Mechanism of pulsar radio emission

M. Gedalin[1], E. Gruman[1], and D.B. Melrose[2]
[1] *Ben-Gurion University, Beer-Sheva 84105, Israel*
[2] *Research Centre for Theoretical Astrophysics, School of Physics, University of Sydney, NSW 2006, Australia*

10 December 2006

ABSTRACT

We propose a pulsar radio emission mechanism that involves a plasma instability not previously considered in the context of pulsars: a nonresonant, beam-driven, hydrodynamical instability in a one-dimensional, highly relativistic, streaming, pair plasma. The growing waves are in a beam mode at frequencies below the frequency of the (known) analogous resonant instability. The instability is analyzed in detail for a cold plasma and a col beam, and the inclusion of a random relativistic spread in momenta does not change the conclusions substantially. The net amplification due to the nonresonant instability is much larger than for the resonant instability due to its broad-band nature implying growth over a much greater distance, as the ratio of the wave frequency to the resonant frequency decreases through the inhomogeneous pulsar magnetosphere. When this ratio reaches unity the beam mode joins onto the L-O mode, and the waves subsequently freely escape from the magnetosphere. Like other beam instabilities, effective growth requires a sufficiently dense beam of not too high energy particles.

Key words: Pulsars – pair plasma – waves – low frequencies – radio emission

1 INTRODUCTION

Pulsars were discovered in 1967, and since then their highly non-thermal (brightness temperatures up to 10^{29} K) pulsed radio emission (Manchester and Taylor 1977) remains one of the most intriguing puzzles of astrophysics. In the standard model for the electrodynamics of pulsars, the superstrong magnetic field (polar field of 10^{12} G) combined with the fast rotation (periods $P \sim 10^{-2} - 10^{0}$ s) results in a very highly relativistic, $\gamma \sim 10^{6} - 10^{7}$, beam of 'primary' particles, and a denser, less relativistic pair plasma produced in an electromagnetic avalanche near the poles (see, e.g., Arons 1983, and references therein). The primary beam, with density, n_p, of the order of the Goldreich-Julian value n_{GJ} (required to maintain corotation), results from acceleration by an electric field parallel to the magnetic field in an 'inner gap' near the magnetic poles. The avalanche produces a secondary pair plasma with a so-called multiplicity factor $M = n_p/n_{GJ} \gg 1$, variously estimated to be within the range $M \sim 10^{2} - 10^{6}$ (Manchester and Taylor 1977; Melrose 1995; Shibata et al. 1998); recent studies (Hibschman and Arons 2001) suggest that a lower value of M may be realistic. The pair plasma is thought to have a flow Lorentz factor of the same order as the spread in random Lorentz factors $\gamma_p \sim \langle\gamma\rangle \sim 10 - 10^{3}$. The superstrong magnetic field causes the particles to radiate away their perpendicular momenta rapidly so that the distribution function becomes one-dimensional in the momentum space. It is assumed that the observed radio emission arises in this 'pulsar plasma'. However, substantial uncertainties in our estimations of the plasma parameters remain (Zhang and Harding 2000; Zhang et al 2000; Hibschman and Arons 2001).

The mechanism of the pulsar radio emission is not adequately understood. A favored scenario involves the excitation of normal modes of the pair plasma in the polar-cap region by an instability, with the growing waves either themselves escaping from the pulsar magnetosphere ('direct' mechanisms) or being partially converted into waves that can escape ('indirect' mechanisms) (see, e.g., Melrose 1993, 1995, and references therein). Models differ in the assumed properties of the natural modes, the properties of the beam, the nature of the beam instability, the spatial location of the growth region, and the relation between the growing and the escaping waves. The favored radio source region, assumed here, is relatively close to an 'inner gap' near the pulsar surface (Cordes 1992), whereas the source region for high energy emission is thought to be in an 'outer gap' near the light cylinder. (There is evidence that some radio features correlate with the high energy emission, suggesting that there might also be radio emission from the outer gap, but we do not consider this here.) The favored type of instability involves a beam of higher energy particles propagating through a lower energy (but still highly relativistic) pair plasma. One type of beam is intrinsic to all models: the 'primary' particles with very high Lorentz factors and density $n \approx n_{GJ}$ stream through the secondary pair plasma. However, the primary beam is inadequate in exciting the waves (Magneville 1990; Asseo and Riazuelo 2000) because of the high Lorentz factor and the low density. A denser and less relativistic beam is needed for the instability to develop effectively. Such a beam may result from nonstationary plasma production (Usov 1987; Asseo and Melikidze 1998). Alternatively, a dense beam can be formed from the tail of the pair plasma distribution even if the pair production itself is stationary, and it has been argued that such a beam can be efficient in producing Langmuir waves (Lyubarskii and Petrova 2000).

Modes in a relativistic one-dimensional pair plasma have been studied extensively (see, e.g., Asseo and Riazuelo 2000, and references therein). In the approximation of the infinite magnetic field, $B \to \infty$ for a wide class of distributions the natural modes include the purely electromagnetic t-mode $\omega = kc$, and two mixed modes with transverse and longitudinal components (Gedalin et al 1998; Melrose and Gedalin 1999): the almost nondispersive subluminal Alfvén mode $\omega = k_\parallel v_A$, where the subscripts $\parallel$ and $\perp$ refer to the direction with respect to the external magnetic field and the Alfvén speed v_A is very close to the light speed, and the generally superluminal ($\omega/k_\parallel > 1$) L-O mode. The L-O mode has a longwavelength cutoff $\omega(k = 0) \neq 0$ and a frequency that may be defined as the relativistic plasma frequency ($\omega_p/\langle\gamma\rangle^{1/2}$ in the notation used below). An important point is that the L-O mode is subluminal for very small angles of propagation above a characteristic frequency, $\sim \omega_p\langle\gamma\rangle^{1/2}$, of order $\langle\gamma\rangle$ times the relativistic plasma frequency. The necessary ingredients for the radio emission mechanism were established by Melrose and Gedalin (1999): (a) only t-wave and the waves which appear on the L-O branch can eventually escape the pulsar magnetosphere as radio emission; (b) of all modes mentioned above only the L-O mode can be produced due to a beam instability; (c) the beam instability is of a hydrodynamical (reactive) type where all particles contribute to the wave excitation, in contrast with the kinetic instabilities where only a group or resonant particles participates in the wave excitation (Machabeli 1991; Lyutikov 1998).

The hydrodynamical instability has two limiting cases, referred to as resonant and nonresonant, depending on whether the frequency shift, $k_\parallel v_b$, associated with the beam is nearly equal or significantly different from the frequency of the growing waves. Previously, only the resonant beam instability has been considered, usually only for parallel propagating waves (Magneville 1990; Lyubarskii 1996; Lyutikov 1998; Asseo and Riazuelo 2000). The parallel resonant instability is inadequate for direct production of the radio emission for at least three reasons. First, it produces waves in a narrow high frequency region, $\sim \omega_p\langle\gamma\rangle^{1/2}$ Melrose and Gedalin (1999), so that subsequent conversion to lower frequencies would be necessary to account for the observations (Lyubarskii 1996). Second, the L-O mode for parallel propagation is purely longitudinal and an additional physical effect needs to be invoked for these longitudinal waves to result in escaping transverse waves. Third, the growth factor is limited due to the change in the characteristic frequency with height causing the waves to get out of resonance as they propagate, and it is difficult to explain how adequate growth can occur. The most efficient way to generate radio emission would be, however, direct excitation of electromagnetic waves in the required frequency range, so that the basic spectrum forms in the excitation region.

