma_r = eB/m\gamma_r$ is the gyro-frequency (m,e being the electron mass and charge and B being the magnetic field strength) and $\gamma_r = (1 - (v/c)^2)^{-1/2}$ (c being the speed of light). For the purposes of this paper, we can safely use the weak-relativistic approximation, $\gamma_r \sim (1 - v^2/2c^2)^{-1}$, because at the temperatures of interest to us ($\sim 10^7$K to $\sim 10^8$K) "fully relativistic" electrons are very small in number. With this approximation, the resonance condition can be pictured as a circle in the $(v_\parallel = v\cos\alpha, v_\perp = v\sin\alpha)$ plane, centered on $(v_c, 0)$ with radius v_r given by:

$$v_c = \frac{N\cos\theta}{sY}c \qquad v_r^2 = v_c^2 - 2\frac{1 - sY}{sY}c^2 = v_c^2 - v_0^2 \tag{2}$$

where $N = kc/\omega$ is the refractive index, $Y = \Omega_0/\omega$ and v_0 is the electron speed. Whether or not the resonance leads to an instability can be determined by calculating the growth rate γ, defined as $\gamma = \dot{E}/2E$ where E is the energy of the wave. The growth rate γ can be thought of as the sum of contributions from a series of harmonics: $\gamma = \sum_s \gamma_s$. The contribution to the growth rate due to harmonic s can be expressed as (Melrose 1986 Eq. 11.20 - note Melrose uses negative γ to mean growth)

$$\gamma_s(\omega,\theta) = \frac{2\pi e^2}{\epsilon_0\omega}F(\omega,\theta)\int\left(\frac{s\Omega}{v_\perp}\frac{\partial f}{\partial p_\perp} + k_\parallel\frac{\partial f}{\partial p_\parallel}\right)\delta(\omega - s\frac{\Omega_0}{\gamma_r} - k_\parallel v_\parallel)v_\perp^2 d^3\mathbf{p} \tag{3}$$

where $\mathbf{p}$ is the momentum vector and the symbols $\perp$ and $\parallel$ represent perpendicular and parallel components with respect to the magnetic field, respectively. δ denotes the Dirac delta function. The wave mode under consideration determines the functional form of the dimensionless function F and the refractive index N. The distribution function $f(\mathbf{p})$ is defined to be number of particles per unit spatial volume, per unit momentum volume. As the energy emitted at the harmonic s is proportional to $(v/c)^{2s}$, v being the typical speed of resonant particles, we will be primarily concerned with emission at $s = 1$. From here on $s = 1$ is assumed.

2.2. Growth rate limits for a loss cone

We now consider the distribution of electrons inside a symmetric magnetic trap of transverse size r, at a distance L from a mirror point. If the loss cone is well defined, meaning that the range of velocity space where $\partial f/\partial p_\perp$ is positive is a small region at the loss cone's edge (i.e. $\psi \ll \alpha_0$ in Fig. 1), then the resonance condition can be additionally restricted by the condition that

$$\frac{v_r}{v_c} \sim \sin\alpha_0 \tag{4}$$

This is asking that the resonance circle passes through a region which contributes to positive growth, without passing through any regions which cause damping (i.e. where $\partial f/\partial p_\perp < 0$, note that $\partial f/\partial p_\parallel \leq 0$ is true everywhere for a loss cone distribution). The use of this condition means that the growth rate discussed in the rest of this paper is really a maximum growth rate for a given particle speed v_0 or equivalently for a given ω, or for a given θ. However, the meaning of our results, and for that matter those of previous authors, are not harmed by this as the growth rates can be shown to decay drastically if the resonance curve enters any region of $\partial f/\partial p_\perp < 0$.

We now wish to illustrate the fact that given a particular v_0 in a distribution with a well-defined loss cone, the preceeding conditions define a frequency ω and corresponding angle θ for large growth. Consider a given v_0, which, by its definition (2), corresponds to a particular ω. Condition (4) tells us that for significant growth, an electron must have pitch angle α_0. With ω, v_0 and α all specified in the resonance condition (1), and with ω, k and θ related through the refractive index, a particular value of θ is determined. A specific v_0 therefore defines a particular ω and θ of maximum growth for that speed. However, large growth is also possible around ω and θ because of the finite width of the loss cone edge. This can be used to estimate a small range of ω (and therefore θ) over which significant growth is possible. Such formulæ can be found in Melrose (1986) p.198 and Melrose and Dulk (1982). However, this is just for *one* v_0. For different v_0, maximum growth occurs at different ωs and θs. So, in general, the range of ω and θ is determined both by $\Delta\alpha$ and by the possible range of v_0. However, we now show that exceptional circumstances are required for the "$\Delta\alpha$-broadening" to be important. The two frequency widths are given by

$$(\Delta\omega)_{\Delta\alpha} \sim \Omega_0 \left(\frac{v_0}{c}\right)^2 \tan\alpha_0 \Delta\alpha$$

$$(\Delta\omega)_{v_0} \sim \frac{\Omega_0}{2}\left(\frac{v_U}{c}\right)^2 \left[1 - \left(\frac{v_L}{v_U}\right)^2\right]$$

where v_U and v_L are the upper and lower limits placed on v_0 - what determines them is discussed below. However, the ratio

$$\frac{(\Delta\omega)_{\Delta\alpha}}{(\Delta\omega)_{\Delta v_0}} \sim 2\left(\frac{v_0}{v_U}\right)^2 \left(1 - \left(\frac{v_L}{v_U}\right)^2\right)^{-1} \tan\alpha_0 \Delta\alpha$$

will generally be small since $v_0 < v_U$ and $\Delta\alpha \ll 1$ is needed for reasonable growth. Also, if v_L approaches v_U then v_0 will be close to violating one of the conditions required for large growth. So in conclusion, unless the loss cone angle α_0 is very large, the angular and frequency distribution of ECM radiation is determined by the range of particle velocities at which resonance can lead to significant growth, and not by the width of the loss cone edge $\Delta\alpha$. We now discuss the physical constraints defining v_L and v_U.

The velocity distribution will have a well-defined loss cone (i.e. few particles are scattered into it) at particle speed v_0, if the time it takes for a particle that is just trapped (i.e. it has pitch angle equal to the loss cone angle α_0) to cross the trap is less than the collisional time, $t_c(v_0)$ (Spitzer 1962):

$$\frac{L}{v_0 \cos\alpha_0} < t_c = \frac{2\pi m_e^2 \epsilon_0^2}{e^4 \ln\Lambda} \frac{v_0^3}{n}$$

where n is the density, ϵ_0 is the permittivity of free space and $\ln\Lambda$ is the Coulomb logarithm (~ 20). Since the collision time rises as v_0^3 and the trap crossing time is proportional to v_0^{-1}, this places a *lower* limit on the particle speed at which a well-defined loss cone exists. Inserting constants, and introducing v_{α_0} to represent this lower limit we have

$$\frac{v_{\alpha_0}}{c} \sim 0.14 \left(\frac{n_{16} L_7}{\cos\alpha_0}\right)^{1/4} \tag{5}$$

where $n_{16} = n/(10^{16}\mathrm{m}^{-3})$ and $L_7 = L/(10^7\mathrm{m})$. Since growth in the ECM can be attributed to $\partial f/\partial p_\perp$ being greater than zero, for speeds less than v_{α_0}, leakage into the loss cone by collisions means that $\partial f/\partial p_\perp$ is reduced, in turn reducing the growth rate. So we do not expect significant growth for $v_0 < v_{\alpha_0}$.

