Plasma Phys. Control. Fusion **45** (2003) 523–534 PII: S0741-3335(03)52128-5

Pulsar emissions

D B Melrose

School of Physics, University of Sydney, NSW 2006, Australia

Received 8 August 2002, in final form 12 November 2002
Published 26 March 2003
Online at stacks.iop.org/PPCF/45/523

Abstract
The standard model for radio pulsars implies that the 'pulsar plasma' in the source region is a highly relativistic, strongly magnetized, one-dimensional, electron–positron pair plasma. Relativistic quantum effects lead to a correction of first order in $\hbar$ to the linear response of a pulsar plasma due to its spin dependence, but this is small for highly relativistic particles. Wave dispersion in the radio range involves the X-mode, Alfvén mode and L–O mode. It is argued that the most plausible radio emission mechanism involves a nonresonant beam instability due to a beam mode that couples to the L–O mode. It is suggested that 'orthogonal modes' in the observed radio emission arise from mode coupling at sharp gradients at the edges of ducts, and that the observed circular polarization is associated with gyrotropy (a difference in the electron and positron distributions) and it develops as a propagation effect.

1. Introduction

Since the first three pulsars were discovered in 1967, the known number of radio pulsars has increased to well over a thousand (Taylor *et al* 1993, Lyne *et al* 2000). With such a large number of examples, the statistical properties of pulsars are well established. A pulsar is a rotating, magnetized neutron star (mass of order a solar mass and radius $R_* \sim 10\,\text{km}$) formed in a supernova explosion. About 1% are associated with known supernova remnants and a few (the 'Crab' in particular) with historical supernovae. After formation, the rotation period, P, of the neutron star increases (from $P \lesssim 10\,\text{ms}$) to $P \gtrsim 1\,\text{s}$; when the pulsar ceases to radiate (it 'dies'). The slowing down, measured by the period derivative, $\dot{P}$, is due to a highly relativistic wind, and is modelled by analogy with magnetic dipole radiation, which provides the standard estimates of both the (polar) magnetic field, $B_* \propto (P\dot{P})^{1/2}$, and of the (spin-down) age of the pulsar $\propto P/\dot{P}$. Most ('ordinary') pulsars have surface magnetic fields of $10^7\,\text{T} \lesssim B_* \lesssim 10^9\,\text{T}$, periods $0.1\,\text{s} \lesssim P \lesssim 1\,\text{s}$, and ages $\lesssim 10^7\,\text{yr}$. A subclass of millisecond pulsars (MSPs) consists of previously dead pulsars spun up in binary systems. MSPs have shorter periods, $1\,\text{ms} \lesssim P \lesssim 10\,\text{ms}$ and weaker fields, $B \sim 10^6\,\text{T}$. Most pulsars are observed only in the radio range, with the radio power a tiny fraction of the spin-down power. Around 1% of pulsars are known to emit high energy photons with a total power a significant fraction of the spin-down

0741-3335/03/050523+12$30.00

power (Thompson *et al* 1996). Statistically, the important criterion for high energy emission appears to be the strength of the magnetic field, $B_{\rm L} \sim B_*(R_*/r_{\rm L})^3$, at the light cylinder radius $r_{\rm L} = 2\pi c/P$. Field lines that extend beyond $r = r_{\rm L}$ are regarded as open, and for a dipolar magnetic field (angle $\theta = 0$ is the dipole axis), the locus of the open field lines defines a polar-cap region, of angular radius $\theta_{\rm pc} = (R_*/r_{\rm L})^{1/2}$ at the stellar surface. The pulsar magnetosphere is conventionally regarded as composed of the two polar-cap regions, both with the same sign for the co-rotation charge density, a co-rotating region with closed field lines and the opposite sign for the co-rotation charge density, and a wind region at $r > r_{\rm L}$.

The electrodynamics of pulsars involve the interplay between two effects. Suppose the pulsar is modelled as a rotating (angular velocity $\boldsymbol{\Omega}$, $\Omega = 2\pi/P$) magnetic dipole. In the absence of any plasma in the pulsar magnetosphere, the vacuum field implies a potential $\Phi = -(\Omega B_* R_*^5/3r^3)P_2(\cos\theta)$, where P_n is a Legendre polynomial, which implies a potential difference $\Delta\Phi \sim \Omega^2 R_*^3 B_*/c$ across the polar cap, and a parallel electric field $E_\parallel = \boldsymbol{E}\cdot\boldsymbol{B}/B = (2\Omega B_* R_*^5/r^4)\cos^3\theta/(3\cos^2\theta - 1)^{1/2}$. In the presence of plasma, the vacuum $E_\parallel$ is neutralized, and a co-rotation field is set up, implying the 'Goldreich–Julian' charge density $en_{\rm GJ} = -\varepsilon_0\,{\rm div}(\boldsymbol{\Omega}\times\boldsymbol{x})$, given explicitly by

$$n_{\rm GJ} = -\frac{\varepsilon_0\boldsymbol{\Omega}\cdot\boldsymbol{B}}{e} = (7\times10^{17}\,{\rm m}^{-3})\left(\frac{B}{10^8\,{\rm T}}\right)\left(\frac{P}{0.1\,{\rm s}}\right)^{-1}\cos\theta. \tag{1}$$

There is no adequate local source of charge to maintain the charge density at the Goldreich–Julian value throughout the polar cap, and all models imply that a parallel electric field of the order of the vacuum value necessarily arises, but the details are model-dependent (e.g. Ruderman and Sutherland (1975), Arons (1983), Muslimov and Tsygan (1992), Mestel and Shibata (1994)). The conventional picture is that this $E_\parallel$ accelerates 'primary' particles, with number density $n_{\rm p} \sim n_{\rm GJ}$, to a Lorentz factor $\gamma_{\rm p} \sim e\Delta\Phi/mc^2$. Due to curvature emission or inverse Compton emission, the primary particles emit high energy photons that decay into pairs at a pair-formation front (PFF). The region below the PFF is referred to as the 'inner gap' to distinguish it from an 'outer gap', also with $E_\parallel \neq 0$, that occurs on field lines for which $n_{\rm GJ}$ passes through zero. At the PFF a pair cascade results in a 'secondary' pair plasma (e.g. Daugherty and Harding (1983), Zhang and Harding (2000)), with number density $n_{\rm s} = Mn_{\rm p}$ and Lorentz factor $\gamma_{\rm s} \sim \gamma_{\rm p}/M$, with a multiplicity $M \gg 1$ that may be much lower ($\sim$10) than previously thought (Hibschman and Arons 2001). Above the PFF, $E_\parallel$ is assumed to be neutralized. The outflowing secondary pairs constitute the 'pulsar plasma' from which the radio emission is assumed to originate; these pairs become the pulsar wind at $r > r_{\rm L}$. In the superstrong B, gyromagnetic emission causes the secondary pairs to radiate away all their perpendicular energy ($p_\perp \to 0$), so that they become a one-dimensional distribution. The current associated with the net charge outflow through the inner gap must return to the star: in one model (Shibata 1997) it closes across field lines, with the $\boldsymbol{J}\times\boldsymbol{B}$ force accelerating the wind (e.g. Okamoto (2002)) and providing the slowdown torque, which must operate at $r \sim r_{\rm L}$ on average (Holloway 1977), and returns to the star through the outer gap.