In the present paper we generalize and develop an idea, proposed in Gedalin et al (2002), that the nonresonant hydrodynamical beam instability can overcome the foregoing difficulties: it can account for generation of transverse waves that can escape directly. The growth occurs in a 'beam mode' with a frequency less than the characteristic frequency, $\omega_p\langle\gamma\rangle^{1/2}$, and the dispersion curve for this beam mode connects to the L-O mode at the characteristic frequency. As ω_p decreases along the ray path, the beam mode continues to grow until its frequency is equal to the characteristic frequency, and which point it becomes an L-O mode and escapes directly.

The paper is organized as follows. In section 2 we discuss the parameters of the region where radio waves are supposed to be generated. In section 3 we consider in detail the oblique beam instability in a cold plasma, including the issues of the relation between the beam and L-O modes and inhomogeneity. We show that qualitative features of the instability are not sensitive to the details of the distribution function. In section 4 we present the mechanism of the radio spectrum formation. We derive the local field line spectrum within a simple approximate model, taking into account the propagation along the curved field lines. The proposed mechanism allows to place constraints on possible parameters of the pair plasma.

2 BACKGROUND

The size, geometry, and other features of the instability region are of primary importance for any model of radio emission generation. In this section we outline the pulsar model adopted here.

2.1 The polar cap model

For simplicity we assume a perpendicular rotator, with the magnetic axis along the x-axis, and the rotation axis along the z-axis. The light cylinder radius is $R_L = cP/2\pi$, where P is the pulsar period. We do not distinguish between the rotating and nonrotating frames, cf. Muslimov & Tsygan (1992); Petrova & Lyubarskii (2000); Gedalin and Melrose (2001). Assuming dipole geometry of the magnetic field, the field lines are given by $\sin\phi/r^{1/2} = \text{const}$, where $x = r\cos\phi$ and $y = r\sin\phi$. The angle between the magnetic field and the magnetic axis is $\tan\psi = 3\sin\phi\cos\phi/(3\cos^2\phi - 1)$. The last open field line has $\phi = 90°$ at $r = R_L$, so that the equation for this line is $\sin\phi = (r/R_L)^{1/2}$. The polar cap zone is defined as the region of open fields lines, and it is limited by the last open field line. The angular size of the polar cap at the pulsar surface is $\sin\phi_p = (R_P/R_L)^{1/2}$, where $R_P \sim 10$ km is the radius of the star. In the polar cap model, the radio waves are assumed to be generated in the region of open magnetic field lines, at some height $r = R_E$, which is usually assumed to satisfy $R_P \ll R_E \ll R_L$. Then the maximum angular width of the emission region is $\sin\phi_E = (R_E/R_L)^{1/2}$. With $R_E/R_L \ll 1$, one has $\phi_E = (R_E/R_L)^{1/2} \ll 1$, and the maximum opening angle of the field line is $\psi_E \approx 3\phi_E/2$. For millisecond pulsars, for which $R_L/R_p \sim 10$, and also for emission from an outer gap in ordinary pulsars, the small angle approximation is not accurate. Our results need modification for millisecond pulsars and for emission from an outer gap, but we do not consider such modifications here.

The inhomogeneity scale of the emission region along the field lines is $L_\parallel \sim R_E$, and the inhomogeneity scale in the perpendicular scale is $L_\perp \lesssim R_E\phi_E$ and may be substantially smaller depending on the details of the pair production. The radio waves may be treated as plane waves with frequency ω and wavevector $\mathbf{k} = (k_\perp, k_\parallel)$ ($\perp$ and $\parallel$ refer to the local magnetic field direction) provided that the geometrical optics conditions, $k_\parallel L_\parallel \gg 1$ and $k_\perp L_\perp \gg 1$, are satisfied. The conditions $k_\parallel L_\parallel \gg 1$ is always well satisfied. Because of relativistic beaming only waves with $\theta = \arctan(k_\perp/k_\parallel) \ll 1$ are of interest. The transverse condition limits the propagation angle θ from below: $\theta \gg \theta_c \sim 1/kL_\perp \gtrsim 1/kR_E\phi_E$. For transverse waves with a nearly vacuum dispersion this gives $\theta_c \gtrsim c/\omega R_E\phi_E$. Technically, this implies that strictly parallel propagation is invalidated by the plasma inhomogeneity. Whether, in practice, waves with $\theta > \theta_c$ can be considered as nearly parallel depends on whether the wave dispersion changes significantly over the range $\theta \lesssim \theta_c$.

Most models of pair plasma production suggest that the secondary, denser pair plasma is generated by a cascade from the pri-

mary beam at a pair production front. The primary particles are accelerated through a potential drop up to the energies corresponding to the Lorentz-factor γ_i, whose value is uncertain (Daugherty and Harding 1982), and which is assumed to be a free parameter here. Beyond the pair production front a plasma with the density $n_p \sim M N_{GJ}$ is produced. The multiplicity factor M is unknown and theoretical estimates vary in the range $M \sim 1 - 10^4$ (Hibschman and Arons 2001). Energy conservation roughly requires $\gamma_p \lesssim \gamma_i/M$, although this depends on the efficiency of the pair production (a part of the primary beam energy is lost to radiation).

2.2 The beam that drives the instability

Beyond the pair production front, the primary particles form a low-density high-energy primary beam propagating through the secondary plasma. However, the primary beam is ineffective in causing radio waves to grow (see, e.g., discussion in Melrose and Gedalin (1999) and references therein). The theory of wave growth requires a denser, less energetic beam. Two suggestion are that this arises due to nonstationarity in the pair production cascade (Usov 1987) or resonant inverse Compton scattering (Lyubarskii and Petrova 2000). Here we simply postulate that a beam with the required properties is present. We assume a beam density n_b and typical Lorentz-factor γ_b, which are free parameters of the theory.

3 BEAM INSTABILITY REVISITED

In this section we discuss the instability caused by a cold relativistic beam penetrating cold relativistic plasma, emphasizing the nonresonant instability that has been previously overlooked.

3.1 Hydrodynamical versus maser theories

An important point that leads to a major simplification is that the instability should be hydrodynamical (or reactive) rather than kinetic (or resistive). Beam instabilities have hydrodynamical (or reactive) rather than kinetic forms, with the hydrodynamical form applying when the growth rate exceeds the bandwidth of the growing waves due to the Doppler spread, and with the kinetic form applying when this inequality is reversed. The Doppler spread is $k_\parallel \Delta v$ where Δv is the velocity spread. For a highly relativistic beam, $\gamma_b \gg 1$, the Doppler spread is intrinsically very small because most of the particles are concentrated in a very small velocity range, $\Delta v \sim c/\gamma_b^2$, just below c. Thus a kinetic instability, with $\omega \approx k_\parallel c$, can have a growth rate no larger than $\sim \omega/\gamma_b^2$, and this is too small to account for effective growth before the waves leave the growth region. Any effective instability must be hydrodynamic.

In a hydrodynamic instability effectively all the particles in the beam interact with the wave in the same way, so that their momentum spread is essentially irrelevant to the instability. This allows one to treat the instability by ignoring the velocity spread entirely, that is, assuming a 'cold' beam.

Another simplification that we make here is to assume that the wave frequency, ω, is much less than the cyclotron frequency of all relevant particles. The zeroth order term in the ratio of the cyclotron to the wave frequency corresponds to the limit of infinite field, $B = \infty$. This limit suffices for identifying the details of the instability, with inclusion of a finite-B leading only to small corrections provided that this ratio is small.

