CSIRO Publishing

Publications of the Astronomical Society of Australia

VOLUME 19, 2002

An international journal of astronomy and astrophysics

For editorial enquiries and manuscripts, please contact:
The Editor, PASA,
ATNF, CSIRO,
PO Box 76,
Epping, NSW 1710, Australia
Telephone: +61 2 9372 4590
Fax: +61 2 9372 4310
Email: Michelle.Storey@atnf.csiro.au

For general enquiries and subscriptions, please contact:
CSIRO Publishing
PO Box 1139 (150 Oxford St)
Collingwood, Vic. 3066, Australia
Telephone: +61 3 9662 7666
Fax: +61 3 9662 7555
Email: publishing.pasa@csiro.au

CSIRO
PUBLISHING

Published by CSIRO Publishing
for the Astronomical Society of Australia

www.publish.csiro.au/journals/pasa

Publ. Astron. Soc. Aust., 2002, 19, 34–38

Coherent Emission in AGN: A Critique

D. B. Melrose

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

Received 2001 July 25, accepted 2001 November 27

Abstract: The inverse Compton (IC) limit, $T_B \lesssim 10^{12}$ K, on synchrotron sources places a severe limit on models for intraday variables. The conventional limit is relaxed for proton synchrotron emission, or when acceleration balances IC losses. Coherent emission avoids the limit entirely but introduces other difficulties that have been inadequately discussed.

Keywords: pulsars — radio emission — brightness temperature — inverse Compton limit

1 Introduction

The very high brightness temperatures, T_B, inferred for intraday variable (IDV) sources pose an unsolved problem. The brightness temperature (in energy units) is determined in terms of the observed flux density, F_ν, by

$$T_B = \frac{c^2}{2\nu^2}\frac{F_\nu}{\Delta\Omega} \tag{1}$$

where $\Delta\Omega$ is the solid angle of the source. Very high estimated values of T_B are for very small estimated values of $\Delta\Omega$. The smallest estimate of $\Delta\Omega$ is for an intrinsic model in which the variations, on a timescale $t_{\rm var}$, are assumed to imply a source size, $r \lesssim ct_{\rm var}$, giving $\Delta\Omega = \pi r^2/L^2$, where L is the distance to the source. This leads to an estimate $T_B \sim 10^{21}$ K in the most extreme case (Kedziora-Chudczer et al. 1997). This is $\sim 10^9$ in excess of the inverse Compton (IC) limit of $T_{\rm IC} \approx 10^{12}$ K for a self-absorbed synchrotron source (Kellermann & Pauliny-Toth 1969). Doppler boosting in a jet implies $T_B = \mathcal{D}^3 T_{\rm IC}$, with $\mathcal{D} = 1/[\gamma(1-\beta\cos\theta)]$, where $\gamma = (1-\beta^2)^{-1/2}$ is the Lorentz factor of the jet flow, and θ the angle between the jet axis and the line of sight. Then $T_B \sim 10^{21}$ K requires $\mathcal{D} \sim 10^3$, which is considered implausibly large for a jet emitting at radio frequencies.

An alternative interpretation is that the temporal variations are due to scintillations in the interstellar medium (ISM) (e.g. Walker 1998). A source scintillates, due to turbulence in a screen at a distance D, only for $\theta \lesssim \theta_{\rm D} = r_{\rm D}/D$, where θ is its angular diameter and $r_{\rm D} = (\lambda D/2\pi)^{1/2}$, $\lambda = c/\nu$ is the Fresnel scale, implying $\Delta\Omega \lesssim \lambda/2D$. In the most extreme case, for $D \sim 1$ kpc, one infers $T_B \sim 10^{15}$ K. In this case Doppler boosting gives $T_B = \mathcal{D}T_{\rm IC}$, and again one requires an implausibly large $\mathcal{D} \sim 10^3$. For a source that is unresolved by scintillations, one has $T_B \propto D$, and hence the estimate of T_B is reduced if the screen is relatively close. There is evidence that, at least in some cases, the screen is indeed relatively close, $D \ll 100$ pc (Dennett-Thorpe & de Bruyn 2000), reducing the required Doppler boost to a more plausible value, $\mathcal{D} \lesssim 30$. Although the present author personally favours a scintillation model, such models are not discussed further here.