The radio emission from pulsars has a relatively steep frequency spectrum, extending from $\lesssim$100 MHz to a high-frequency cut-off $\sim$3–30 GHz. It is intrinsically very bright (brightness temperature $T_{\rm B} \gtrsim 10^{25}$ K), and highly (mainly linearly) polarized (e.g. Han *et al* (1998), Crawford *et al* (2001)). The radio emission is thought to originate in the secondary pair plasma well above the PFF, at $\sim$10–100R_* (Cordes 1978). There is observational evidence for 'core' and 'conal' radio emissions, with different spectral characteristics (Rankin 1983), but these difference are probably not due to different emission mechanisms for core and conal components (e.g. Kramer *et al* (1999)). The simplest assumption is that a single radio emission mechanism operates in core and conal components and in both ordinary pulsars and

MSPs. The frequency of the radio emission is plausibly associated with the plasma frequency corresponding to $n_{\rm GJ}$, which depends only on the ratio B/P. The high-energy emission, which probably originates from the outer gap (Romani 1996), is not discussed further in this paper.

In section 2 the linear response of a pulsar plasma is discussed, with emphasis on relativistic plasma dispersion functions (RPDFs). In section 3 the results are used to describe wave dispersion in a pulsar plasma. The most widely favoured type of pulsar emission mechanism is discussed in section 4; this involves a beam instability and the emphasis in section 4 is on the nature of this instability. In section 5, implications of the observed properties of the polarization of pulsar radio emission are discussed, including the role of propagation through inhomogeneous plasma leading to a mixture of two orthogonal modes, and gyrotropy and cyclotron absorption imposing some circular polarization (CP).

2. The response of a pulsar plasma

The pulsar plasma from which the radio emission originates is assumed to be a one-dimensional, strongly magnetized, pair plasma with a highly relativistic beaming motion, a highly relativistic spread in momenta and significant differences in the electron and positron distributions. The linear response tensor for such a pulsar plasma is traditionally treated using a relativistic theory ignoring quantum effects. Other conventional simplifying approximations are identical electron and positron distributions ('non-gyrotropic' approximation), and $B \to \infty$ ('infinite-B' approximation) (Arons and Barnard 1986). In the treatment here, relativistic quantum effects are discussed, and both gyrotropy and a finite B are included explicitly.

2.1. Quantum effect in the response for a one-dimensional pair plasma

Conventionally, quantum effects are ignored in treating dispersion in a pulsar plasma. The most familiar quantum effect is the quantum recoil, and its neglect is well justified in the radio range. A second quantum, called 'anharmonicity' in the Russian literature, results from the quantization of the gyromagnetic motion, involving the ratio $B/B_{\rm c} = \hbar eB/m^2c^2 = B/1.4 \times 10^9$ T), which is not necessarily small in pulsars. An electron has energy eigenvalues $\varepsilon_n(p_z) = [m^2c^4(1+2nB/B_{\rm c})+p_z^2c^2]^{1/2}$, where n labels the Landau levels. The anharmonicity is that the frequency of cyclotron transitions, in which n changes by unity, depends on n. In a pulsar plasma the transition $n = 0 \to 1$ is the only one usually considered important, and the non-quantum approximation to the transition frequency needs to be corrected for pulsars with large $B/B_{\rm c}$. A third quantum effect is the spin. It is known that for unpolarized electrons (equal populations with $s = \pm 1$) the response tensor is an even function of $\hbar$, but this is not the case for a pulsar plasma. A spin-dependence arises because the ground state has a unique spin ($n = 0$, $s = -1$), whereas the other energy eigenvalues are doubly degenerate (spins $s = \pm 1$). The (Hermitian part of the) response tensor for an unpolarized electron gas, that is, with equal populations of the two spin states, is known to be an even function of $\hbar$, but in a spin-dependent system the first quantum correction is of order $\hbar$; it is of interest to know whether the correction of order $B/B_{\rm c}$ to the classical response is important in a pulsar plasma. A fourth quantum effect arises from dispersion due to one-photon pair creation, whose effect on the wave dispersion needs clarification in view of a result given by Svetozarova and Tsytovich (1962). A fifth quantum effect is the vacuum polarization, which is not discussed here.

The one-dimensional assumption corresponds to all the electrons ($\epsilon = 1$) and positrons ($\epsilon = -1$) being in their ground state. In a quantum description the distributions are described by their occupation numbers $n^\epsilon(p_z)$, where p_z is the physical momentum along $\boldsymbol{B}$, normalized such that the number densities are $n^\epsilon = [eB/(2\pi\hbar)^2] \int \mathrm{d}p_z\, n^\epsilon(p_z)$. Using this quantum-mechanical notation one can rewrite the classical (nonquantum) expression for the dielectric tensor for a one-dimensional pair plasma (e.g. Lyutikov (1998), Gedalin *et al* (1998), Melrose *et al* (1999)) in terms of the occupation number:

$$K_{ij}(\omega,\boldsymbol{k}) = \delta_{ij} + \sum_{\epsilon=\pm1} \frac{\mu_0 c^2 e^3 B}{m(2\pi\hbar\omega)^2} \int_{-\infty}^{\infty} \mathrm{d}p_z\, n^{\epsilon}(p_z) d_{ij}^{\epsilon}, \tag{2}$$