3.2 The cold, $B = \infty$ limit

Let the plasma density and Lorentz-factor be n_p and γ_p, with the corresponding quantities for the beam being n_b and γ_b. The dispersion relation for the L-O mode in the limit $B = \infty$ may be written in the form

$$\frac{\tan^2 \theta}{z^2 - 1} = \epsilon_\parallel, \tag{1}$$

where $z = \omega/k_\parallel$, θ is the propagation angle with respect to the magnetic field direction, and hereafter $c \equiv 1$ for convenience. For our purposes it is more convenient to write the parallel dielectric constant in the form $\epsilon_\parallel = 1 - (\omega_p^2/k_\parallel^2)W(z)$, with $\omega_p^2 = 4\pi n_0 e^2/m$, where n_0 is the sum of the number densities of electrons and positrons. The relativistic plasma dispersion function, $W(z)$, involves a sum over contributions, $s = \pm 1$, of electrons and positrons, each with one-dimensional distribution function f_s. One has

$$W(z) = \sum_s \alpha_s \int \frac{df_s}{v - z}, \tag{2}$$

with $\alpha_s = n_s/n_0$, $v = (1 - 1/\gamma^2)^{1/2}$. We consider only forward propagating modes, $k_\parallel > 0$, and then $W(z)$ is well-defined for $\mathrm{Im}\, z > 0$. One has to analytically continue $W(z)$ into the half-plane $\mathrm{Im}\, z < 0$. With this notation, the dispersion relation (1) becomes

$$\frac{k_\parallel^2}{\omega_p^2} = \frac{W(z^2 - 1)}{z^2 - z_0^2}, \tag{3}$$

with $z_0 = 1/\cos\theta$.

The cold plasma-beam system is given by $f_s = \delta(v - v_s)$. Then one has

$$W(z) = \frac{1}{\gamma_p^3(z - v_p)^2} + \frac{\alpha}{\gamma_b^3(z - v_b)^2}, \tag{4}$$

with $\alpha = n_b/n_p$. It is convenient to write $z = v_b + \eta$. The large ratio $(\gamma_b/\gamma_p)^3$ makes the contribution of the beam negligible provided $|\eta| \ll v_b$. Then (4) gives

$$W(z) = 4\gamma_p + \frac{\alpha}{\gamma_b^3 \eta^2}. \tag{5}$$

Substituting $z = v_b + \eta$ and expanding in small quantities, including $\theta \ll 1$, one finds

$$\left(\frac{k_\parallel^2}{\omega_r^2}\right)\left[1 - \frac{\theta^2 \gamma_b^2}{2\gamma_b^2 \eta - 1}\right] = 1 + \frac{\alpha}{4\gamma_p \gamma_b^3 \eta^2}, \tag{6}$$

where, in the absence of the beam, the frequency $\omega_r = 2\omega_p \gamma_p^{1/2}$ separates the region $\omega \ll \omega_r$ where the L-O mode is highly superluminal and nearly longitudinal from the region $\omega \gg \omega_r$ where it is nearly transverse with vacuum-like dispersion. For small propagation angles, $\theta \gamma_b \ll 1$, the dependence on the angle is negligible for $|2\gamma_b^2 \eta - 1| \gg 1$.

3.3 The growth rate

The dispersion equation (6) is a cubic equation for $\eta = z - v_b$, and instability occurs when two of the solutions are a complex conjugate pair. The growth rate is then given by $\Gamma = k_\parallel v_b\, \mathrm{Im}\, z = \omega\, \mathrm{Im}\, \eta$.

The nonresonant growth rate applies at low frequencies, $k_\parallel \ll$
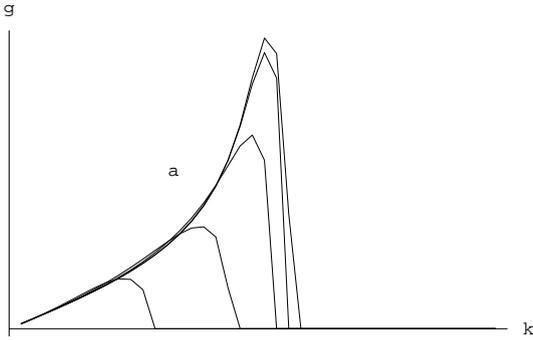
4   *Gedalin et al.*

Figure 1. Growth rate vs frequency for the cold-cold case and different angles of propagation. The instability range squeezes as the propagation angle increases: higher growth rates and larger regions of instability correspond to smaller propagation angles.

ω_r, when the left hand side of (6) can be neglected. The solution of the resulting quadratic equation gives

$$\frac{\Gamma_n}{\omega} = \left(\frac{n_b}{n_p}\right)^{1/2} \frac{1}{2\gamma_p^{1/2}\gamma_b^{3/2}}, \qquad (7)$$

where the approximation $k_\parallel/\omega_r \ll 1$ is used.

The instability becomes resonant for $k_\parallel = \omega_r$, and the growth rate is then

$$\frac{\Gamma_r}{\omega_r} = \frac{\sqrt{3}}{2\sqrt{2}}\left(\frac{n_b}{n_p}\right)^{1/3}\frac{1}{2\gamma_p\gamma_b}. \qquad (8)$$

The general solution of (6) reduces to (7) and to (8) in the respective limits. For higher frequencies the three solutions are real, implying no growth for $\omega > \omega_r$.

The ratio of the nonresonant to the resonant growth rate,

$$\frac{\Gamma_n/\omega}{\Gamma_r/\omega_r} \approx \left(\frac{n_b}{n_p}\right)^{1/6}\left(\frac{\gamma_p}{\gamma_b}\right)^{1/2}, \qquad (9)$$

is not especially small unless n_b/n_p and γ_p/γ_b are very small. These results apply for sufficiently small θ, and for larger θ the term $\propto \theta^2$ in (6) tends to reduce the instability range, in the sense that the maximum frequency at which growth is possible decreases as a function of increasing θ.

Figure 1 shows the dependence of the growth rate on frequency for various angles of propagation. Units are not shown since the parameters are chosen for convenience of presentation and do not correspond to real pulsar plasma parameters. It is seen that the instability range squeezes as the propagation angle increases. The maximum growth rate occurs near the resonance $\omega = \omega_r$.

3.4   The amplification factor for a cold beam

In a homogeneous plasma the fastest growing instability should dominate, and it is for this reason that the resonant instability has been assumed to dominate in pulsar plasma. However, the situation is different in the pulsar plasma which is flowing outwards, so that the density drops as $n \propto R^{-3}$. We now argue that in this more realistic case it is the nonresonant instability that dominates.

Let the resonance width be $\Delta\omega_r$, that is, the wave is in resonance if $|\omega - \omega_r| < \Delta\omega_r$. This width can be estimated from the resonance condition $|\omega - k_\parallel v_b| < \Gamma_r$ and the dispersion for the wave, $\omega = k_\parallel v_p + \omega_p/\gamma_p^{3/2}$. One has $\Delta\omega_r/\omega_r \sim$

$(n_b/n_p)^{1/3}(\gamma_p/\gamma_b)^{1/2} \sim (\Gamma_n/\Gamma_r)^2$. The energy going into growing waves is proportional to the product of the growth rate and the bandwidth of the growing waves. The bandwidth of growing waves is the resonance width and the bandwidth for the nonresonant instability is much broader, of order the frequency itself. Hence the ratio of the bandwidths is of order $\Delta\omega_r/\omega_r$, and the ratio of the energy going into resonant and nonresonant waves is $\sim \Gamma_r\Delta\omega_r/\Gamma_n\omega_r \sim \Gamma_n/\Gamma_r$. It follows that the effect of the bandwidth of the growing waves dominates over the growth rate, and that for $\Gamma_n/\Gamma_r \ll 1$, contrary to intuition, the nonresonant waves dominate over the resonant waves.