In this paper several suggestions on how intrinsic values of $T_{\rm IC}$ well in excess of 10^{12} K might be achieved are discussed. In Section 2 ways in which the limit $T_{\rm IC} \lesssim 10^{12}$ K can be relaxed within the synchrotron hypothesis are considered. After outlining a derivation of the conventional IC limit, the changes resulting from (a) postulating that the synchrotron radiating particles are protons rather than electrons, and (b) invoking acceleration that balances synchrotron losses by electrons are discussed. In Section 3 coherent emission mechanisms are reviewed briefly, and in Section 4 their possible application to IDVs is discussed critically.

2 The Inverse Compton Limit

The IC limit is usually taken to be $T_B \lesssim T_{\rm IC} \approx 10^{12}$ K, and it is relevant to consider conditions under which this limit might be relaxed.

2.1 Derivation of the 10^{12} K Limit

There are three assumptions that lead directly to the $T_B \lesssim 10^{12}$ K limit (e.g. Kellermann & Pauliny-Toth 1969; Kardashev 2000): the source is optically thick, the frequency of observation is at the peak of the synchrotron spectrum, and the energy density in synchrotron photons is equal to the energy density in the magnetic field. The derivation is insensitive to factors of order unity, and it suffices to express the first two assumptions in the form

$$T_B \approx E, \quad \nu \approx \nu_B (E/mc^2)^2, \tag{2}$$

respectively, where E is the energy of an electron radiating at the peak of the spectrum, ν is the frequency of emission, and $\nu_B = eB/2\pi m$ is the cyclotron frequency. The third assumption is implied by considering the rate of energy loss by an electron due to synchrotron emission and IC emission. This may be written

$$\frac{dE}{dt} = -\frac{8\pi}{3} r_0^2 c W_{\rm mag}\left(\frac{E}{mc^2}\right)^2 \times \left(1 + \frac{W_{\rm syn}}{W_{\rm mag}} + \frac{W_{\rm syn}^2}{W_{\rm mag}^2} + \cdots\right), \tag{3}$$

 10.1071/AS01111 1323-3580/02/01034$05.00

https://doi.org/10.1071/AS01111 Published online by Cambridge University Press

where $r_0 = e^2/4\pi\varepsilon_0 mc^2$ is the classical radius of the electron, $W_{\rm mag} = B^2/2\mu_0$ is the energy density in the magnetic field, and $W_{\rm syn}$ is the energy density in synchrotron photons. The unit term inside the parentheses describes synchrotron losses, the next term describes IC losses due to scattering off the synchrotron photons, the next term describes IC losses due to scattering off the photons produced in this initial scattering, and so on. The geometric series in (3) may be summed to give $1/(1 - W_{\rm syn}/W_{\rm mag})$, which diverges for $W_{\rm syn} \geq W_{\rm mag}$, leading to the third assumption. One has

$$W_{\rm mag} = mc^2 \frac{\pi \nu_B^2}{2r_0 c^2}, \quad W_{\rm syn} \approx \frac{4\pi \nu^3}{c^3} T_B. \tag{4}$$

Then setting $T_B = E$ and $W_{\rm mag} = W_{\rm syn}$ gives $(E/mc^2)^7 \approx c/8r_0\nu_B$. Thus the factors of order unity appear only to the one seventh power on solving for E or T_B and to the two sevenths power on solving for ν, and can be neglected (cf. Kardashev 2000). Thus one finds

$$T_B = E = mc^2 \left(\frac{c}{8r_0\nu_B}\right)^{1/7}$$

$$\approx (1.0 \times 10^{12}\,{\rm K}) \left(\frac{B}{10^{-4}\,{\rm T}}\right)^{-1/7}, \tag{5}$$

with $10^{-4}\,{\rm T} = 1\,{\rm G}$, and where $T_B = mc^2$ corresponds to $T_B = 0.5 \times 10^{10}$ K. The frequency at the peak of the self-absorbed spectrum is also determined by the model, and is referred to here as the optimum frequency:

$$\nu_{\rm opt} = \nu_B \left(\frac{c}{8r_0\nu_B}\right)^{2/7}$$

$$= (1.2 \times 10^{11}\,{\rm Hz}) \left(\frac{B}{10^{-4}\,{\rm T}}\right)^{5/7}. \tag{6}$$

One has $T_B \propto \nu^{1/2}$ (intensity $I \propto \nu^{5/2}$) for $\nu \ll \nu_{\rm opt}$ in the optically thick regime, and $T_B \propto \nu^{-2-\alpha}$ ($I \propto \nu^{-\alpha}$) for $\nu \gg \nu_{\rm opt}$ in the optically thin regime.