with, for $\boldsymbol{B}$ along the z-axis and $\boldsymbol{k} = (k_\perp, 0, k_z)$ in the x–z plane,

$$d_{11} = d_{22} = -\frac{(\omega - k_z v_z)^2}{(\omega - k_z v_z)^2 - (\Omega/\gamma)^2}, \qquad d_{13} = d_{31} = -\frac{\gamma v_z k_\perp(\omega - v_z k_z)}{(\omega - k_z v_z)^2 - (\Omega/\gamma)^2},$$
$$d_{12} = -d_{21} = \frac{\mathrm{i}\epsilon(\omega - k_z v_z)\Omega/\gamma}{(\omega - k_z v_z)^2 - (\Omega/\gamma)^2}, \qquad d_{23} = -d_{32} = -\frac{\mathrm{i}\epsilon v_z k_\perp \Omega}{(\omega - k_z v_z)^2 - (\Omega/\gamma)^2}, \tag{3}$$
$$d_{33} = -\frac{\omega^2}{\gamma^2(\omega - k_z v_z)^2} - \frac{v_z^2 k_\perp^2}{(\omega - k_z v_z)^2 - (\Omega/\gamma)^2},$$

with $\Omega = eB/m$, $\gamma = (1 + p_z^2/m^2c^2)^{1/2}$, $v_z = p_z/\gamma m$. The simplifying assumptions mentioned above are apparent from (3): the $\epsilon = \pm1$ contributions to d_{12} and d_{23} cancel in the non-gyrotropic approximation, and $\Omega \to \infty$ implies that the only nonzero term is d_{33} in the infinite-B approximation.

The response tensor derived by Svetozarova and Tsytovich (1962) includes relativistic quantum effects in the long-wavelength limit, $\boldsymbol{k} \to 0$. In a notation similar to (3) their result is (after correcting their 12-term)

$$K_{11} = K_{22} = 1 - \frac{\mu_0 e^3 B}{(2\pi\hbar)^2} \int_{-\infty}^{\infty} \frac{\mathrm{d}p_z}{\varepsilon_1} \left\{ \frac{(\varepsilon_1 - \varepsilon_0)^2}{(\hbar\omega)^2 - (\varepsilon_1 - \varepsilon_0)^2} - \frac{(\varepsilon_1 + \varepsilon_0)^2}{(\hbar\omega)^2 - (\varepsilon_1 + \varepsilon_0)^2} \right\} \bar{n},$$
$$K_{12} = -K_{21} = \frac{\mathrm{i}\mu_0 e^3 B}{(2\pi\hbar)^2} \int_{-\infty}^{\infty} \frac{\mathrm{d}p_z}{\varepsilon_1} \left\{ \frac{\hbar\omega(\varepsilon_1 - \varepsilon_0)}{(\hbar\omega)^2 - (\varepsilon_1 - \varepsilon_0)^2} + \frac{\hbar\omega(\varepsilon_1 + \varepsilon_0)}{(\hbar\omega)^2 - (\varepsilon_1 + \varepsilon_0)^2} \right\} n^{\mathrm{d}}, \tag{4}$$
$$K_{33} = 1 - \frac{\mu_0 e^3 B}{(2\pi\hbar)^2} \int_{-\infty}^{\infty} \frac{\mathrm{d}p_z}{\varepsilon_0} \frac{4m^2c^4}{(\hbar\omega)^2 - 4\varepsilon_0^2} \bar{n},$$

with $\bar{n} = n^+ + n^-$, $n^{\mathrm{d}} = n^+ - n^-$, and where arguments $(\omega, \boldsymbol{k})$ and (p_z) are omitted for simplicity in writing. The final expressions inside the curly brackets in the first two of equations (4) describe dispersion due to one-photon pair creation, and it is a surprising fact that these terms cannot be neglected in deriving the correct nonquantum limit. Specifically, these terms are required to ensure that the integrands go to zero for $\omega \to 0$.

For the $i, j = 1, 2$ terms, the long-wavelength limit of (3) and the small-B/B_{c} limit of (4) are equivalent. The anharmonicity leads to a correction to the cyclotron frequency in (3) when one expands in $B/B_{\mathrm{c}} = \hbar\Omega/mc^2$,

$$\varepsilon_1 - \varepsilon_0 = \frac{\hbar\Omega}{\gamma}\left(1 - \frac{\hbar\Omega}{8\gamma^2 mc^2} + \cdots\right), \tag{5}$$

in (4). Thus, the anharmonicity modifies (from the nonquantum value) the frequency at which the cyclotron resonance occurs. This correction appears as a first-order correction in B/B_{c} in the contribution to the response due to particles in their ground state. However, the correction is small for highly relativistic particles, $\gamma \gg 1$, and it is likely to be important in a pulsar plasma only for mildly relativistic particles in fields $B \sim B_{\mathrm{c}}$.

The dispersion in the K_{33} term in (4) is at best misleading: it describes dispersion associated with one-photon absorption into the lowest Landau level for the electron and positron, and this is not relevant to dispersion in the classical limit. To reproduce the part of the d_{33} in (3) that is independent of Ω one needs to retain a term that is absent from (4). This term corresponds to virtual transitions between p_z and $p'_z = p_z - \hbar k_z$, both in the ground Landau state. The exact relativistic quantum counterpart of this term is (Melrose and Weise 2002)

$$K_{33} = 1 + \frac{\mu_0 c^4 e^3 B e^{-\hbar k_\perp^2/2eB}}{4\pi\hbar\omega^2} \int_{-\infty}^{+\infty} \frac{\mathrm{d}p_z}{2\pi\hbar} \left(1 + \frac{p_z p'_z c^2 - m^2 c^4}{\varepsilon_0(p_z)\varepsilon_0(p'_z)}\right) \frac{\bar{n}_0(p_z) - \bar{n}_0(p'_z)}{\hbar\omega - \varepsilon_0(p_z) + \varepsilon_0(p'_z)}. \quad (6)$$

The long-wavelength limit of (6) is artificially set to zero in (4), due to $\bar{n}_0(p_z) - \bar{n}_0(p'_z) \to 0$ for $k_z \to 0$, whereas a nonzero result is obtained by setting $\bar{n}_0(p_z) - \bar{n}_0(p'_z) \to \hbar k_z\, \mathrm{d}n(p_z)/\mathrm{d}p_z$ and partially integrating to obtain the nonquantum expression included in d_{33}. Thus, the long-wavelength limit to K_{33} in (4) is not useful in describing the low-frequency dispersion in a pulsar plasma, and the appropriate approximation to the exact expression (6) reproduces the classical result (3).