The amplification factor for the waves is given by integrating the growth rate along the ray path. The foregoing argument implies that the amplification factor for nonresonant growth exceeds that due to resonant growth due to the much longer path over which the waves grow. This effect may be estimated by assuming that the resonant frequency varies with radius, as $\omega_r \propto \Delta\omega_r \propto R^{-3/2}$, and integrating along the ray path. Suppose that the wave is in the resonance at $R = R_0$. For resonant growth it would remain in the resonance only over a distance $|R - R_0|$ given by $|R - R_0|/R_0 \sim \Delta\omega_r/\omega_r \ll 1$. Assuming linear growth, such that the wave amplitude grows according to $(da/dt) = \Gamma a$, the growth in wave energy over a path $R_1 < R < R_2$ is determined by the amplification factor

$$G(\omega, \theta) = 2\int_{R_1}^{R_2} \Gamma(R, \omega, \theta)dR/c. \qquad (10)$$

The contribution of the resonance is typically small, and is at most comparable with that of the nonresonant part. The energy going into resonant waves certainly does not dominate that going into nonresonant waves.

3.5   The escaping radiation

Beyond the resonant frequency the instability ceases and the L-O waves are in the transparency range. Figure 2 shows the behavior of the normal modes in the transparency region. The left panel shows the range $-\gamma_p^2 < z - 1 < \theta^2$. The right panel magnifies the range $-\gamma_b^2 < z - 1 < \theta^2$ which cannot be resolved in the left panel, since $1/\gamma_b^2 \ll 1/\gamma_p^2$. The superluminal mode with $z > 1 + \theta^2/2$ ($\theta \ll 1$) is the L-O mode. It corresponds to $\omega = kc$ in the high frequency limit, $\omega/\omega_r \gg 1$. Four other modes have $\omega = k_\parallel v_p$ or $z - 1 = 1/2\gamma_p^2$ (two) and $\omega = k_\parallel v_b$ or $z - 1 = 1/2\gamma_b^2$ (other two) asymptotically (that is, for $k_\parallel \to \infty$), and stand in the plasma and beam frame, respectively. We note that the "beam mode" $z = v_b$ is electromagnetic: the polarization is determined by relation as

$$E_x/E_z = \tan\theta/(1 - z^2), \to E_\perp/E_\parallel \approx \theta/(1/2\gamma_b^2 + \theta^2). \quad (11)$$

The last ratio is large for $1/\gamma_b^2 \ll \theta$.

The instability occurs in the range $1 - 1/2\gamma_p^2 < z < 1 - 1/2\gamma_b^2$, where there is a gap between the dispersion curve and the z axis. The instability ceases in the lowest point of this curve, where two complex conjugate solutions of the dispersion relation become a double real solution. The unstable mode should proceed further along one of this bifurcating branches.

3.6   Effect of a relativistic spread in momenta

To justify our neglect of the spread in velocity or momentum in the beam, we now include a spread and show that it does not affect our conclusions significantly.
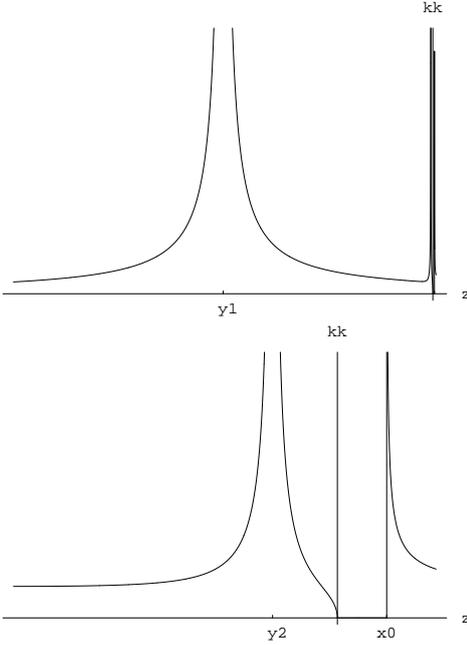
© 0000 RAS, MNRAS **000**, 000–000

Figure 2. The transparency range for the cold-cold case (magnified in the right panel). See explanation in text.

As an example of a hot plasma distribution we consider the double waterbag distribution of the following form (here $y = z - 1$, see notation in Appendix): $f = n_p A_p H(y_1 - y) H((y - y_2) + n_b A_b H(y_3 - y) H(y - y4)$, where $H(y)$ is the step function, $H(y \geqslant 0) = 1$, $H(y < 0) = 0$, the ordering of the parameters $y_i = 1/2\gamma_i^2$ is $y_1 > y_2 > y_3 > y_4$, and the normalizing parameters A_p and A_b are obtained from the normalization condition as follows: $A_p = 1/\sqrt{2y_2} - 1/\sqrt{2y_1}$ (A_b respectively). The function $W(z)$ is

$$W(z) = \frac{\omega_p^2 A_p (y_1 - y_2)}{(x + y_1)(x + y_2)} + \frac{\omega_b^2 A_b (y_3 - y_4)}{(x + y_3)(x + y_4)}. \quad (12)$$

Figure 3 shows $K(x)$ in the transparency range. The parameters used are $\alpha = 0.1$, $\theta = 0.1°$, $\gamma_1 = 10$, $\gamma_2 = 50$, $\gamma_3 = 500$, and $\gamma_4 = 1000$.

The corresponding growth rate is shown in Figure 4 as a function of the propagation angle. As anticipated, the growth rate is similar to what is found for the cold case.

The double-waterbag distribution is discontinuous at the edges. It has been shown (Gedalin et al 1998)that the Alfvén and beam mode are expected to damp for more realistic distributions (although damping depends on the details of the distribution function behavior and may be weak). The only mode which is expected to propagate freely beyond the resonance point, is the L-O mode. To illustrate what happens in this case we consider the "double soft bell" distribution

$$f = n_p A_p (y-y_1)^2 (y-y_2)^2 H(y_1-y)H(y-y_2) + n_b A_b (y-y_3)^2 (y-y_4)^2 H(y_3-y)H(y-y_4) \quad (13)$$

The exact expressions for the normalization parameters $A_{p,b}$ and for $W(z)$ are cumbersome and we do not quote them here. Figure 5 shows $K = \sqrt{(\text{Re } F) \cdot H(F)}$ and $\text{Im } F$, with $F = k_\parallel^2/\omega_p^2$, for the double soft distribution with $\gamma_1 = 5$, $\gamma_2 = 60$, $\gamma_3 = 50$, $\gamma_4 = 200$, $n_b/n_p = 0.3$, and $\theta = 0.5°$. The distribution function is shown in Figure 6. In this case the distribution and its first derivative are continuous, and there are no gaps ($\gamma_3 < \gamma_2$).
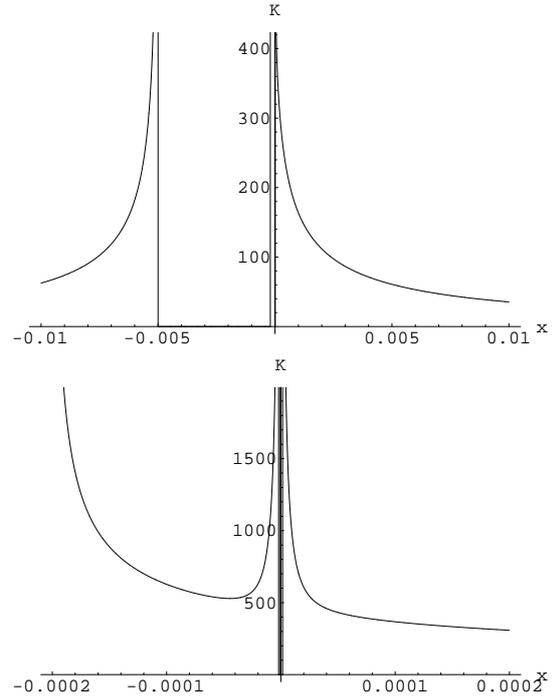