In order for the maser mechanism to operate, the distance the radiation travels in a growth time must not exceed the size of the region which is undergoing masering. In general this size will be the smaller of: the size of the trap in the direction determined by angle θ; and the coherence length in the same direction, which will be determined by the variation of magnetic field across the trap. Assuming, for simplicity that the magnetic field strength is constant across the trap, and given that the ECM is usually (though not always, see Figs. 2a–d and 4a–d) in the perpendicular direction to the magnetic field, it is most appropriate to use the transverse trap dimension as the size of the masering region. Therefore, the following condition must be satisfied for particles of speed v_0 to participate in the maser

$$\gamma(v_0) > \frac{c}{r} \tag{6}$$

So, even if (3) gives a growth rate that is positive, the radiation energy density cannot grow at arbitrarily small rates. At small $v_0 \leq v_{\alpha_0}$, where $\partial f/\partial p_\perp$ is reduced due to loss cone filling, one might expect this condition to be broken. Also, at large v_0, where f has decreased and flattened out, once might also expect $\partial f/\partial p_\perp$ to be small enough to violate this condition. Another possible limit comes from collisional damping of the maser instability (Zheleznyakov 1970, Chap. 25). In general, however, because the coherence limit on the growth rate is larger than the collisional frequency, collisional damping is only important in exceptional circumstances, e.g. in very dense hot plasmas.

A third constraint can be placed on v_0 indirectly: N must allow a solution for real θ from the resonance condition (1) and the loss cone condition (4). Specifically, substituting (2) into (4) we have

$$\cos\theta = \frac{\sqrt{2Y(1-Y)}}{N(\omega,\theta)\cos\alpha_0} \tag{7}$$

If N becomes small (e.g. when the frequency is nearing a cutoff) then this equation might not give a real solution for θ. The frequency at which this occurs, w_N say, dictates a limit on the speed, which we will denote by v_N. If the cutoff is approached from above (below) in frequency, then v_N is a lower (upper) limit.

2.3. A loss cone distribution in an otherwise thermal plasma

Inserting the distribution

$$\begin{aligned} f(p,\alpha) &= 0 && \alpha < \alpha_0 - \Delta\alpha \quad (8)\\ f(p,\alpha) &= \left(1+\tfrac{\alpha-\alpha_0}{\Delta\alpha}\right)\tfrac{n}{(\pi P^2)^{3/2}}e^{-p^2/P^2} && \alpha_0-\Delta\alpha<\alpha<\alpha_0\\ f(p,\alpha) &= \tfrac{n}{(\pi P^2)^{3/2}}e^{-p^2/P^2} && \alpha>\alpha_0 \end{aligned}$$

into (3) with $s = 1$ gives

$$\begin{aligned} \gamma(\omega,\theta) \;=\; & \frac{\Omega_0}{16(\Delta\alpha)^{1/2}}\frac{XF(X,Y,\theta)}{Y} \qquad (9)\\ & \left(1-\frac{N\cos\theta}{Y}\frac{v_0}{c}\frac{\sin^2\alpha_0}{\cos\alpha_0}\right)\left(\frac{\pi mc^2}{kT}\right)^{3/2}\\ & e^{-mv_0^2(\omega)/2kT}\frac{v_0(\omega)}{c}\sin^{3/2}2\alpha_0 \end{aligned}$$

where $P^2 = 2mkT$, T being the temperature of the original distribution and k being Boltzmann's constant. $X = ne^2/m_e\epsilon_0\omega^2 = \omega_p^2/\omega^2$ and $Y = \Omega_0/\omega$. Note that v_0 is now expressed as a function of ω, i.e. from (2)

$$\left(\frac{v_0}{c}\right)^2 = 2\frac{1-Y}{Y} = 2\frac{\omega-\Omega_0}{\Omega_0} \qquad (10)$$

In full generality, the dimensionless function F is algebraically quite unwieldy. However, it can be greatly simplified by using the formulæ for either the quasi-circular or quasi-planar approximations, given in Melrose (1986) p.172. Further simplifications arise if $w_p^2 \ll \Omega_0^2$. Usually, however, the variation of F over the narrow frequency range of possible emission does not have much impact on the growth rate, whose behaviour as a function of frequency is dominated by $v_0(\omega)$ in the exponential. This is not true if the frequency is nearing a cutoff, and in particular when $w_p \sim \Omega_0$. In this case F can tend to zero, significantly reducing the growth rate. The effect of F in this regime is discussed later in relation to the growth rate plots.

The analytic result (9) lets us simply discuss the behaviour of the maser. In particular, we can now specify the range of particle speeds, and therefore the range of ω (or equivalently θ) for which the maser instability can grow. The exponential decay of the growth rate with increasing v_0, together with the restriction involving the size of the masering region (6), gives us a maximum for the particle speed, which we shall denote by v_U. Although the growth rate also falls to zero as v_0 becomes much less than the thermal speed, this condition does not set a lower limit on the growth rate in conditions of interest to us because either the collisional or the cutoff lower limits are higher. (6) can only set a lower limit if the magnetic field varies significantly across the trap, or if the plasma is extremely hot, tenuous or strongly magnetised. The lower limit, denoted by v_L, is therefore given by the larger of v_{α_0} (collisional) or v_N (cutoff), as defined in (5) and (7) respectively. To a good approximation when $\omega_p^2 \ll \Omega_0^2$, the latter limit can simply be taken at the frequency at which $N(\omega_N) = 0$. Whether or not the cutoff limit is important depends on both the wave-mode of interest and also on the ambient density and magnetic field. The x- and o-mode magnetoionic waves have lower cutoffs at $\Omega_0/2 + \sqrt{\Omega_0^2/4+\omega_p^2}$ and ω_p respectively. So, for increasing density, or decreasing magnetic field strength the lower cutoff limit v_N increases, and once larger than v_{α_0}, it will have a profound affect on the magnitude of the growth rate. The maximum possible growth will always be near the lower velocity limit (assuming that the thermal velocity is less than the lower limit, which is always true in temperatures of interest here). So from (9), when the cutoff limit applies we find that the maximum possible growth rate varies as

$$\exp\left\{-\frac{mc^2}{2kT}\left(\sqrt{1+\frac{4\omega_p^2}{\Omega_0^2}}+1\right)\right\} \quad \text{and}$$

$$\exp\left\{-\frac{mc^2}{kT}\left(\frac{w_p}{\Omega_0}-1\right)\right\}$$

for the x- and o-modes respectively. So, for the x-mode, it is clear that a large n and a small B can greatly reduce the growth rate. In fact, for small ω_p^2/Ω^2, the largest possible growth rate decreases with an exponent proportional to n/B^2. For the o-mode, the cutoff condition can never apply until w_p has become larger than Ω_0 because the form of (2) requires that $\omega > \Omega_0$. So, it is not until $\omega_p^2/\Omega^2 = 1$ that the condition applies, and even then the exponential decay is not as harsh as for the x-mode. This provides a simple explanation for the results found from the detailed investigations of Sharma and Vlahos (1984) (albeit in a somewhat different, non-thermal scenario). They found that if $w_p/\Omega_0 < 0.35$ then maximum growth is in the x-mode of the wave, whilst for $0.35 < w_p/\Omega_0 < 1$, maximum growth is in the o-mode. For values in excess of unity, they found that growth can only occur in the whistler or electrostatic upper hybrid wave modes. Similar arguments to those just presented also apply to these cases.