2.2. *Relativistic plasma dispersion functions*

The dielectric tensor in the general case, with quantum effects neglected, may be written in terms of three RPDFs. Let the one-dimensional distribution function, $f(u)$, be normalized by $\int \mathrm{d}u\, f(u) = 1$, with $u = p_z/mc = \gamma v$, $v = v_z/c$, and let the phase speed be written as $z = \omega/k_z c$. A definition of the three RPDFs is (Melrose *et al* 1999)

$$W(z) = \left\langle \frac{1}{\gamma^3(v-z)^2} \right\rangle, \qquad R(z) = \left\langle \frac{1}{\gamma(v-z)} \right\rangle, \qquad S(z) = \left\langle \frac{1}{\gamma^2(v-z)} \right\rangle, \quad (7)$$

with $\langle K \rangle = \int \mathrm{d}u\, K f(u)$. Then, one has (Gedalin *et al* 1998, Melrose *et al* 1999)

$$K_{11}(k) = K_{22}(k) = 1 - \frac{\omega_\mathrm{p}^2}{\omega^2}\frac{1}{1+y^2}\left[\left\langle\frac{1}{\gamma}\right\rangle + \frac{(z-z_+)^2 R(z_+) - (z-z_-)^2 R(z_-)}{z_+ - z_-}\right],$$

$$K_{33}(k) = 1 - \frac{\omega_\mathrm{p}^2}{\omega^2}\left\{z^2 W(z) + \frac{\tan^2\theta}{1+y^2}\left[\left\langle\frac{1}{\gamma}\right\rangle + \frac{z_+^2 R(z_+) - z_-^2 R(z_-)}{z_+ - z_-}\right]\right\},$$

$$K_{12}(k) = -K_{21}(k) = -\mathrm{i}\eta\frac{\omega_\mathrm{p}^2}{\omega^2}\frac{y}{1+y^2}\left[\frac{(z-z_+)S(z_+) - (z-z_-)S(z_-)}{z_+ - z_-}\right], \quad (8)$$

$$K_{13}(k) = K_{31}(k) = -\frac{\omega_\mathrm{p}^2}{\omega^2}\frac{\tan\theta}{1+y^2}\left[-\left\langle\frac{1}{\gamma}\right\rangle + \frac{(z-z_+)z_+ R(z_+) - (z-z_-)z_- R(z_-)}{z_+ - z_-}\right],$$

$$K_{23}(k) = -K_{32}(k) = \mathrm{i}\eta\frac{\omega_\mathrm{p}^2}{\omega^2}\frac{y\tan\theta}{1+y^2}\left[\frac{z_+ S(z_+) - z_- S(z_-)}{z_+ - z_-}\right],$$

with $y = \Omega/k_\parallel c$, $z_\pm = [z \pm y(1+y^2-z^2)^{1/2}]/(1+y^2)$, $\eta = (n^+ - n^-)/(n^+ + n^-)$, and where θ is the angle between $\boldsymbol{k}$ and $\boldsymbol{B}$. (The term $\langle 1/\gamma \rangle$ was omitted in the expression for K_{13} written down by Melrose *et al* (1999), but was included correctly in all their calculations.) The anharmonicity cannot be included in (8) in an elementary manner because the frequency shift from the classical value, cf (5) depends on γ.

2.3. *One-dimensional Jüttner distribution*

The pulsar plasma is streaming outwards relativistically, and its momentum spread in its rest frame is expected to be highly relativistic. A thermal (Jüttner) distribution in the rest frame is favoured by recent calculations of pair creation (Arendt and Eilek 2002). The one-dimensional Jüttner distribution is $f(u) = \mathrm{e}^{-\rho\gamma}/2K_1(\rho)$, where $\rho = mc^2/T$ is the inverse temperature in units of the rest energy and K_n is a modified Bessel (Macdonald) function of order n. The analytic properties of $W(z)$ for this case have been discussed by Asseo and Riazuelo (2000) and by Kennett *et al* (2000), who related it to the RPDF, $T(z, \rho)$ explored in detail by Godfrey *et al* (1975):

$$W(z) = \frac{1}{2K_1(\rho)} \frac{\partial T(z, \rho)}{\partial z}, \qquad T(z, \rho) = \int_{-1}^{1} \mathrm{d}v \frac{\mathrm{e}^{-\rho\gamma}}{v - z}. \tag{9}$$

2.4. Properties of $W(z)$

Except near the cyclotron resonance the dispersion in (8) is determined primarily by the K_{33} term, and more specifically by the function $z^2W(z)$. In a highly relativistic plasma, $\langle\gamma\rangle \gg 1$, one has $\lim_{z\to 0} z^2W(z) = 0$, with $z^2W(z)$ negative for $z \ll 1$, reaching a maximum negative value just below $z = 1 - 1/\langle\gamma\rangle$, and then rising to a very high, sharp peak near $z = 1 - 1/\langle\gamma\rangle$. The height and width of the peak are sensitive to the detailed form of the distribution function; however, damping is strongest in this range of z, and the detailed shape of the peak is not important in determining the properties of weakly damped waves. Above the peak, $z^2W(z)$ decreases and crosses $z = 1$ with $W(1) = 4\langle\gamma\rangle$. There is no damping for $z \geqslant 1$, with $W(z) \approx \langle\gamma^{-3}\rangle/(z^2 - 1)$ for $z \gg 1$. The cut-off frequency, ω_c, often interpreted as the relativistic plasma frequency, depends on $\lim_{z\to\infty} z^2W(z) = \langle\gamma^{-3}\rangle \sim 1/\langle\gamma\rangle$ in the form $\omega_c = \omega_p\langle\gamma^{-3}\rangle^{1/2}$.