Figure 3. (Left) K as a function of x for the double waterbag case with $\alpha = 0.1$, $\theta = 0.1°$, $\gamma_1 = 10$, $\gamma_2 = 50$, $\gamma_3 = 500$, and $\gamma_4 = 1000$; (Right) The instability region magnified.
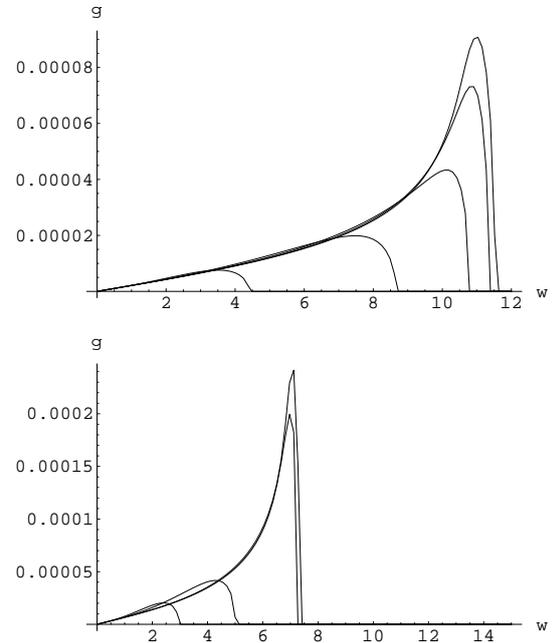

Figure 4. Left: growth rate vs frequency for the double waterbag distribution with $\gamma_1 = 10$, $\gamma_2 = 50$, $\gamma_3 = 500$, and $\gamma_4 = 1000$, and propagation angles $\theta = 0.01°$, $0.05°$, $0.1°$, $0.2°$, and $0.5°$. Right: same for $\gamma_1 = 5$, $\gamma_2 = 20$, $\gamma_3 = 200$, and $u_\gamma = 500$, and $\theta = 0.05°$, $0.1°$, $0.5°$, $1.0°$. The instability range progressively diminishes towards larger propagation angles as in Figure 1.

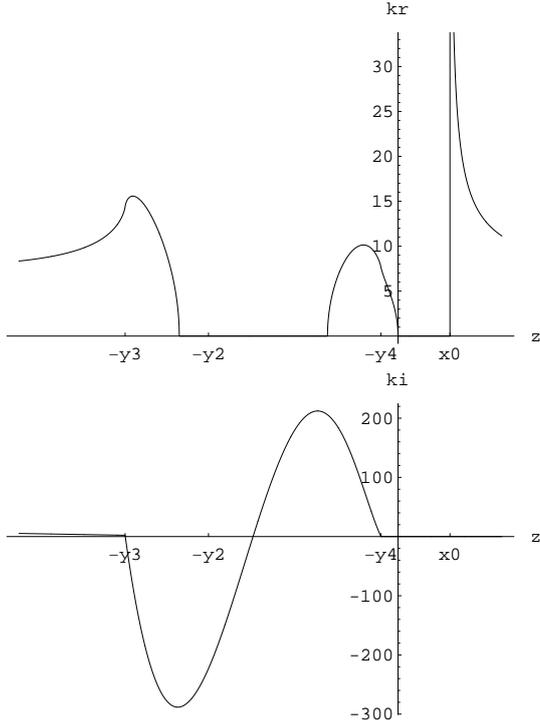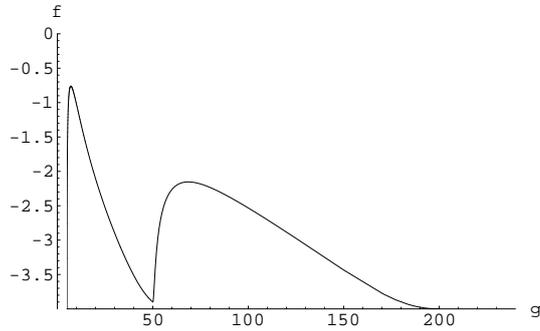**Figure 5.** Transparency region (see text).

Figure 6. Overlapping double soft bell distribution.

Figure 7. Phase velocity (left) and normalized growth rate (right) for double soft bell distribution case.

It is seen that above a certain frequency only L-O mode with $z > 1/\cos\theta$ can propagate, that is, the plasma is not transparent for other modes. Figure 7 shows the phase velocity and growth rate of the unstable mode. The maximum growth rate occurs at the frequency $\omega = K\omega_p < \omega_p\sqrt{\gamma_2}$, and below this maximum, the growth rate $\Gamma/\omega = \text{Im}\, z$ is almost independent of ω. Thus, the effective resonance frequency ω_r decreases relative to the frequency predicted by the cold plasma-beam analysis. The instability ceases at $\omega = \omega_c \approx 2\omega_p\sqrt{\gamma_2}$, which also corresponds to the frequency beyond which only the L-O mode can propagate.

In an inhomogeneous pulsar plasma the point $\omega = \omega_c$ would correspond to the point of refractive conversion of the unstable beam mode into L-O mode. Indeed, the propagating unstable mode grows until $\omega \approx \omega_r$ and then propagates almost without change up to the critical radius where $\omega = \omega_c$. This mode cannot propagate any further. Since the polarizations of the unstable mode and L-O mode are the same at this point, the unstable mode simply joins onto L-O mode (see Appendix B). At the point of conjunction, ω and $k_\perp$ do not change. The unstable mode has $\omega \approx k_\parallel$ while the L-
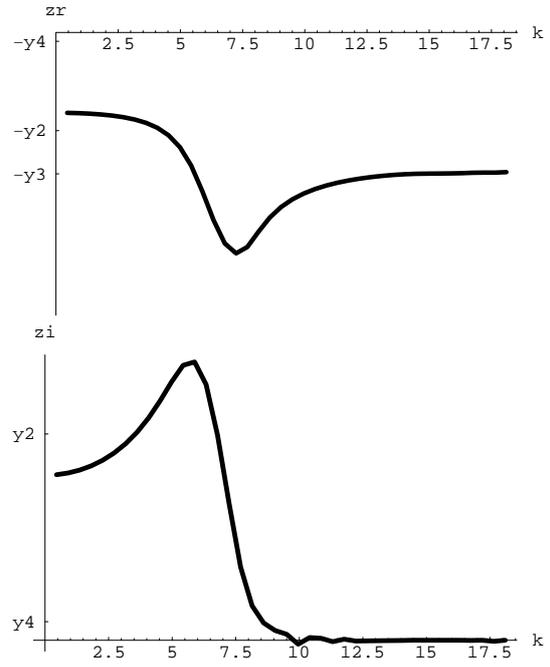
O mode has $\omega = k$. Thus, the propagation angle of the L-O mode is related to the propagation angle of the unstable mode by the relation $\theta' = \theta/\sqrt{1-\theta^2} \approx \theta$, where we assumed $\theta \ll 1$. Thus, to within our approximations, the propagation angle as the growing mode joins onto the L-O mode.