Figs. 2a–d and 4a–d show the growth rates as a function of propagation angle θ for the x- and the o-modes respectively. All results are plotted for a single set of parameters ($\alpha_0 = 30°$, $L = 10^7$m, $r = 10^6$m) as their effects can be easily seen from the analytic results discussed earlier. Curves are only plotted over ranges that satisfy the various conditions given earlier in Sect. 2.2. For this reason, some curves are missing because growth could not occur in any range for the given set of parameters. In general, as expected, the higher T and B, and the lower n, the larger the range of angles, and therefore frequencies over which large growth is possible. It is also clear that increasing n and decreasing B have a more dramatic impact on the x-mode than the o-mode. As a rule, the maximum growth in the o-mode ($\sim 10^7\text{s}^{-1}$) is much less than that of the x-mode ($\sim 10^9\text{s}^{-1}$) because the function F is generally of order unity for the x-mode but is several orders of magnitude smaller for the o-mode. In order that the kinetic maser growth rate theory applied here is valid, the growth rates must be smaller than the range of frequencies at which significant growth occurs. Inspection of Fig. 2a–d (and 4a–d), together with the frequency ranges depicted in Fig. 3a and b (and 5a and b) show that this is always the case for the x-mode (and o-mode). However, for the very

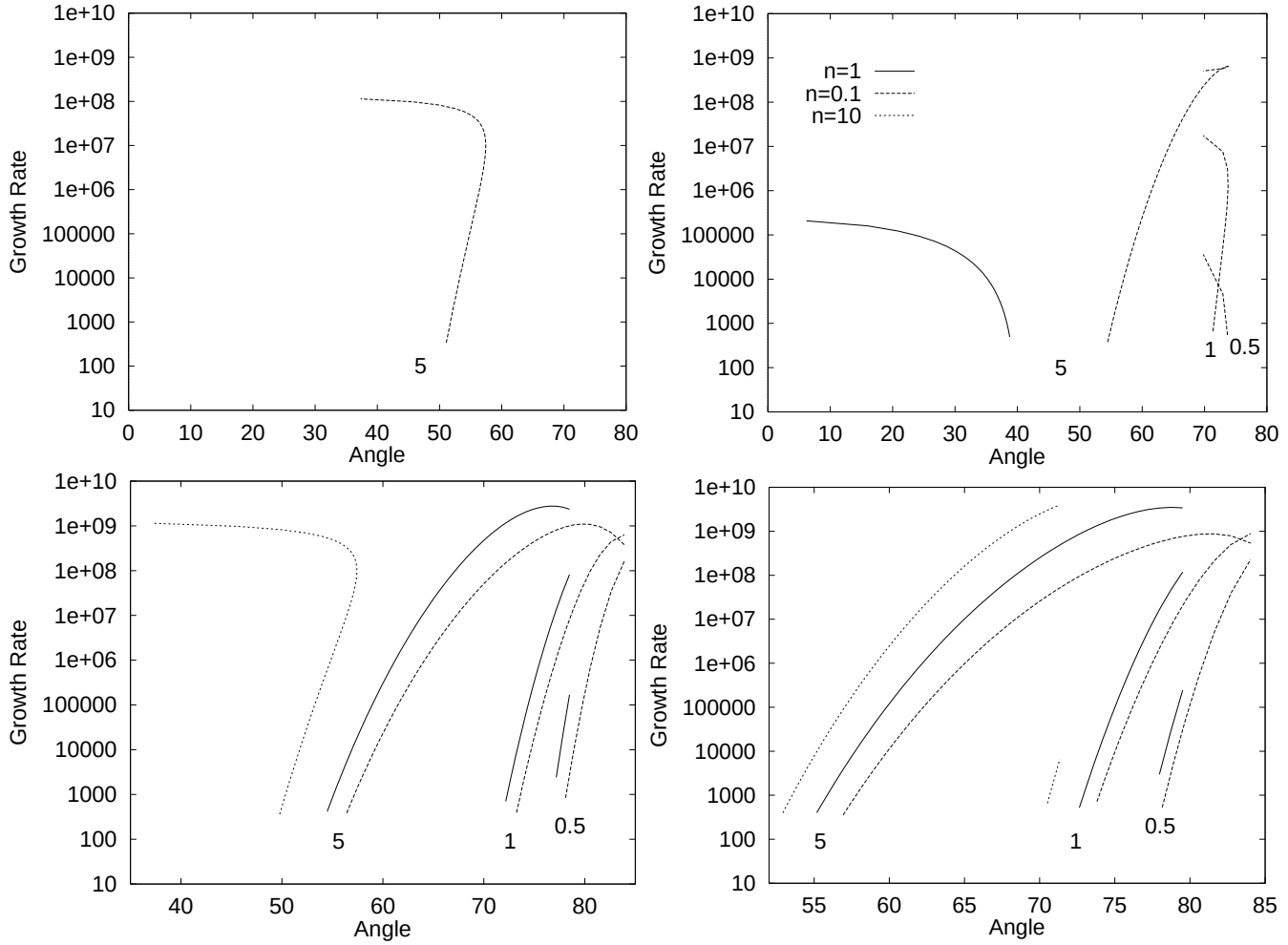

Fig. 2a–d. Plots of the growth rate γ (s^{-1}) for the ECM in the x-mode for different magnetic field strengths (in Tesla), densities (in units of $10^{16}m^{-3}$) and temperatures (in units of 10^7K). **a** (upper left) $B = 0.05$T ($\Omega_0 = 8.79$ GHz), **b** (upper right) $B = 0.1$T ($\Omega_0 = 17.6$GHz), **c** (lower left) $B = 0.5$T ($\Omega_0 = 87.9$ GHz), **d** (lower right) $B = 1$T ($\Omega_0 = 176$GHz). Curves for three temperatures $T_7 = 0.5, 1, 5$ are labeled by the numbers appearing in each graph. Note that higher temperature curves always have their lowest growth rates at a smaller angle. For lower temperatures and magnetic field strengths no growth is possible for the higher density cases, hence no curve is plotted.

highest growth rates in the x-mode, the 'kinetic' approach, with the implicit random phase assumption, may begin to lose its validity, but the results nevertheless suffice to show a significant regime where instability occurs.