The shape of $z^2W(z)$ is illustrated in figure 1 for Jüttner distributions, showing how it changes with decreasing $\rho = mc^2/T$. In the familiar nonrelativistic case, illustrated for $\rho = 100$, there is a negative region near $z = 0$, associated with Debye shielding, a positive peak near the thermal speed, $z \sim 1/\rho^{1/2}$, and a region where damping is weak for $z \gg 1/\rho^{1/2}$. As the temperature becomes more relativistic, the peak moves closer to $z = 1$ and becomes much more prominent. In the highly relativistic regime, $\rho \ll 1$, this peak becomes so sharp that a different representation is needed to illustrate its properties clearly (cf Melrose *et al* (1999)). Wave dispersion in a relativistic plasma is not particularly sensitive to the functional form of $f(u)$: although the peak is sensitive to $f(u)$ it is in the region of relatively strong (Landau) damping. In the region $z < 1$ above this peak there is a region where damping is weak, and it is zero for $z \geqslant 1$. In this region the properties of $W(z)$ are sensitive only to moments of the distribution specifically to $\langle\gamma\rangle$ for $z \approx 1$ and $\langle\gamma^{-3}\rangle$ for $z \gg 1$.

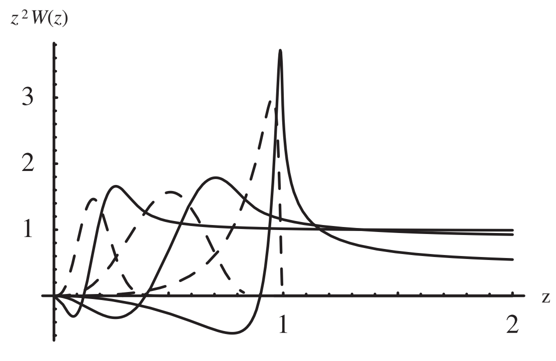

Figure 1. The real (——) and imaginary (- - - -) parts of the function $z^2W(z)$ are plotted for a one-dimensional Jüttner distribution with temperature $T = 0.01mc^2$ (leftmost curves), $T = 0.1mc^2$ (middle curves) and $T = mc^2$ (rightmost curves). For $T \gg mc^2$ the peak just below $z = 1$ becomes higher, sharper and closer to $z = 1$ ($z - 1 \approx (T/mc^2)^{-1}$).

3. Wave dispersion in a pulsar plasma in its rest frame

The simplest useful approximation for describing the dispersion of low-frequency waves in a pulsar plasma is of the form (8) for a non-gyrotropic plasma ($\eta = 0$) expanded to first order in $1/y^2$. This has three independent nonzero components (Melrose *et al* 1999):

$$K_{11} = a + \frac{1-b}{z^2}, \qquad K_{13} = -\frac{1-b}{z^2}\tan\theta,$$
$$K_{33} = 1 - \frac{\omega_p^2}{\omega^2} z^2 W(z) - \frac{1-b}{z^2}\tan^2\theta, \tag{10}$$

with $a = 1 + 1/v_A^2$, $b = 1 - \Delta v^2/v_A^2$, where in $v_A^2 = \langle\gamma^{-1}\rangle\Omega^2/\omega_p^2$ the plasma frequency, $\omega_p = [\mu_0 c^2 e^2 \bar{n}/m]^{1/2}$, is defined in terms of the actual number density, $\bar{n}$ of electrons plus positrons, and where $\Delta v^2 = 1 - \langle\gamma^{-1}\rangle/\langle\gamma\rangle$ characterizes the spread in velocities.

The full dispersion equation factorizes into two dispersion equations

$$a - \frac{b - \sin^2\theta}{z^2} = 0, \qquad \frac{\omega^2}{\omega_p^2} = \frac{z^2 W(z)(z^2 - z_A^2)}{z^2 - z_A^2 - b\tan^2\theta}, \tag{11}$$

with $z_A^2 = b/a$, which describe the X-mode and the L–O and Alfvén modes, respectively. The X-mode is magnetosonic-like, with phase speed strictly subluminal and with strictly perpendicular (to $\boldsymbol{B}$) polarization.

For $\theta = 0$ equations (11) imply $z^2 = z_A^2$, corresponding to the parallel Alfvén mode, and $\omega^2/\omega_p^2 = z^2 W(z)$, corresponding to a parallel Langmuir (L) mode. The L-mode has a cut-off, $z \to \infty$, at $\omega_c = \omega_p\langle\gamma^{-3}\rangle^{1/2}$ (which continues to apply for arbitrary θ); it crosses $z = 1$ at $\omega_1 = \omega_p[\langle\gamma\rangle(1 + \Delta v^2)]^{1/2} \sim \omega_c\langle\gamma\rangle$, it crosses the Alfvén line $z = z_A$ at $\omega_{co} = \omega_p z_A [W(z_A)]^{1/2} > \omega_1$, it has a maximum frequency slightly above ω_{co}, and is strongly damped at smaller z. As illustrated in figure 3, for $\theta \neq 0$ the modes reconnect, with the L–O mode joining onto the parallel-Alfvénic for $\omega > \omega_{co}$, and the Alfvén mode joining onto the strongly damped portion.

Dispersion curves for these modes (for $\theta \neq 0$) are illustrated in figure 2, where $\langle\gamma\rangle \approx 2$ is chosen to be unrealistically small for illustrative purposes. As $\langle\gamma\rangle$ is increased, the dispersion

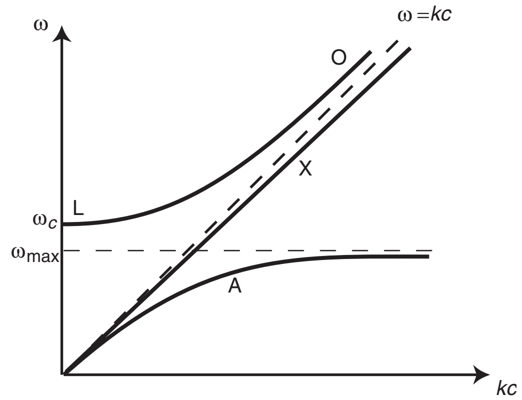

Figure 2. Dispersion curves for a (non-gyrotropic) pulsar plasma at frequencies well below the cyclotron frequency for a slightly oblique angle of propagation. The X-mode is strictly transverse. The Alfvén (A) mode has a maximum frequency, ω_{max}, and is heavily damped as this frequency is approached. The L–O mode is approximately longitudinal (L) near its cut-off, ω_c (the 'relativistic plasma frequency'), and approximately transverse (O) at higher frequencies.