3.7   Validity of the cold-plasma, cold-beam model

Finally, we show that the instability growth rate may be taken from the cold-cold approximation even if the true distribution is far from begin cold. Let us assume that the plasma and beam distributions are well separated in velocity space (this is not necessary as we have seen above, but simplifies the analysis). In other words, we assume that the plasma body is contained within $\gamma_1 < \gamma < \gamma_2$ and the beam is contained within $\gamma_3 < \gamma < \gamma_4$, with $\gamma_3 \gg \gamma_2$. This does not necessarily mean that there are no particles in the regions $\gamma < \gamma_1$, $\gamma_2 < \gamma < \gamma_3$, and $\gamma > \gamma_4$. However, we do assume that number of these particles and their contribution are negligible. We also assume that the distribution function is steep at the low- and high-energy ends. These criteria are by no means rigorous and should, in principle, be verified in each individual case. We do not make any a priori assumption about the behavior of the distributions inside the defined intervals. The behavior of the unstable mode is determined by the function $W(z)$ when $\text{Im}\, z > 0$. As we have seen from above, in the nonresonant instability regime $z = v_n + i\sigma$, where $v_n \lesssim v_3$, and $1/\sigma_2^2 \gtrsim \sigma \gtrsim 1/\gamma_3^2$. With this one finds

$$W(z) \approx \langle\gamma\rangle_p - \frac{\alpha\langle\gamma^{-3}\rangle_b}{\sigma^2}, \qquad (14)$$

where $\langle\ldots\rangle = \int f(\ldots)d\gamma$. Comparing (14) with the expressions for the cold case one can see that the only change which has to be done is the substitution $\gamma_p \to \langle\gamma\rangle_p$, and $\gamma_b^{-3} \to \langle\gamma^{-3}\rangle_b \sim 1/\gamma_3^3$. With these substitutions the dispersion relation is reduced to that of the cold case. To this one has to add that the only non-damping waves are the superluminal L-O mode $z - 1 > x_0$.

From the above we conclude that the beam instability develops in a qualitatively similar manner irrespective of the distribution functions of the relativistic pair plasma and of the beam. The basic features of the instability are: (a) the growth rate for slightly oblique propagation does not differ much from the growth rate for the parallel propagation; (b) unstable oblique waves are transverse, whereas unstable parallel waves are longitudinal; (c) the nonresonant instability dominates over resonant instability in the outflowing pulsar plasma; (d) at the point where the instability ceases the unstable mode joins onto L-O mode.

It is worth mentioning that for the distributions found in numerical analyses (Daugherty and Harding 1982; Hibschman and Arons 2001) ($f \propto \gamma^{-3/2}$ or $f \propto \gamma^{-2}$) the resonance frequency depends only weakly on the maximum Lorentz factor of the plasma ($\langle \gamma \rangle \propto \gamma_2^{1/2}$ or $\langle \gamma \rangle \propto \ln \gamma_2$).

We suggest that the oblique nonresonant instability with the immediate conversion of the unstable mode into the propagating L-O wave provides the mechanism for local generation of radio waves. In the following section we consider the implications of this mechanism for the pulsar magnetosphere.

4 MECHANISM

The emission mechanism proposed here involves three stages: (a) radio noise, from the background or from incoherent emission, is amplified in the nonresonant, oblique-angle instability, with amplification occurring over a broad frequency range; (b) as the waves propagate outward, the ratio, ω/ω_r, of the wave frequency to the characteristic frequency increases, and when this ratio reaches unity the amplification ceases; (c) at the point where the ratio reaches unity the waves, which grow in a beam mode, evolve into L-O mode waves which then escape freely, at least until they reach the region where the wave frequency is equal to the cyclotron frequency of the ambient particles Luo and Melrose (2001). In this section we discuss some further details of this proposed mechanism.

4.1 Parallel and oblique propagation

It is usually assumed that waves propagating parallel to the magnetic field grow (see, e.g. Lyubarskii and Petrova (2000) and references therein). This is the case only if the beam instability is most efficient for $\theta = 0$. We now argue that this is not the case.

We have shown (Gedalin et al 2002) that the efficiency of oblique wave generation does not change much for small angles of propagation. For the same growth rate the efficiency of the instability, in terms of the wave energy generated, is proportional to the phase space available to the instability. The foregoing calculations imply that the beam instability can be considered as (quasi)parallel only for $\theta \lesssim 1/\gamma_b^2$; in the range $1/\gamma_b^2 < \theta \lesssim 1/\gamma_b$ the instability is oblique in character. The ratio of the available phase spaces for the oblique case is of order $\sim \gamma_b$ greater than for the parallel case. It follows that the waves that grow are dominated by those propagating at oblique angles.

4.2 Model for the amplification factor

A generic model for any form of coherent emission involves a statistical collection of elementary bursts, each of which is proportional to a growth factor, G. For a wide variety of sources it seems that G obeys log-normal statistics Cairns et al (2001). Here we are not concerned with the details of the statistics, but it is important to estimate the typical expected value of G to show that it is large enough to account for effective growth.

Let us consider a single field line, or, more precisely, a narrow magnetic tube in which the electromagnetic waves are amplified in a burst of wave growth. Let $I(\omega, \theta, R)$ be the wave intensity in the unstable region. Then one has $dI/dt = 2\Gamma(\omega, \theta, R)I$. This radiation transfer equation can be rewritten as

$$\frac{dI}{dR} + \frac{d\theta}{dR}\frac{dI}{d\theta} = 2\Gamma I, \qquad (15)$$

where the group velocity of the excited waves is $v_b \approx 1$, and the evolution of the angle θ between the wavevector and the magnetic field direction is described by (Barnard and Arons 1986)

$$\frac{d\theta}{dR} = \frac{3\phi}{4R} - \frac{3\theta}{2R}. \qquad (16)$$

For a field line $\phi/R^{1/2} = \text{const.}$ Eq. (15) is valid only in the instability region, $\omega < \omega_c(R, \theta)$. Let $R = R_r(\theta, \omega)$ be the radius where growth rate ceases (where $\omega = \omega_r$). Beyond this radius the mode $\omega \approx k_\parallel v_b$ propagates up to the radius R_c where $\omega = \omega_c$, so that the angle θ continues to grow, while the intensity does not change. Eq. (15) implies that the intensity at R_r is

$$I(\omega, \theta, R_r) = GI(\omega, \theta_0, R_0),$$
$$G = \exp 2 \int_{R_0}^{R_r} \Gamma(\omega, \theta'(R), R)dR, \qquad (17)$$

where $\theta'(R)$ is the solution of (16) with the boundary condition $\theta' = \theta$ at $R = R_r$, and R_0 is the lowest radius at which the instability sets in:

$$\theta' = (\theta - \frac{3\phi_r}{8})\left(\frac{R_r}{R}\right)^{3/2} + \frac{3\phi_r}{8}\left(\frac{R}{R_r}\right)^{1/2}. \qquad (18)$$