All the growth curves start at a growth rate of about 300 because this corresponds to the coherence time - i.e. the lowest angle of each curve therefore represents the upper velocity limit set by (6). This lowest angle only slightly decreases with increasing density, but decreases markedly with increasing temperature. This can be understood physically by examining the plots in Fig. 2a–d. Firstly, from (9) it can be seen that the growth rate is proportional to n (contained in the dimensionless plasma frequency parameter X). This linear dependence simply comes from the fact that the number of resonant electrons is proportional to the number density of electrons. Therefore, one should expect that the curves of the same temperature are vertically separated by a factor roughly equal to the factor by which the density has changed. In the $B = 1$ and $B = 0.5$ plots this holds true for all but the $B = 0.5$, $n_{16} = 10$, $T_7 = 5$ curve, which will be discussed shortly. The temperature dependence obviously comes mainly through the exponential in (9), so that a linear decrease in temperature corresponds to an exponential decrease in electrons at speeds relevant to electron-cyclotron resonance, and thus a similar decrease in the growth rate. Similar arguments apply for the o-mode. The density is also important in determining the high angle cutoff of the growth rate, which generally occurs at lower angles for increasing densities, because the collisional limit sets the lower velocity limit for large B.

The reason that the simple explanations just given apply only to large B is because the frequencies involved are well above the cutoff frequencies. In these cases, the collisional limit in velocity always gives a *higher* lower limit on the frequency than the cutoff frequency. This means that F and N, which appear in (9), are roughly constant, and that θ is related to ω in a rather simple way in (7). Smaller magnetic field strengths (x-mode only)

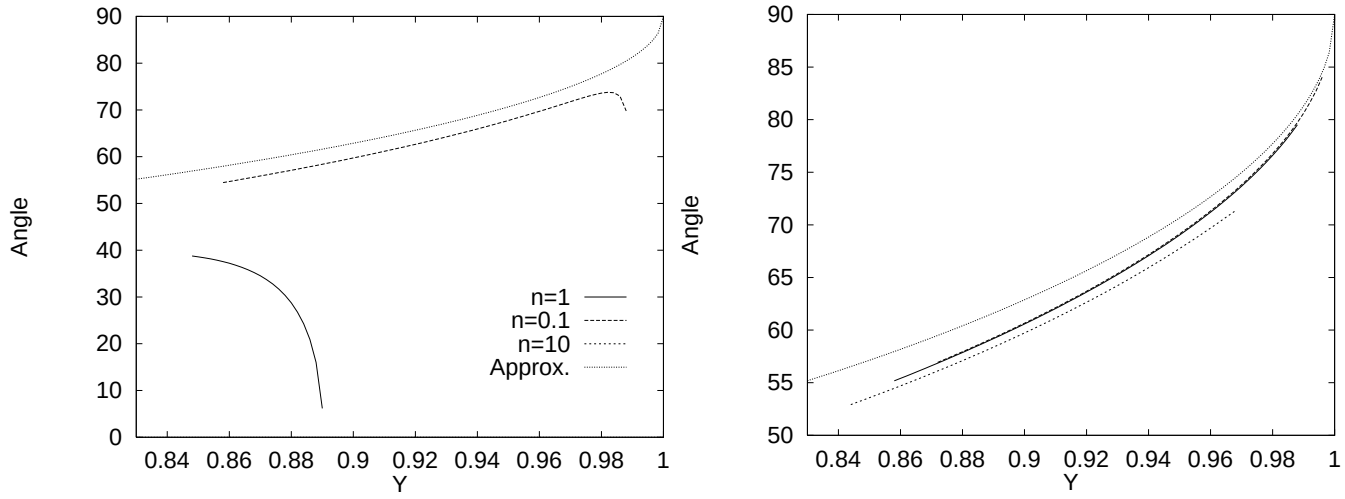

Fig. 3a and b. Plots of $Y = \Omega_0/\omega$ vs. the propagation angle θ for the ECM in the x-mode for two magnetic field strengths (in Tesla) and densities (in units of $10^{16}\mathrm{m}^{-3}$). **a** (left) $B = 0.1\mathrm{T}$ ($\Omega_0 = 17.6\mathrm{GHz}$), **b** (right) $B = 1\mathrm{T}$ ($\Omega_0 = 176\mathrm{GHz}$). As the relationship between frequency and propagation angle does not involve temperature, we only plot curves for $T_7 = 5$. For $B = 0.1$ no growth is possible for the highest density case. Note that the approximate $\theta - \omega$ relation obtained from (7) with $N = 1$ is always an upper limit on θ, and is most accurate when ω_p/Ω_0 is small.

and larger densities increase the cutoff frequencies, drastically reducing F whilst taking N well below unity. In such cases, growth becomes restricted to smaller frequency/angle ranges, and in the smallest B cases, is only possible for the higher temperature/density cases. The "knee" in the x-mode growth rate graphs appears because the small N forces θ to smaller angles, as can be appreciated from (7) and the $\omega - \theta$ plots for $B = 0.1$ in Fig. 3a and b. It is clear from the o-mode growth rate graphs and $\omega - \theta$ plots in Fig. 5a and b that there is no such "knee" for o-mode - this is because the cutoff condition does not depend on B.

3. Energy constraints

We place constraints on the energy released by the ECM by placing upper and lower limits on the energy of final distribution. The minimum amount of energy released is calculated by finding the distribution with the largest total energy that has $\partial f/\partial p_\perp \leq 0$. The maximum energy emitted is calculated by assuming that the final distribution is Maxwellian, with temperature determined by the initial excess of perpendicular energy. In both cases, for simplicity, we assume all the emitted photons are emitted near to perpendicular, i.e. $\theta \sim 90°$. The case of emission far from the perpendicular is discussed qualitatively at the end of this section.

3.1. The minimum released energy

Starting with the initial distribution (8), our task is to find the distribution with the maximum energy that has $\partial f/\partial p_\perp \leq 0$ everywhere, so that the ECM can no longer operate. Strictly speaking the ECM saturates before this, when the filling of the loss cone causes growth at all frequencies to fall below the coherence timescale (6). However, given that that the growth can initially be many orders of magnitude above this limit, the $\partial f/\partial p_\perp \leq 0$ approximation should be accurate in many situations. The other constraint is that particles should only lose energy and, specifically, lose it in the perpendicular direction. Again this approximation is a good one in many situations, as will be discussed later.