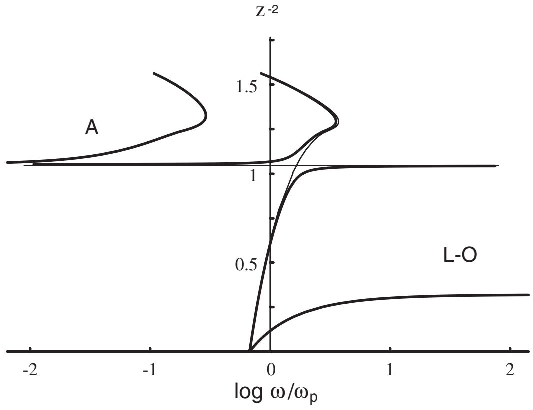

Figure 3. Dispersion curves for the Alfvén (A) and L–O modes are plotted as a function of $1/z^2$ for a bell-type distribution $f(u) \propto (u_m^2 - u^2)^3$ with $v_m = 0.9$, and with $z_A = 0.95$, for $\theta = 0$ (faint continuous curve and faint horizontal line with $z = z_A$), $\theta = 0.1$ (central pair of curves) and $\theta = 1$ (outer pair of curves). Note that for sufficiently small θ (e.g. $\theta = 0.1$ here) the O mode is subluminal ($z_A < z < 1$) at high frequencies, and superluminal for higher θ (e.g. $\theta = 1$ here).

curves are drawn out along the light line $\omega = k_z c$, such that the O-mode approaches (crosses for sufficiently small θ) the light line at a much higher frequency, specifically, near $\omega_{co} \approx \omega_c \langle\gamma\rangle$.

An important detail is that for $0 < \theta < \theta_{max} = (1 + \Delta v^2)^{1/2}/v_A$ there is a subluminal portion of the L–O mode, with $z_A \leqslant z < 1$ over a finite frequency range that shrinks to zero at $\theta = \theta_{max}$. This is illustrated in figure 3. In particular, particles with a given speed $v > z_A$ can resonate with L–O mode waves with sufficiently small θ over a small range of frequencies above ω_1.

4. Beam instability and pulsar radio emission

Pulsar emission must be due to an instability, and a beam instability is plausible. The foregoing discussion implies that the conditions for beam instability are satisfied for a beam with speed $v_b > z_A$ (in the rest frame of the background plasma). Moreover, the waves are in the L–O mode, which extends to arbitrarily large ω. This suggests an emission mechanism in which the primary beam propagates through the secondary pair plasma causing L–O mode waves to grow, in the small allowed range of the θ–ω space, and these waves escape to produce the observed radiation. However, this simple model encounters overwhelming difficulties: the growth rate is far too small (Larroche and Pellat 1987) for plausible parameters, the frequency range is too high to account for the lowest frequencies observed (Melrose and Gedalin 1999), and the polarization is transverse linear, orthogonal to $\boldsymbol{B}$, whereas the observed emission is a mixture of orthogonal modes (Stinebring *et al* 1994), which are sometimes elliptical (Radhakrishnan and Rankin 1990). How can these difficulties be overcome? An obvious requirement is a faster instability, and Alfvén waves are known to grow much faster than the L–O mode (Egorenkov *et al* 1983, Lominadze *et al* 1986, Arons and Barnard 1986, Lyutikov 2000). However, these waves encounter a region of strong damping and a stop band (cf figure 2) along any prospective escape path, and some secondary process that converts them into X or L–O mode waves is required for the growth of Alfvén waves to constitute an emission mechanism. Direct growth of escaping L–O mode waves is the simplest emission mechanism and an important question is whether or not the growth can be fast enough for it to be viable.

4.1. The growth rate

Wave growth can be due to negative absorption, in a kinetic instability driven by a distribution with $\mathrm{d}f(u)/\mathrm{d}u > 0$, or due to a reactive instability involving an intrinsically growing wave mode (e.g. Melrose (1986)). The kinetic version of a beam instability applies when the growth rate, Γ, is less than the Doppler width of the growing waves, and the reactive instability when this inequality is reversed. In a highly relativistic plasma the Doppler spread is intrinsically small ($\Delta\omega \approx \omega\Delta v$ for $\omega \approx kc$), due to the bulk of the particles being concentrated very close to the speed of light with an intrinsically small ($\Delta v \approx \Delta\gamma/\gamma^3$). As a result, simple estimates imply that a growth rate of the order of the Doppler width is impossibly slow for highly relativistic particles. Consequently, effective instability must involve a reactive instability (Egorenkov *et al* 1983, Lominadze *et al* 1986). The Doppler spread is then unimportant in a reactive instability. This allows a major simplification: a cold-beam model suffices when treating reactive instabilities.

The simplest model for the instability involves a cold, relativistic, one-dimensional beam moving through a cold background one-dimensional plasma in the limit $B \to \infty$ ($v_{\mathrm{A}} \to \infty$). In the pulsar frame, the background plasma is assumed to be streaming with speed v_{p} and $\gamma_{\mathrm{p}} = (1 - v_{\mathrm{p}}^2)^{-1/2} \gg 1$. This model suffices for discussing the reactive instability provided that one adjusts the parameters of the background to ensure that the 'resonant' frequency, ω_{r}, corresponds to the frequency, $\approx\omega_{\mathrm{co}}$, at which the waves actually grow for a realistic background distribution. The dispersion equation for the L–O mode is

$$1 - \frac{\omega_{\mathrm{p}}^2}{\omega^2} z^2 W(z) = \frac{1}{z^2 - \tan^2\theta}, \qquad W(z) = \frac{1}{\gamma_{\mathrm{p}}^3 (z - v_{\mathrm{p}})^2} + \frac{n_{\mathrm{b}}/n_{\mathrm{p}}}{\gamma_{\mathrm{b}}^3 (z - v_{\mathrm{b}})^2}, \tag{12}$$

where $n_{\mathrm{b}}/n_{\mathrm{p}} \ll 1$ is the ratio of the densities in the beam and the background pair plasma. An instability corresponds to a complex solution of (12). In two limits the relevant complex solution implies a growth rate (Gedalin *et al* 2002)