From Figure 4 the highest unstable frequency ω_r remains almost the same for $\theta < \theta_c \sim 1/\gamma_b$, and decreases approximately linearly with the angle increase for $\theta > \theta_c$. Here we consider a simplified model with

$$\Gamma(R, \theta') = \omega \left(\frac{n_b}{n_p}\right)^{1/2} \left(\frac{1}{\gamma_p \gamma_b^{3/2}}\right) H\left(\omega_r(R) - \omega\right) H(\theta_c - \theta'). \qquad (19)$$

This approximation implies either $\theta = \theta_c$ and $\omega_r(R_c) > \omega$ or $\theta < \theta_c$ and $\omega_r(R_c) = \omega$. The only dependence on the radius in (19) comes from the step functions. Let the instability onset occur at $R = R_0$ on the field line with $\phi(R = R_0) = \phi_0$, and let the initial wave parameters be ω, θ_0. Then the amplification factor is

$$G = 2\omega \left(\frac{n_b}{n_p}\right)^{1/2} \left(\frac{1}{\gamma_p \gamma_b^{3/2}}\right) R_c, \qquad (20)$$

where R_c is the minimum of the two radii found from the conditions $\omega_r(R_c) = \omega$ or $\theta(R_c) = \theta_c$. For $\omega_r(R_c) = \omega$ the final propagation angle is

$$\theta = (\theta_0 - \frac{3\phi_0}{8})\left(\frac{R_0}{R_c}\right)^{3/2} + \frac{3\phi_0}{8}\left(\frac{R_c}{R_0}\right)^{1/2} < \theta_c. \qquad (21)$$

Neglecting in (21) the rapidly decreasing term $\propto R^{-3/2}$, one gets

$$\ln G = 4 \left(\frac{n_b}{n_p}\right)^{1/2} \frac{\omega_{p0} R_0}{c \gamma_b^{3/2}} \times \begin{cases} \xi^{1/3} - \xi & \eta < \xi^{1/3} \\ \eta^{-2} - 1 & \eta > \xi^{1/3} \end{cases}, \qquad (22)$$

where $\xi = \omega/\omega_{r0}$ and $\eta = 3\phi_0/8\theta_c$. Here we use $\omega_r \propto R^{-3/2}$. Figure 8 shows the dependence of the gain G on (dimension-
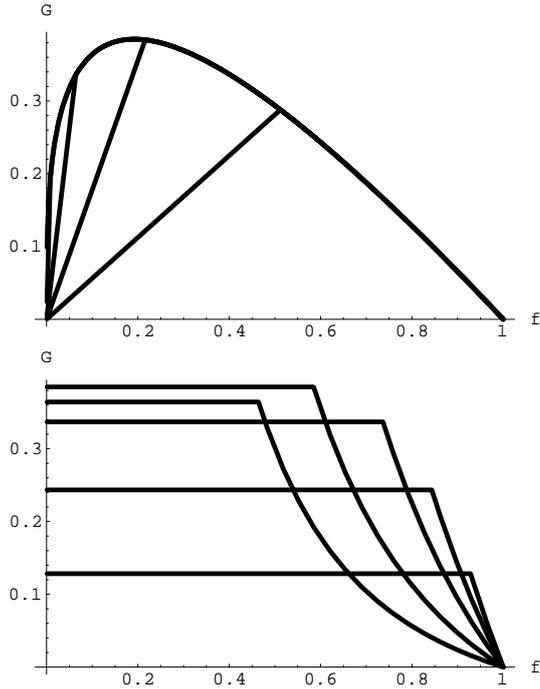
Figure 8. Gain as a function of frequency, $\xi = \omega/\omega_{r0}$ (left panel), and opening angle, $\eta = 3\gamma_b\phi_0/8$ (right panel).

less) frequency ξ and on the (normalized) field line opening angle η. It is worth mentioning that the wave propagates at the angle $\theta + 3\phi_c/2 \lesssim 5\theta_c$ to the magnetic axis at the point where the instability ceases. The maximum beam half-width at the decoupling site (beyond which the waves propagate along straight lines) is $\gtrsim 5\theta_c$, which is $\sim 5°$ for $\gamma_b \sim 50$. The simplified model used here neglects the wings $\theta > \theta_c$, so that we expect that the resulting width to be somewhat larger. The pulse half-peak-intensity width should be determined by $\omega_r = \omega$, so that we expect $R \propto \omega^{-2/3}$ and the width $\propto R^{1/2} \propto \omega^{-1/3}$.

5   DISCUSSION AND CONCLUSIONS

The above analysis shows that the nonresonant oblique beam instability may be quite efficient in generation of low-frequency electromagnetic waves for very wide class of particle distributions. In the plasma which flows outwards in the diverging magnetic field the nonresonant form of the instability is more important than the previously considered resonant form. The difficulties with the resonant form included too small a growth factor and too high a frequency, and both these problems are alleviated by invoking the nonresonant instability. The unstable waves are in a beam mode below the resonant frequency, and as the waves propagate outward the beam mode joins onto the L-O mode, so that the waves can freely leave the pulsar magnetosphere.

The most important requirement of any model for the wave growth is that it be capable of accounting for the very bright emission observed. We estimate that an amplification is required from an initial effective temperature, $T_i \sim \gamma_p m_e c^2$, of the waves to $10^{23} - 10^{26}$K. For $\gamma_p \sim 10$ one has $T_i \sim 10^{11}$K, so that an amplification factor $G \sim 10^{12} - 10^{15}$ is required. From the analysis in this paper, this requires that the plasma parameters satisfy

$\omega_{p0}R_0/c\gamma_b^{3/2} \gtrsim 100$ at the site of onset of the instability. It should be understood that this estimate constrains any beam instability scenario, and no convincing case has been made that any known instability can satisfy it.

Yet another constraint follows from the observation that the maximum intensity should be observed at the frequency $\sim 0.1\omega_{p0}\sqrt{\gamma_p}$. Together with the previous estimate one finds that a frequency ω can be efficiently generated for $\omega R_0/c\gamma_b^{3/2} \gtrsim 100$. For $\gamma_b \sim 10^2$ and $\omega \sim 10^9$s^{-1} one finds $R_0 \gtrsim 3 \cdot 10^6$cm, which is within the expected region.

The intrinsic frequency spectrum of a single burst of wave growth possesses some of the features of the observed pulsar radio spectrum. The intrinsic spectrum decreases towards higher and lower frequencies, and the slope also increases towards higher and lower frequencies. The derived shape is not a power-law spectrum. However, this is a single-field-line spectrum, and the observed power law spectrum results from an integration over a very large number of elementary bursts of wave growth from a number of field lines. Hence the observed spectrum should be determined by the distribution of elementary bursts, e.g., as a function of the frequency maximum and other relevant parameters. Thus, it is inappropriate to attempt to explain the observed spectrum in terms of the properties of the instability alone.

The mechanism predicts efficient generation of electromagnetic waves propagating at angles $\theta \lesssim 1/\gamma_b$, which, for most accepted estimates, is within several degrees and typically less than the observed pulse widths. The high-frequency, $\omega \gg \omega_p\sqrt{\gamma_p}$, L-O mode propagates along straight lines, so that additional refraction zones are needed to increase the pulse width, unless the effective γ participating in the instability are sufficiently low.

In the treatment here, the waves which are excited by the nonresonant beam instability, are completely linearly polarized. This results from our neglect of gyrotropy resulting from any difference in the distributions of electrons and positrons. Inclusion of gyrotropy is essential when considering circular polarization, but a detailed investigation of this point is beyond the scope of the present paper.