The distribution that meets the above conditions is

$$\begin{aligned} f_M(p_\parallel, p_\perp) &= f_i(p_\parallel, p_\perp = p_1) \quad p_\perp < p_1 \\ f_M(p_\parallel, p_\perp) &= f_i(p_\parallel, p_\perp) \qquad p_\perp \geq p_1 \end{aligned} \tag{11}$$

where f_i is the distribution function given in (8) and $p_1 > p_0 = p_\parallel \tan\alpha_0$ is given by

$$e^{(p_1^2 - p_0^2)/P^2} = \frac{p_1^2}{P^2}$$

A proof of this is given in the Appendix. Subtracting the final from the initial distribution and integrating up the energy over the range of momentum space where the ECM can grow gives the minimum released maser energy as

$$(\Delta E)_{\min} = \frac{nkT}{4\pi^{1/2}} \int_{u_L}^{u_U} e^{-(u_\parallel + u_1)} (2u_0(1+u_1) - u_1^2) \, u_\parallel^{-1/2} du_\parallel$$

where $u = p^2/P^2$ and the L and the U subscripts refer to the limits discussed earlier in Sect. 2.2.

3.2. Maximum released energy

As the maser will always saturate before a Maxwellian state is reached, we can use this end state to calculate an upper limit on the energy emitted by the maser. Again, assuming that all the energy is emitted in the perpendicular direction, we can specify

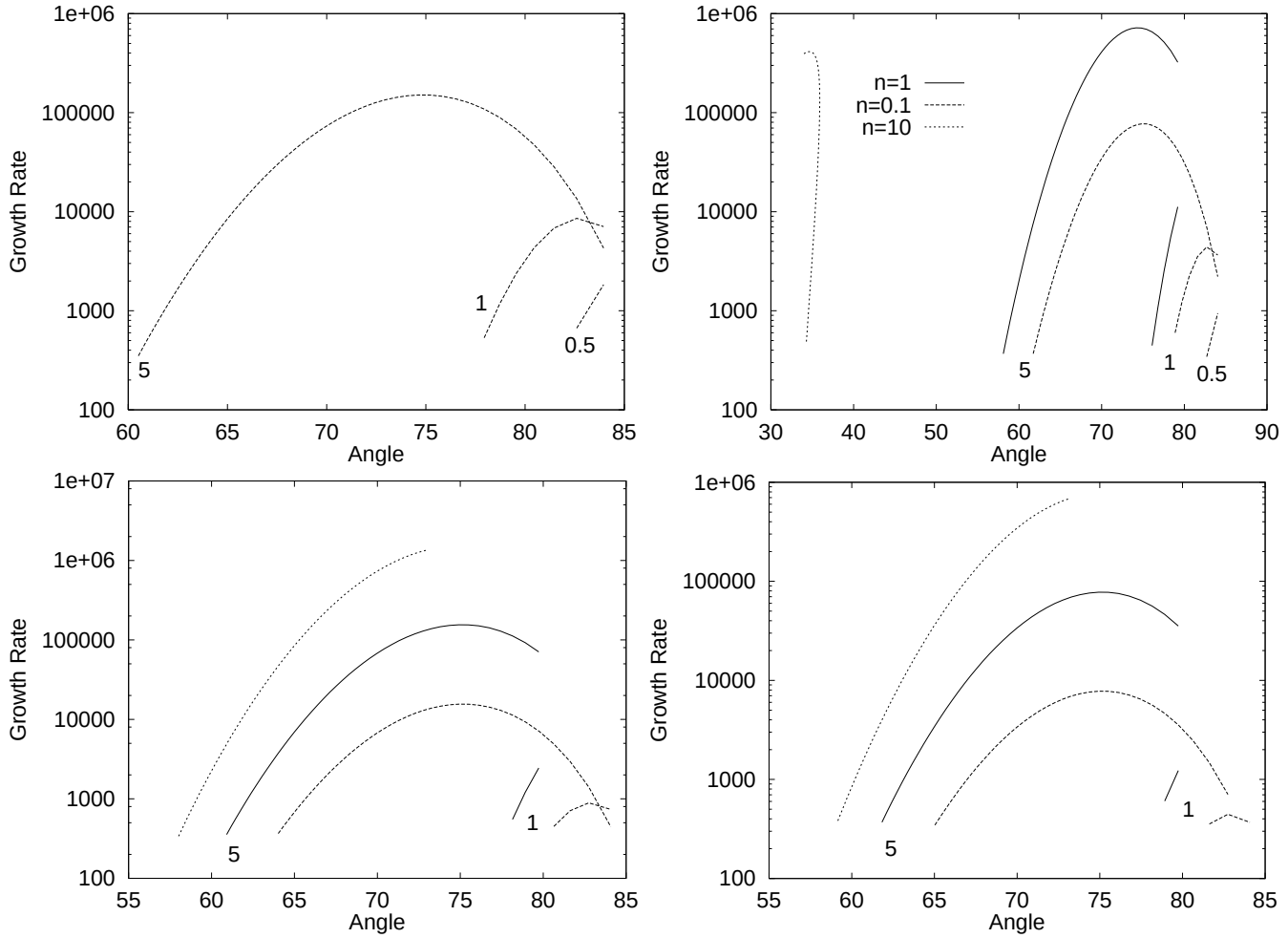

Fig. 4a–d. Plots of the growth rate γ (s^{-1}) for the ECM in the o-mode for different magnetic field strengths (in Tesla) densities (in units of $10^{16}m^{-3}$) and temperatures (in units of 10^7K). **a** (upper left) $B = 0.05$T ($\Omega_0 = 8.79$ GHz), **b** (upper right) $B = 0.1$T ($\Omega_0 = 17.6$GHz), **c** (lower left) $B = 0.5$T ($\Omega_0 = 87.9$ GHz), **d** (lower right) $B = 1$T ($\Omega_0 = 176$GHz). It is clear that o-mode can sustain growth at lower B than the x-mode.