$$\frac{\Gamma}{\omega} = \begin{cases} \left(\dfrac{n_{\mathrm{b}}}{n_{\mathrm{p}}}\right)^{1/2} \dfrac{1}{2\gamma_{\mathrm{p}}^{1/2}\gamma_{\mathrm{b}}^{3/2}}, & \text{for } \omega \ll \omega_{\mathrm{r}}, \\ \dfrac{\sqrt{3}}{2\sqrt{2}} \left(\dfrac{n_{\mathrm{b}}}{n_{\mathrm{p}}}\right)^{1/3} \dfrac{1}{2\gamma_{\mathrm{p}}\gamma_{\mathrm{b}}}, & \text{for } \dfrac{|\omega - \omega_{\mathrm{r}}|}{\omega_{\mathrm{r}}} \lesssim \dfrac{\Delta\omega_{\mathrm{r}}}{\omega_{\mathrm{r}}} \approx \left(\dfrac{n_{\mathrm{b}}}{n_{\mathrm{p}}}\right)^{1/3} \left(\dfrac{\gamma_{\mathrm{p}}}{\gamma_{\mathrm{b}}}\right)^{1/2}, \end{cases} \tag{13}$$

with $\omega_{\mathrm{r}} = 2\omega_{\mathrm{p}}\gamma_{\mathrm{p}}^{1/2}$ the 'resonant frequency'. The two regimes in (13) are referred to as nonresonant and resonant, respectively. Nonresonant growth occurs in a beam mode, whereas resonant growth occurs in the L–O mode of the pair plasma. The beam mode joins on smoothly to the L–O mode at the resonant frequency.

The growth rate is too small to be effective for the primary beam (Larroche and Pellat 1987) due to its very large γ_{b}. There are several suggestions as to how beams with much lower γ_{b} might form (Usov 1987, Lyubarskii and Petrova 2000), the simplest being that the secondary pair creation process favours lower γ (Hibschman and Arons 2001). It is assumed here that the instability is driven by a low-γ beam.

The conventional assumption that the growth is resonant leads to two serious difficulties (e.g. Melrose and Gedalin (1999)). First, the resonant frequency is too high: one can account for observed emission at $\lesssim$100 MHz only by invoking plasma parameters that are considered implausible. Second, the spatial gradient causes ω_{p} to change with height, r, and growth at a fixed frequency, ω, is restricted to a small range of heights over which ω_{p} changes by order $\Delta\omega_{\mathrm{r}}$. Both problems are alleviated by appealing to the nonresonant instability. In fact, due to its much greater growth path, the gain factor for growth is larger for the nonresonant instability than for the resonant instability (Gedalin *et al* 2002), implying that most of the energy should go into nonresonant waves. Nonresonant growth due to such a beam with a relatively low γ seems the most plausible candidate for the pulsar emission mechanism.

4.2. Orthogonal polarizations

The foregoing theory suggests that pulsar emission should be completely linearly polarized in the sense of the L–O mode. Although this is consistent with the characteristic sweep of the linear polarization across a pulse (Radhakrishnan and Coocke 1969), it is not consistent with two other important features of the polarization: orthogonal modes and CP.

In some pulsars the sweep of linear polarization is punctuated by jumps of 90° in the plane of polarization (Stinebring *et al* 1994). The interpretation assumes that the emission is a mixture of two orthogonal modes, with the difference in the intensities of the two modes changing sign at such a jump (e.g. McKinnon and Stinebring (2000)). One interpretation might be that the instability generates both L–O mode and X-mode waves, with similar growth rates. However, this is not possible without a major change in the theory. The X-mode polarization has no component along $\boldsymbol{B}$, so that there is no coupling to a beam; although gyrotropy (and possibly other effects) can modify the polarization, allowing the X-mode to grow, its growth rate remains much smaller than for the L–O mode.

4.3. Circular polarization

Many pulsars have significant CP (Radhakrishnan and Rankin 1990, Han *et al* 1998). It has been known since the late 1960s that the CP can be large in single pulses, and variable from one pulse to another. Recent results confirm this phenomenon (Karastergiou *et al* 2001), and unpublished data (Johnston (2002), private communication) comparing the statistics of Stokes V with $|V|$ (for at least five of 30 pulsars at 21 and 6 cm) show that $\langle |V| \rangle \gg |\langle V \rangle|$, confirming that the handedness changes sign frequently. Another feature recognized recently (von Hoensbroech and Lesch 1999) is that the CP increases to a high value at the highest observed frequencies (in about five of 200 pulsars observed, Wielebinski (2002), private communication).

Gyrotropy, due to different distributions of electron and positron (Kennett *et al* 2000, Gedalin *et al* 2001), causes the polarizations of the natural modes to become elliptical, and gyrotropy combined with relativistic streaming implies that the handedness of the CP of a natural mode reverses at $\theta \sim 1/\gamma_{\rm p}$. This provides the basis for a potential model for reversals in handedness, which should occur most frequently in pulsars in which the polarization is determined in a region where the condition $\theta \sim 1/\gamma_{\rm p}$ is satisfied. Gyrotropy also implies that cyclotron damping is different for electrons and positrons, and such cyclotron absorption should occur inside the light cylinder for a subset of pulsars (Luo and Melrose 2001). The observed increase in CP at the highest frequencies is plausibly due to cyclotron absorption. A test of this suggestion is that the pulsars for which this is observed should be the ones for which cyclotron absorption becomes important at the lowest heights (cf Luo and Melrose (2001)). A systematic investigation of CP and its possible interpretations, including these ideas, is in progress.

4.4. Mode coupling

The observed polarization of pulsar radio emission implies a mixture of two modes, and any instability strongly favours only one mode. Mode coupling can produce a mixture of the two modes (Petrova 2001). There is observational evidence suggesting that the plasma is concentrated into columns (Deshpande and Rankin 1997), and this has various theoretical implications (e.g. Asseo and Khechinashvili (2002)). A speculation is that if the emission is generated in underdense ducts, between the columns, reflection off the edges of the ducts can convert waves in a single mode into a mixture of the two modes in a much more effective way. Such a ducting process is thought to occur for radio waves in the solar corona (Duncan 1979), and it can account for the mixture of observed modes in this case, where the emission mechanism also operates only for a single mode (e.g. Melrose (1986)).