2.2 *Discussion of the 10^{12} K Limit*

The foregoing model is remarkably robust because of its insensitivity to factors of order unity. For example, the assumption $W_{\rm syn} = W_{\rm mag}$ arises from the infinite series in (3), and in practice the scattering cross section decreases from its classical (Thomson) value in the Klein-Nishina limit for sufficiently energetic photons. In a synchrotron self-Compton model for an AGN there may be only one generation of IC photons before this limit is reached. It would then be appropriate to replace the assumption $W_{\rm syn}/W_{\rm mag} = 1$ by $W_{\rm syn}/W_{\rm mag}$ equal to the observationally determined ratio of the power in IC emission to the power in synchrotron emission. However, for every order of magnitude that this ratio exceeds unity, it increases the limit on $T_{\rm IC}$ only by a factor of 1.4.

The assumption $W_{\rm syn} = W_{\rm mag}$ was also questioned by Readhead (1994), who replaced it by the equipartition condition $W_{\rm el} = W_{\rm mag}$, where $W_{\rm el}$ is the energy density in the relativistic electrons. The resulting limit on T_B is smaller than the IC limit by a factor of 3 or so, and this is not exceeded by most self-absorbed sources (Readhead 1994). An argument in favour of the equipartition condition is that it minimises the total energy required to account for a given synchrotron power. A counterargument (K. I. Kellermann 2001, private communication; also Kellermann & Pauliny-Toth 1969) is that in IDVs one is observing a source near where the energy is being injected, and that such a source is unlikely to have relaxed too close to the equipartition model. However, the total energy requirement is a steeply increasing function of T_B, and to obtain T_B even a factor 2 or 3 above the Readhead (1994) limit requires an energy content very much greater than the equilibrium value.

2.3 *The Proton Synchrotron Limit*

Kardashev (2000), cf. also Jukes (1967), suggested that the high values of T_B might be explained in terms of proton synchrotron emission, rather than electron synchrotron emission. With $\nu_B \propto 1/m$, it is apparent from (5) that the limit on T_B scales with the mass of the particle as $T_B \propto m^{9/7}$. The ratio of the mass of the proton to the mass of the electron to this power gives $(m_p/m_e)^{9/7} = 1.6 \times 10^4$, implying that the 10^{12} K limit for electrons is replaced by 1.6×10^{16} K for protons.

The optimum frequency scales as $m^{-3/7}$, so that the factor 1.3×10^{11} Hz in (6) for electrons is replaced by 5×10^9 Hz for protons. The optimum frequency also scales as $B^{5/7}$, so that emission at a fixed frequency requires that for protons the magnetic field be a factor $(m_p/m_e)^{3/5} = 95$ stronger than for electrons.

Any model based on synchrotron emission by protons must avoid a problem with synchrotron absorption by electrons. If relativistic electrons are present in the source region or along the ray direction, then synchrotron absorption by them tends to reduce the brightness temperature to the self-absorption limit for electrons, which is $\lesssim 10^{12}$ K. The proton synchrotron model requires that the source be optically thick to proton synchrotron emission and optically thin to electron synchrotron emission.