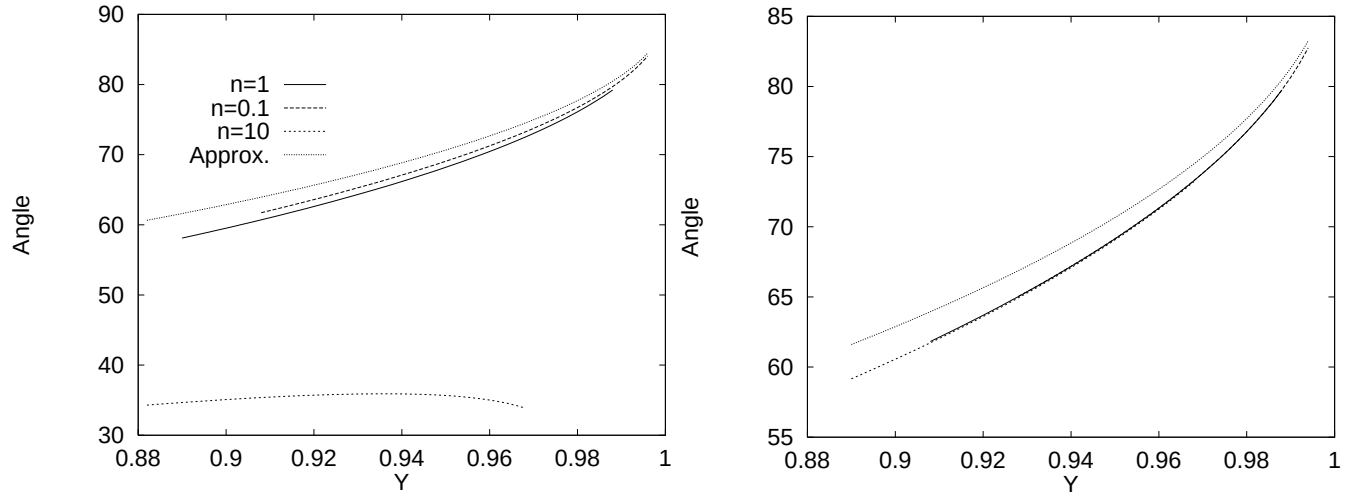

Fig. 5a and b. Plots of $Y = \Omega_0/\omega$ vs. the propagation angle θ for ECM in the o-mode for different magnetic field strengths (in Tesla) and densities (in units of $10^{16}m^{-3}$). **a** (left) $B = 0.1$T ($\Omega_0 = 17.6$GHz), **b** (right) $B = 1$T ($\Omega_0 = 176$GHz). The approximate $\theta - \omega$ relationship obtained using (7) with $N = 1$ holds better for the o-mode, so much so in fact, that it is difficult to distinguish the various density cases on each graph.

the final distribution's temperature T_1 by realising that the total parallel energy component (i.e. $p_\parallel^2/2m_e$ integrated over all particle momentum) will remain unchanged. In other words, we need to find the temperature T_1 of a normal Maxwellian distribution such that it has the same parallel energy as the initial loss cone distribution. (Note that we define the initial distribution to have developed from an original Maxwellian plasma of temperature T.) It is therefore given by

$$T_1 = \frac{T}{2}(1+\cos^3\alpha_0)$$

The maximum released maser energy is therefore

$$(\Delta E)_{\max} = \frac{nkT}{\pi^{1/2}}\int_{u_L}^{u_U}\left[\left(\frac{P_1^2}{P^2}\right)^{1/2} e^{-u_\parallel P^2/P_1^2} - (1+u_\parallel\tan^2\alpha_0)e^{-u_\parallel/\cos^2\alpha_0}\right] u_\parallel^{-1/2}du_\parallel$$

3.3. Interpreting the energy constraints

Figs. 6a–d and 7a–d show how the minimum and maximum limits on the emitted radiation for the x- and the o-mode (resp.) vary with n, B and T. Note that, as for the growth curves earlier, only curves whose parameters satisfy all the conditions given in Sect. 2.2 are plotted. It is apparent that the energies involved are comparable for the x- and o-modes, though the x-mode cannot release much energy for the $B = 0.05$ case.

These results all assume that the emitted photons travel perpendicular to the local magnetic field. In reality, where photons leave at smaller angles to the magnetic field, more photons will be needed to carry away the same amount of perpendicular momentum. Since there can be little variation in individual photon energy, more energy must be emitted in total. This means that in the general case of arbitrary θ, the minimum limit is an absolute minimum and that the maximum can be exceeded by a factor of $\sim 1/\sin\bar{\theta}$, where $\bar{\theta}$ is the typical angle of emission.

Another, more tacit assumption is that the distribution *can* actually relax to a stable state by shedding excess perpendicular momentum as the ECM operates. However, a curious possibility arises if significant growth occurs at $\theta < \alpha_0$, but is suppressed for $\theta > \alpha_0$ (e.g. see Fig. 2a–d for $B = 0.1$, $n_{16} = 1$, $T_7 = 5$). In this extreme case maximum growth takes place at a smaller angle to the magnetic field than the boundary of the loss cone; in subsequent quasilinear relaxation, conservation of energy and momentum between electrons and photons would appear to imply that wave growth will *strengthen* the anisotropy, and at least that saturation must be dominated by wave trapping. In other words the maser becomes self perpetuating, because the particles in the distribution outside the loss cone lose momentum, because of the maser, in such a way that they move *away* from the loss cone edge in momentum space. Judging by the growth rates, this action takes place on a timescale of $<\sim 10^{-6}$s - much less than either the loop crossing time or collisional time. Electrons in this range of momentum space will therefore keep on losing energy unless some other instability develops that greatly alters the distribution. For example, if the loss cone fills up on either side of the range of velocities that are experiencing a self-perpetuating maser, then an instability involving a positive $\partial f/\partial p_\parallel$ could develop. If no such instability develops, then a quasi-loss cone distribution will be maintained in some steady state in which the maser action, the normal trapping mechanism and and the effect of collisions are mutually balanced.

4. Discussion and possible applications

We conclude that the electron cyclotron maser can have large growth rates in a thermal plasma given the following conditions. (1) The density must be low enough and the magnetic field strength large enough so that $w_p < \Omega_0$. (2) A well defined loss cone must exist at electron energies not too much larger than the thermal energy - this requires a sufficiently high temperature and sufficiently low density. (3) The properties of the plasma must be such that growth can occur in much less than the time a photon takes to cross the maser region. (4) In the case where a magnetic trap is responsible for the loss cone distribution, the magnetic field strength must not greatly vary across the trap.

We have also shown that the frequency and angular distributions are determined by the the range of resonant velocities where the loss cone edge is well defined. Only in exceptional circumstances, e.g. where the loss cone angle is very large, does the actual width of the loss cone edge determine the frequency and angular distribution of the radiation.

In deriving these results we have made two main assumptions. Firstly, the temperature must not be so high (less than about 10^8K) that the weak relativistic approximation becomes invalid - at higher temperatures the ECM can still operate but requires an involved numerical calculation of the growth rate, like that of Aschwanden (1990). Secondly, the magnetic field must not be so large (e.g. neutron stars) that quantum effects become important (Melrose et al 1982).

There are two possible ways that this theory could find application in astrophysical plasmas, though we note that our original motivation was in applying it to flaring loops in the solar corona. Firstly, the ECM may be observable directly as a radio source with a very high brightness temperature (Melrose and Dulk 1982), with a distinctive polarisation determined by the local plasma conditions in the masering region. Also, the directionality of the radiation could well provide evidence that the ECM mechanism is at work. For example, on the Sun, the distribution of radio bursts over the solar disk would be markedly affected by the narrow directionality of the radiation. As the mechanism requires plasmas hotter than the ambient solar coronal temperature of 2×10^6K, the location of such bursts should correspond to coronal hot spots visible in soft X-rays.