Although the mechanism proposed here alleviates some of the problems with other proposed beam mechanisms, it requires a relatively high-density, low-Lorentz-factor beam. The required beam must flow through the background pair plasma, requiring that it have an even lower Lorentz factor. There have been several suggestions as to how such dense beams might form Usov (1987); Asseo and Melikidze (1998); Lyubarskii and Petrova (2000), and although the mechanism proposed here is not sensitive to the details, such a beam is an essential pre-requirement.

6   ACKNOWLEDGMENTS

REFERENCES

Arons, J., ApJ, 266, 175 (1983).
Asseo, E., and Melikidze, G.I., 1998, MNRAS, 301, 59.
Asseo, E., and Riazuelo, A., 2000, MNRAS, 318, 983.
Barnard, J.J., and Arons, J., 1986, ApJ, 302, 138.
Cairns, I.H., Johnston, S., and Das, P., 2001, ApJ, 563, L65.

Cordes, J.M., in *The Magnetospheric Structure and Emission Mechanisms of Radio Pulsars* (eds T.H. Hankins, J.M. Rankin and J.A. Gill, Pedagogical University Press, Zielona Góra, Poland, 1992), p. 253.

Daugherty J.K. and Harding A.K., 1982, ApJ, 252, 337.

Gedalin, M., and Melrose, D.B., 2001, Phys. Rev. E, 6402, 7401.

Gedalin, M., Gruman, E., and Melrose, D.B., New mechanism of pulsar radio emission, 2002, Phys. Rev. Lette., 88, 121101.

Gedalin, M., Melrose, D.B., and Gruman, E., 1998, Phys. Rev. E, 57, 3399.

Hibschman, J.A., and Arons, J., 2001, ApJ, 554, 624; 2001, ApJ, 560, 871.

Luo, Q., and Melrose, D.B., 2001, MNRAS, 325, 187.

Lyubarskii, Yu.E., 1996, Astron. Astrophys., 308, 809.

Lyubarskii, Y.E., and Petrova, S.A., 2000, Astron. Astrophys., 355, 406.

Lyutikov, M., 1998, MNRAS, 293, 447.

Machabeli, G.Z., 1991, Plasma Phys. Contr. Fus. 11, 1227.

Magneville, A., 1990, J. Plasma Phys., 44, 213.

Manchester, R.N. and Taylor, J.H., Pulsars, Freeman & Co, San Francisco, 1977.

Melrose D.B., in Pulsars as Physics Laboratories, (eds R.D. Blandford, A. Hewish, A.G. Lyne and L. Mestel, Oxford University Press, 1993) p. 105.

Melrose D.B., 1995, J. Astrophys. Astron., 16, 137.

Melrose, D.B., and Gedalin, M., 1999, ApJ, 521, 351.

Muslimov, A.G., and Tsygan, A.I., 1992, MNRAS, 255, 61.

Shibata S., Miyazaki J. & Takahara F., 1998, MNRAS, 295, L53.

Petrova, S.A., and Lyubarskii, Y.E., 2000, Astron. Astrophys., 355, 1168.

Usov, V.V., 1987, ApJ, 320, 333.

Zhang, B., and Harding, A.K., 2000, ApJ, 535, L51.

Zhangh, B., Harding, A.K., and Muslimov, A.G., 2000, ApJ, 531, L135.

APPENDIX A: PLASMA FUNCTION FOR ULTRARELATIVISTIC DISTRIBUTIONS.

We consider ultrarelativistic distributions with $\gamma^2 \gg 1$, $v = \sqrt{1 - 1/\gamma^2} = 1 - 1/2\gamma^2$. Writing $y = 1/2\gamma^2$ and $x = z - 1$, one obtains the following approximation for the dielectric function:

$$\epsilon_s = \frac{\omega_s^2}{k^2} W_s, \quad W_s = \int \frac{df_s}{y + x}, \quad \epsilon = 1 + \epsilon_p + \epsilon_b, \quad \text{(A1)}$$

where f_s is regarded as a function of y, and $df = (df/dy)dy$. The normalization condition is $\int f dy/(2y)^{3/2} = 1$. The corresponding dispersion relation is

$$K^2 = \frac{x(x+2)W}{(x-x_0)(x+x_0+2)}, \quad W = W_p + \alpha W_b, \quad \text{(A2)}$$

with $K = k/\omega_p$, $\alpha = n_b/n_p$, and $x_0 = 1/\cos\theta - 1$. In what follows we restrict our consideration to forward propagating waves, $K > 0$, so that W is well defined for $\text{Im}\, x > 0$ and should be analytically continued to $\text{Im}\, x \leqslant 0$. Waves with $\text{Re}\, x < 0$ are subluminal ($\omega/k < 1$) and those with $\text{Re}\, x > 0$ are superluminal.

APPENDIX B: DIRECT REFRACTIVE CONVERSION.

Consider growing waves approaching the point $\omega_c = \omega$ where the growth rate vanishes. As shown above (see Fig. 5), the only
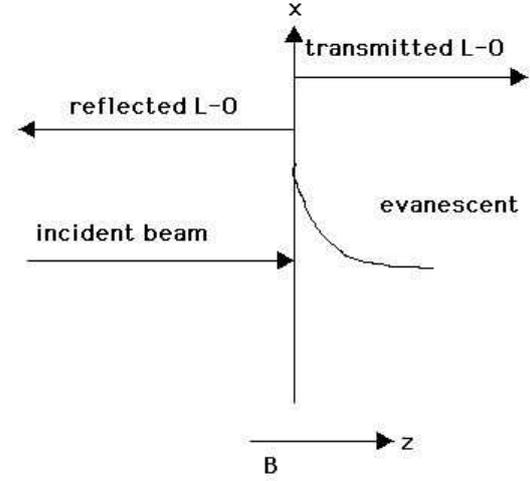

Figure B1. Matching at the critical point $\omega_c = \omega$.

mode which can propagate outwards (forward) beyond this point is the L-O mode. Just before this point the only backward propagating mode is also the L-O mode. Thus, at the point $\omega_c = \omega$ the incident beam mode (ω, k_{ix}, k_{iz}) and backward L-O (ω, k_{bx}, k_{bz}) mode should match the forward L-O mode (ω, k_{fx}, k_{fz}) and the evanescent mode (ω, k_{ex}, k_{ez}). The system is shown in Figure B1. We describe the evanescent mode by putting $k_{ez} = -i\kappa$ and further $\kappa \to \infty$. The electric and magnetic fields must be continuous at the matching point. It follows that only nonzero components are E_x, E_z, B_y. One also has $k_{ix} = k_{bx} = k_{fx} = k_{ex}$. For all these modes one has

$$\frac{E_x}{E_z} = \frac{k_x k_z}{k_z^2 - \omega^2}, \quad B_y = z E_x, \quad \text{(B1)}$$

with $z = \omega/k_z$. There is an evanescent mode for $|k_z| \to \infty$ which has $B_y = E_x = 0$, so that it should not be taken into account in the continuity conditions for B_y and E_x, and continuity of E_z always can be satisfied. Thus, one has

$$E_{ix} + E_{bx} = E_{fx}, \quad z_i E_{ix} - |z_b| E_{bx} = z_f E_{fx}, \quad \text{(B2)}$$

where we take into account that for the backward wave $z_b < 0$. Eq. (B2) gives

$$\frac{E_{fx}}{E_{ix}} = \frac{z_i + |z_b|}{z_f + |z_b|} \quad \text{(B3)}$$

In our case $z_i \approx z_f \approx |z_b| \approx 1$, so that $E_{fx} \approx E_{ix}$, that is, the conversion of the unstable mode into the forward propagating L-O mode is almost lossless.