2.4 *Initial Injection of Relativistic Electrons*

The conventional IC limit is based on a steady-state model. It was pointed out by Slysh (1992), cf. also Kardashev (2000), that a higher $T_{\rm IC}$ is allowed in a model in which the losses are offset by acceleration. A model, which in a sense is the opposite extreme to the steady-state model, involves assuming (a) a constant injection of electrons at arbitrarily high energies, and (b) IC losses that greatly exceed synchrotron losses due to $W_{\rm syn} \geq W_{\rm mag}$. These authors assumed $W_{\rm syn} \propto E$, and with these assumptions, integrating (3) retaining only the first order IC losses gives $E \propto t^{-1/2}$ after a sufficiently long time such that one has $E \ll E_0$. For this model Kardashev (2000) estimated

$$T_{\rm IC} = (3 \times 10^{13}\,{\rm K}) \left(\frac{\nu}{30\,{\rm GHz}}\right)^{3/2} \left(\frac{t}{1\,{\rm day}}\right)^{-1/2}, \tag{7}$$

https://doi.org/10.1071/AS01111 Published online by Cambridge University Press

with emission at $\nu = 30$ GHz after $t = 1$ day corresponding to $E = 3.5 \times 10^9$ eV.

The assumption $W_{\rm syn} \propto E$ seems artificial: as noted by Kardashev (2000) implicitly it requires that $B \propto t$ increase with time. For a model in which B does not depend on t one has $W_{\rm syn} \propto E^7$, and the long-time behaviour then gives $E \propto t^{-1/8}$. For a more general model with $W_{\rm syn} \propto E^n$, the long-time behaviour gives $T_B \approx E \propto t^{-1/(n+1)}$, $\nu \propto t^{-2/(n+1)}$, and $B \propto t^{(7-n)/3(n+1)}$. Slysh's model corresponds to $n = 3$ and the constant B model corresponds to $n = 7$. All such models allow arbitrarily high $T_{\rm IC} < E_0$ for a characteristic time $\Delta t \sim t$ after the postulated initial injection.

Slysh's model ignores second and higher order generations of IC scattering, and with $W_{\rm syn} \gg W_{\rm mag}$ these should dominate according to (3). Thus it would appear that the model is only valid if these higher order scatterings are suppressed due to the Klein-Nishina effect. The implications of this point do not seem to have been considered.

2.5 *Acceleration Balancing IC Losses*

Slysh (1992) argued that in the presence of acceleration and IC losses, the electrons tend to pile up at the energy where the acceleration balances the losses. Slysh analysed this model for the case where the acceleration has no energy dependence, $dE/dt =$ constant, and the IC losses of the form $dE/dt \propto -E^3$ assumed in the initial injection model. This leads to $T_{\rm IC} \propto \nu^{-2/5}$, and with one choice of parameters Slysh estimated $T_{\rm IC} = 3 \times 10^{14}$ K at $\nu = 1$ GHz. This model generalises to any situation where (a) acceleration balances IC losses, (b) synchrotron losses are negligible compared with IC losses, and (c) the source is optically thick to synchrotron emission. One then has $T_{\rm IC} \approx E$, with E the energy at which acceleration balances IC losses.

3 Coherent Emission Mechanisms

Coherent emission mechanisms are invoked for AGN to overcome the high T_B problem. There are only three known coherent radiation mechanisms in astrophysical plasmas (four if one includes molecular line masers), and none of these is readily adaptable to AGN. In this section I describe the known coherent emission mechanisms then summarise some of the suggested coherent emission mechanisms for AGN.

3.1 *Known Coherent Emission Mechanisms*

There are two reasonably well understood coherent emission mechanisms in astrophysical plasmas (e.g. Melrose 1986): plasma emission and electron cyclotron maser emission (ECME).

Plasma emission is the emission mechanism for most solar radio bursts and for planetary bow shocks. It is characterised by emission at the local plasma frequency, ν_p, or its second harmonic. The emission of escaping radiation is 'indirect' in the sense that it occurs in a two stage process. In the first stage a beam-driven plasma instability causes Langmuir waves (longitudinal plasma waves) to grow, and in the second stage nonlinear processes in the plasma convert these into escaping radiation with little change in frequency (except for frequency doubling). A high effective temperature, T_L say, for the Langmuir waves results from the beam instability, and one has $T_B \lesssim T_L$ for the escaping radiation.

ECME is a 'direct' emission process in the sense that a maser-like instability produces (X-mode) radiation that can escape directly without any second stage conversion process. The emission occurs very near the electron cyclotron frequency, and is driven by an anisotropy in the electron pitch angle distribution. ECME is the accepted emission process for the Earth's auroral kilometric radiation and Jupiter's decametric radiation, and is the favoured mechanism for solar spike bursts and for the very bright radio emission from some flare stars.