The second application of this theory is to cross field energy transport. In plasmas with properties suitable for masering, particles cannot cross the field, yet in the solar corona, for example, many cases of simultaneous brightenings and flaring are observed on distinct sets of field lines. Melrose and Dulk (1984) have already suggested that an ECM originating in non-thermal distributions would be capable of transferring energy across field lines in sufficient quantities to explain such obser-

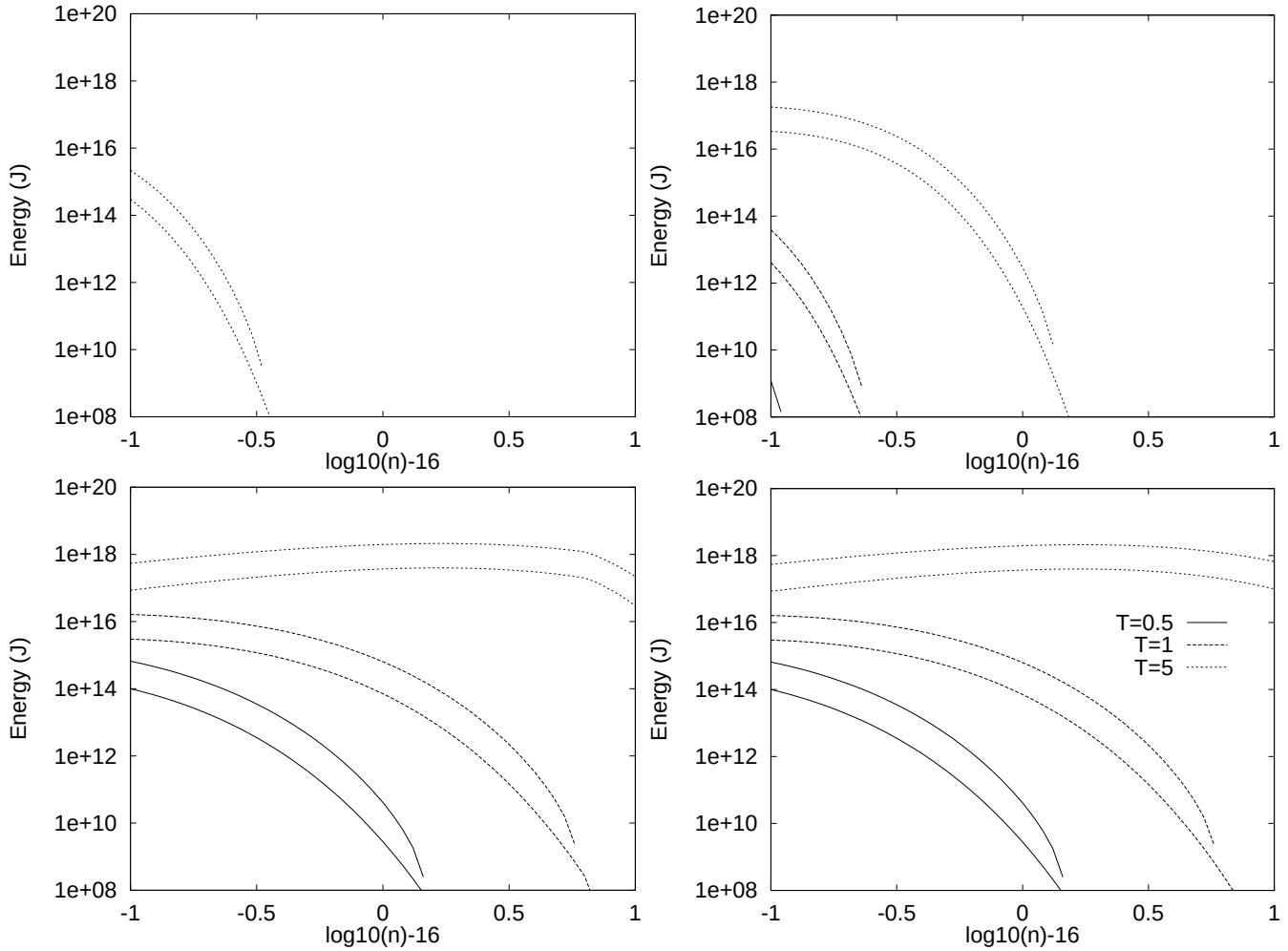

Fig. 6a–d. Plots of the minimum and maximum energies released in the x-mode over a range of densities for different magnetic field strengths (given units of Tesla) and temperatures (in units of 10^7K). **a** (upper left) $B = 0.05$T ($\Omega_0 = 8.79$ GHz), **b** (upper right) $B = 0.1$T ($\Omega_0 = 17.6$GHz), **c** (lower left) $B = 0.5$T ($\Omega_0 = 87.9$ GHz), **d** (lower right) $B = 1$T ($\Omega_0 = 176$GHz). These energies are calculated simply by multiplying the calculated energy density by a volume of 10^{19}m^3, which is typical of solar coronal loops. For reference, the total electron thermal energy in this volume is $\sim 2.1 \times 10^{19} n_{16} T_7$J.

vations. Here we suggest that hot thermal plasmas in solar flare loops could remotely heat regions by the ECM. The energies that can be released are significant, reaching up to 10% of the total electron thermal energy in the loop - see Figs. 6a–d and 7a–d. Specifically, for typical solar flare loop parameters, we have shown that one "relaxation" by the maser instability can release energies in excess of $\sim 10^{18}$J. One relaxation a second could easily satisfy the energy requirements of the hard X-ray above-the-loop-top source in the Masuda flare (Masuda 1995). The details of the energy transport from the flare loop to the remote site could be similar to the non-thermal scenario suggested in Sprangle and Vlahos (1983). In our case, an ECM originating in the thermal plasma could be absorbed in a nearby plasma, accelerating electrons that are in resonance with the wave. This could produce the kinds of anisotropic electron distributions that have already been suggested (Fletcher 1995 and Conway et al. 1998) as a possible explanation for above the loop top hard X-rays.

The conditions found in sufficiently hot coronæ of other stars could provide an even more interesting application. For example, red dwarf flare stars are believed to have ambient coronal temperatures in excess of 10^7K which would allow the thermal ECM to operate almost routinely, contributing to the quiescent radio emission. In fact, in discussing models of such quiescent emission, Kundu et al. (1987) suggested that the ECM was operating in such stars, though they only referred to it as a mechanism for producing enhanced pitch angle diffusion.

Acknowledgements. The authors would like to thank PPARC for funding and provision of the Starlink computer facilities. AJC thanks Richard Barrett, Darren MacDonald, Scott McIntosh and John Brown for many useful discussions. We also thank an anonymous referee for suggesting improvements to the paper.