5. Discussion and conclusion

The interpretation of pulsar radio emission provides us with a challenge to adapt plasma concepts and theories developed for laboratory and space plasmas to an exotic regime. The pulsar plasma consists of one-dimensional, highly relativistic pairs in a superstrong magnetic field. The basic theory of wave dispersion in such a plasma is relatively well established, but there are several effects that need to be explored further. Gyrotropy, which is important for the interpretation of the observed CP, needs to be modelled in detail. The role of relativistic quantum effects is discussed in this paper, where it is shown that the anharmonicity and spin-dependence of the ground state combine to give a correction of first order in $B/B_{\rm c}$, which, however, is unimportant for highly relativistic particles due to a multiplicative factor $1/\gamma^2$. Relaxation of the one-dimensional approximation to include particles in higher Landau states is needed to discuss possible differences between millisecond and ordinary pulsars and possible radio emission from outer gaps.

A major outstanding problem in pulsar physics is a clear identification of the radio emission mechanism. Amongst the various suggested mechanisms, the most widely favoured involve beam instabilities, but these have been found either to grow too slowly and to occur at too high a frequency (for L–O mode waves that can escape) or to favour (Alfvén) waves which cannot escape. A recent suggestion that alleviates these difficulties involves nonresonant (reactive) instability of a beam mode that couples directly to the L–O mode wave allowing free escape of the radiation (Gedalin *et al* 2002). Alternative emission mechanisms, several of which have been proposed (curvature maser emission, linear acceleration emission, free electron maser emission, curvature-drift instability, anomalous Doppler instability), cannot be ruled out until it can be shown that a beam instability can account for the observed properties of pulsar radio emission. Evidence for ‘orthogonal modes’ seems to require that mode coupling occur as the radiation escapes. It is speculated here that this might occur due to reflection of the walls of underdense ducts, and detailed modelling is required to develop observation tests for this idea.

Acknowledgments

I thank Qinghuan Luo, Michael Gedalin, Jeanette Weise, Simon Johnston and Richard Wielebinski for useful discussions.

References

Arendt P N Jr and Eilek J A 2002 astro-ph/0207638
Arons J 1983 *Astrophys. J.* **266** 215
Arons J and Barnard J J 1986 *Astrophys. J.* **302** 120
Asseo E and Khechinashvili D 2002 *Preprint* (astro-ph/0203129)
Asseo E and Riazuelo A 2000 *Mon. Not. R. Astron. Soc.* **318** 983
Cordes J 1978 *Astrophys. J.* **222** 1006

Crawford F, Manchester R N and Kaspi V M 2001 *Astrophys. J.* **122** 2007
Daugherty J K and Harding A K 1983 *Astrophys. J.* **273** 761
Deshpande A A and Rankin J M 1997 *Astrophys. J.* **524** 1008
Duncan R A 1979 *Solar Phys.* **63** 389
Egorenkov V D, Lominadze D G and Mamramdze P G 1983 *Astrofizika* **19** 753
Gedalin M, Gruman E and Melrose D B 2001 *Mon. Not. R. Astron. Soc.* **325** 715
Gedalin M, Gruman E and Melrose D B 2002 *Phys. Rev. Lett.* **88** 121101
Gedalin M, Melrose D B and Gruman E 1998 *Phys. Rev.* E **57** 3399
Godfrey B B, Newberger B S and Taggart K A 1975 *IEEE Trans. Plasma Sci.* **PS-3** 60
Han J L, Manchester R N, Xu R X and Qiao G J 1998 *Mon. Not. R. Astron. Soc.* **300** 373
Hibschman J A and Arons J 2001 *Astrophys. J.* **560** 871
Holloway N J 1977 *Mon. Not. R. Astron. Soc.* **181** 9P
Karastergiou A *et al* 2001 *Astron. Astrophys.* **379** 270
Kramer M *et al* 1999 *Astrophys. J.* **526** 957
Kennett M P, Melrose D B and Luo Q 2000 *J. Plasma Phys.* **64** 333
Larroche O and Pellat R 1987 *Phys. Rev. Lett.* **59** 1104
Lominadze D G *et al* 1986 *Sov. J. Plasma Phys.* **12** 712
Luo Q and Melrose D B 2001 *Mon. Not. R. Astron. Soc.* **325** 187
Lyne A G *et al* 2000 *Mon. Not. R. Astron. Soc.* **312** 698
Lyubarskii Y E and Petrova S A 2000 *Astron. Astrophys.* **355** 406
Lyutikov M 1998 *Mon. Not. R. Astron. Soc.* **293** 447
McKinnon M M and Stinebring D R 2000 *Astrophys. J.* **529** 435
Melrose D B 1986 *Instabilities in Space and Laboratory Plasmas* (Cambridge: Cambridge University Press)
Melrose D B and Gedalin M 1999 *Astrophys. J.* **521** 351
Melrose D B, Gedalin M E, Kennett M P and Fletcher C S 1999 *J. Plasma Phys.* **62** 233
Melrose D B and Weise J I 2002 *Phys. Plasmas* **9** 4473
Mestel L and Shibata S 1994 *Mon. Not. R. Astron. Soc.* **271** 621
Muslimov A G and Tsygan A I 1992 *Mon. Not. R. Astron. Soc.* **255** 61
Okamoto I 2002 *Astrophys. J.* **573** L31
Petrova S A 2001 *Astron. Astrophys.* **378** 883
Radhakrishnan V and Cooke D J 1969 *Astrophys. J. Lett.* **3** 225
Radhakrishnan V and Rankin J M 1990 *Astrophys. J.* **352** 258
Rankin J M 1983 *Astrophys. J.* **274** 333
Romani R W 1996 *Astrophys. J.* **470** 469
Ruderman M A and Sutherland P G 1975 *Astrophys. J.* **196** 51
Shibata S 1997 *Mon. Not. R. Astron. Soc.* **287** 262
Stinebring D R *et al* 1994 *Astrophys. J. Suppl.* **55** 247
Svetozarova G I and Tsytovich V N 1962 *Izv. Vyssh. Uchebn. Zaved. Radiofiz.* **5** 658
Taylor J H, Manchester R N and Lyne A G 1993 *Astrophys. J. Suppl.* **88** 529
Thompson D J *et al* 1996 *Astrophys. J. Suppl.* **101** 259
Usov V V 1987 *Astrophys. J.* **320** 333
von Hoensbroech A and Lesch H 1999 *Astron. Astrophys.* **342** L57
Zhang B and Harding A K 2000 *Astrophys. J.* **532** 1150