Common features of these coherent emission mechanisms are (a) they are associated with natural frequencies in a nonrelativistic plasma, and (b) they are driven by very weak, maser-like instabilities operating near marginal threshold. The maser-like instabilities involve a weak pump (faster electrons outpacing slower electrons to set up a beam-type distribution, and propagation into an increasing B to set up a pitch angle anisotropy, respectively) providing the source of free energy, balanced by a very large number of highly localised, sporadic outbursts of wave growth converting this free energy into wave energy.

Neither of these coherent emission mechanisms is a plausible candidate for AGN, which appears to be at $\nu \gg \nu_p, \nu_B$.

3.2 *Pulsar Radio Emission*

The pulsar radio emission mechanism is not understood. It is generally accepted that the radio emission arises from the polar cap region, where the 'pulsar plasma' is a highly relativistic pair plasma, which is one dimensional (zero gyrational motion). Possible mechanisms include relativistic forms of plasma emission, free electron maser emission, maser curvature emission, and several others (Melrose 1991, 1995). For all those mentioned the only form of free energy in a pulsar plasma is in a distribution of electrons or positrons that is an increasing function of γ, $df(\gamma)/d\gamma > 0$, over some range of γ.

Pulsar radio emission is an implausible basis for a coherent emission mechanism for AGN, which do not have superstrong magnetic fields (at least in the radio source regions) nor the other characteristic properties of a pulsar plasma.

3.3 *Suggested Coherent Emission Mechanisms for AGN*

A specific plasma emission type model for coherent emission mechanism from AGN was proposed by Colgate (1967). Colgate argued that the observed emission could be due to photons, created at $2\nu_p$, interacting with

https://doi.org/10.1071/AS01111 Published online by Cambridge University Press

longitudinal waves, with $\nu \to \nu \pm \nu_p$ in each interaction, causing them to diffuse in frequency, until they escape when their mean free path is comparable with the size of the source. More recent suggestions for coherent emission from AGN, including those of Baker et al. (1988), Sol, Pelletier, & Asséo (1989), Weatherall & Benford (1991), Lesch & Pohl (1992), Krishan & Wiita (1994), and Benford & Lesch (1998), have tended to concentrate on the physics of the emission process, typically in highly relativistic plasmas. However, the seemingly overriding importance, for coherent emission in other astrophysical contexts, of how the microphysics is related to a macroscopic model has been given inadequate attention.

4 Critique of Coherent Emission in AGN

The argument in favour of a coherent emission mechanism is that it can account naturally for the high T_B. The arguments against coherent emission include a variety of criticisms of specific models, especially in connection with the relation between the microphysics and the source model, and in connection with the escape of the radiation. A more qualitative criticism is that the observed emission looks like synchrotron emission, and is not like any other known form of coherent emission, favouring a synchrotron-based interpretation.

4.1 Marginal Wave Growth

In known astrophysical sources of coherent emission, the emission originates from highly localised, transient bursts of emission within a large envelope. A balance between the pumping mechanism, tending to drive a system unstable, and the back reaction (quasilinear relaxation), tending to restore stability, leads to a marginally stable distribution (e.g. Manheimer & Lashmore-Davies 1989). The back reaction tends to cause the growth rate to decrease to a small value which, however, needs to be large enough to overcome various damping and loss processes. This evidently occurs in highly localised, highly sporadic bursts. A simple statistical model, called stochastic growth theory (e.g. Robinson 1995), based on the gain factor being a random variable, accounts well for the statistical distribution of bursts of emission. An important implication for coherent emission in astrophysical plasmas is that the observed properties of the emission depend only in part on the instability, and depend strongly on the statistical distribution of a very large number of highly localised, highly sporadic bursts of wave growth. These important ideas have not been taken into account in existing models for coherent emission in AGN.