Appendix: Proof of maximum energy distribution

We wish to prove that f_M (11) is indeed the final distribution with no positive $\partial f/\partial p_\perp$, and the maximum possible energy,

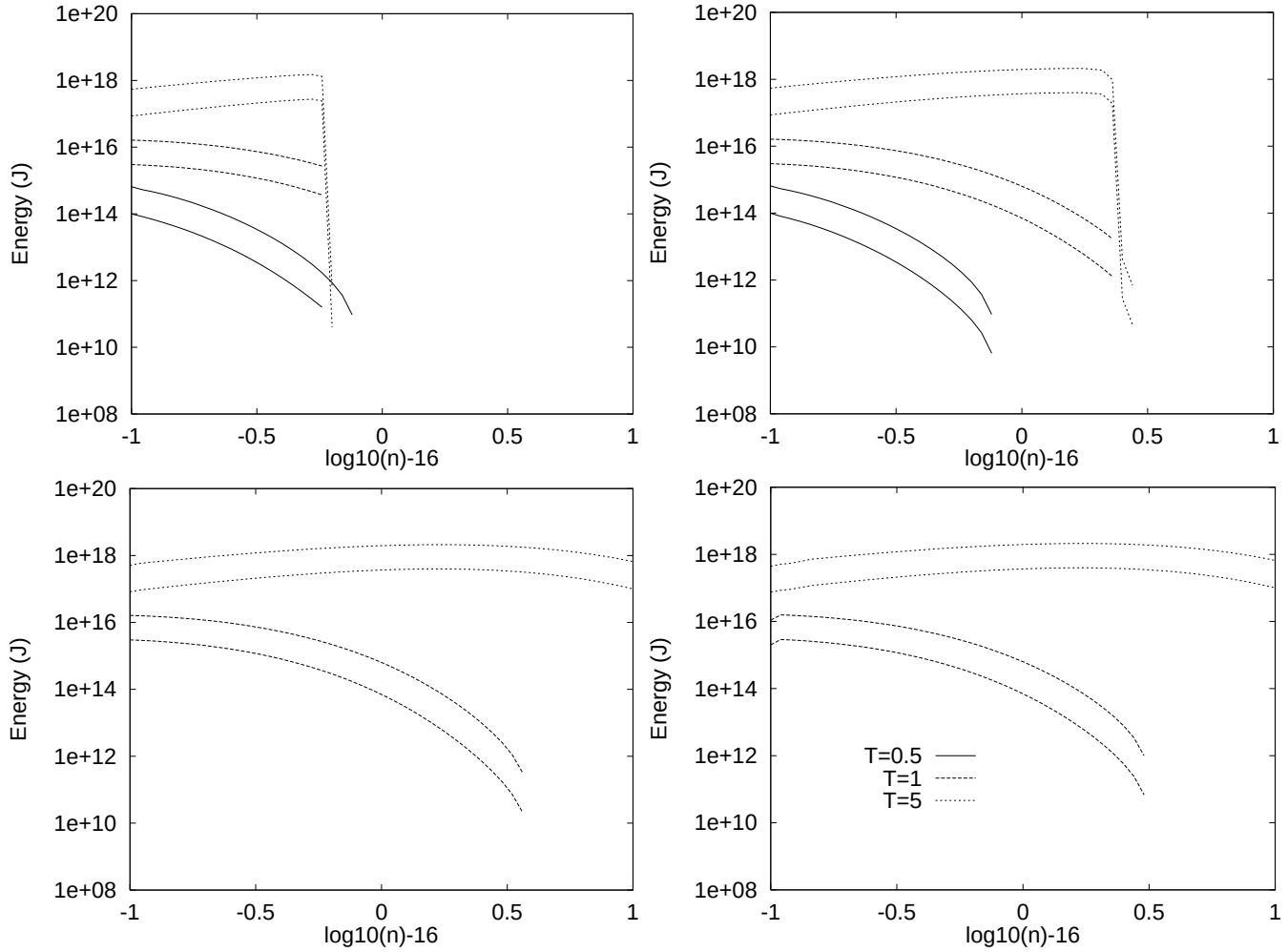

Fig. 7a–d. Plots of the minimum and maximum energys released by the electron cyclotron maser in the o-mode over a range of densities for different magnetic field strengths (given units of Tesla) and temperatures (in units of 10^7K). **a** (upper left) $B = 0.05$T ($\Omega_0 = 8.79$ GHz), **b** (upper right) $B = 0.1$T ($\Omega_0 = 17.6$GHz), **c** (lower left) $B = 0.5$T ($\Omega_0 = 87.9$ GHz), **d** (lower right) $B = 1$T ($\Omega_0 = 176$GHz). These energies are calculated simply by multiplying the calculated energy density by a volume of 10^{19}m^3, which is typical of solar coronal loops. For reference, the total electron thermal energy in this volume is $\sim 2.1 \times 10^{19} n_{16} T_7$J.

evolved from distribution f_i (8), by particles only losing momentum in the perpendicular direction. Since only the perpendicular momentum is changing, we need only consider the distributions at one $p_\parallel$, so for brevity we drop the $\perp$ subscript and omit integrals over $p_\parallel$.

Mathematically, we can state the problem as finding a function f such that

$$\text{MAX} \quad \left\{ \int_0^\infty f p^3 dp \right\} \tag{A1}$$

$$\int_0^\infty f p dp \; = \; n \tag{A2}$$

$$\int_{p_a}^{p_b} f p dp \; \leq \; \int_{p_a}^{\infty} f_i p dp \qquad \forall p_b \geq p_a \geq 0 \tag{A3}$$

$$\partial f / \partial p_\perp \; \leq \; 0 \tag{A4}$$

Note that (A3) is the mathematical translation of the fact that individual particles can only *lose* energy. Our aim is to show that there is no non-zero function δ, where $f = f_M + \delta$, that can make f meet the above requirements.

Firstly, if $\delta < 0$ for all $p \geq p_1$, the energy of distribution f must be less than f_M because individual particles can only lose energy. If $\delta > 0$ for some range of $p \geq p_1$, then according to condition (A3), this can only be achieved by bringing particles down from higher energies: f again has less energy than f_M. Therefore we have proved that δ must be zero for $p \geq p_1$.

For $p < p_1$, given that f_M is constant, (A4) means that

$$\frac{d\delta}{dp} = \frac{df}{dp} \leq 0$$

Since $\delta(p_1) = 0$, we therefore have that $\delta \geq 0$ for $p < p_1$. Now, (A2) together with the fact that $\delta = 0$ for all $p \geq p_1$ implies that

$$\int_0^{p_1} \delta p dp = \int_0^\infty \delta p dp = 0$$

Therefore δ must be zero for all $p < p_1$. We have therefore proved that f_M is indeed the distribution satisfying all the above conditions.

References

Aschwanden M.J., 1995, A& AS 85, 1141-1117
Conway A.J., MacKinnon A.L., Brown J.C., MacArthur G.K., 1998, A& A 331,1103
Fletcher L., 1995, A& A 303, L9
Kundu M.R., Jackson P.D., White S.E. and Melozzi M., 1987, ApJ 312, 822
Masuda S., Kosugi T., Hara H., Sakao T., Shibata K., Tsuneta S., 1995, PASJ 47, 677
Melrose D.B., Hewitt R.G., Parle A.J., 1982, Proc. A& A 4, 359
Melrose D.B., Dulk G., 1982, ApJ 259, 844
Melrose D.B., Dulk G., 1984, ApJ 282, 308
Melrose D.B. 1986, 'Instabilities in space and laboratory plasmas', C.U.P., Cambridge
Sharma R.R., Vlahos L., 1984, ApJ 280, 405
Spitzer L. 1962, 'Physics of fully ionized gases', John Wiley & Sons, New York
Sprangle P., Vlahos L., 1983, ApJ 273, L95
Wu C.S., Lee L.C., 1979, ApJ 230, 621
Zheleznyakov V.V. 1970, 'Radio Emission of the Sun and Planets', Pergamon Press, Oxford