4.2 Escape of the Radiation

A serious potential difficulty with any coherent emission model for AGN is that radiation with sufficiently high T_B may not escape. Induced Compton scattering by thermal electrons prevents escape in a spherical source of radius R if the number density n_e satisfies $\sigma_T n_e L(T_B/5 \times 10^9\ \mathrm{K}) \gg 1$, where σ_T is the Thomson cross section. This condition is relatively easily satisfied for the very high values, $T_B \gtrsim 10^{18}$ K, for which coherent emission mechanisms are invoked. Lesch & Pohl (1992) suggested that this difficulty might be overcome due to the anisotropy of the photon distribution. However, a different mechanism can be effective in preventing escape in the highly anisotropic case: the radio-wave beam acts like a particle beam in generating Langmuir waves, which scatter the beam. This effect is the accepted mechanism for the occultation of several radio pulsars for which the line of sight passes through a relatively low density stellar wind (Gedalin & Eichler 1993; Luo & Melrose 1995), and its relevance to AGN was pointed out by Levinson & Blandford (1995). Also, induced scattering occurs in the highly relativistic plasma in which the radiation is assumed to be generated, e.g. the discussion by Wilson & Rees (1978) for the Crab Pulsar. Such effects impose limits on T_B that have not been taken into account adequately in existing coherent emission models for AGN. Benford & Tzach (2000) argued that laboratory evidence suggests that high T_B radiation does escape, contrary to such suggestions. However, laboratory experiments on coherent emission involve systems that are being driven extremely hard, compared to the near-marginal astrophysical systems, with the emission dominated by transient and boundary effects that are not relevant in astrophysical contexts.

5 Discussion and Conclusions

Although the high T_B problem for IDVs is unresolved, the solution is likely to be found within the context of the synchrotron hypothesis. A combination of scintillation due to a nearby screen in the ISM, Doppler boosting in a jet, and some relaxation of the 10^{12} K IC limit can account for the observed source properties. (Relaxation of the IC limit due to acceleration balancing IC losses requires that the Klein-Nishina effect suppress higher order IC scattering, and the implications of this have yet to be taken into account.) The appeal to coherent emission overcomes the high T_B problem but seems to raise several questions that have yet to be adequately addressed, including the source of the free energy, the emission mechanism itself, and the escape of the radiation.

References

Baker, D. N., Borovsky, J. E., Benford, G., & Eilek, J. A. 1988, ApJ, 326, 110
Benford, G., & Lesch, H. 1998, MNRAS, 301, 414
Benford, G., & Tzach, D. 2000, MNRAS, 317, 497
Colgate, S. A. 1967, ApJ, 150, 163
Dennett-Thorpe, J., & de Bruyn, A. G. 2000, ApJ, 529, L65
Gedalin, M., & Eichler, D. 1993, ApJ, 406, 629
Jukes, J. D. 1967, Nature, 216, 461
Kardashev, N. S. 2000, ARep, 44, 719
Kedziora-Chudczer, L., et al. 1997, ApJ, 490, L9
Kellermann, K. I., & Pauliny-Toth, I. I. 1969, ApJ, 155, L71
Krishan, V., & Wiita, P. J. 1994, ApJ, 423, 172
Lesch, H., & Pohl, M. 1992, A&A, 254, 29
Levinson, A., & Blandford, R. 1995, MNRAS, 274, 717
Luo, Q., & Melrose, D. B. 1995, ApJ, 452, 346

https://doi.org/10.1071/AS01111 Published online by Cambridge University Press

Manheimer, W. M., & Lashmore-Davies, C. N. 1989, MHD and Microinstabilities in Confined Plasmas (Bristol: Adam Hilger)
Melrose, D. B. 1986, Instabilities in Space and Laboratory Plasmas (Cambridge: Cambridge University Press)
Melrose, D. B. 1991, ARA&A, 29, 31
Melrose, D. B. 1995, JApA, 16, 137
Readhead, A. 1994, ApJ, 426, 51
Robinson, P. A. 1995, PhPl, 2, 1466
Slysh, V. I. 1992, ApJ, 391, 453
Sol, H., Pelletier, G., & Asséo, E. 1989, MNRAS, 237, 411
Walker, M. A. 1998, MNRAS, 294, 307
Weatherall, J. C., & Benford, G. 1991, ApJ, 378, 543
Wilson, D. B., & Rees, M. J. 1978, MNRAS, 185, 297