Spectral and variability constraints on compact sources

P. W. Guilbert Joint Institute for Laboratory Astrophysics, University of Colorado, Boulder, CO 80309, USA
A. C. Fabian Institute of Astronomy, Madingley Road, Cambridge CB3 0HA
M. J. Rees Institute for Advanced Study, Princeton, NJ 08540, USA and Institute of Astronomy, Madingley Road, Cambridge CB3 0HA

Received 1983 February 18

Summary. The spectra of compact sources which radiate significantly above 511 keV may be modified by the electron—positron pairs created by photon—photon collisions. The Thomson opacity due to pairs may easily exceed unity in sources such as active galactic nuclei; most of the observed luminosity is likely to be due to mildly relativistic pairs. The high-energy spectral cut-off is inversely related to source luminosity. General constraints on variability involving rest-mass to energy conversion efficiency, geometry and the effects of pairs are discussed. Spectral relativistic corrections may be large, even in an isotropic source.

1 Introduction

Rapid variability has been claimed from a number of active galactic nuclei (Tananbaum et al. 1978; Matilsky, Shrad & Tananbaum 1982; Wolstencroft, Gilmore & Williams 1982; Schwartz & Ku 1982) and in particular from NGC 6814 which shows ~ 100 s X-ray variations (Tennant et al. 1981). Extreme variations have been observed in γ-ray bursts such as the 1979 March 5 event for which the rise time was less than 0.3 ms (see Cline 1981). Such sources are compact in the sense that the observed variations take place on a time-scale close to the light crossing time of the emission region. General arguments that constrain the change in luminosity, ΔL, occurring on a time-scale, Δt, yield physical parameters of the sources (Elliott & Shapiro 1974; Lightman, Giaconi & Tananbaum 1978; Cavallo & Rees 1978; Fabian 1979; Fabian & Rees 1978; Rees 1982).

The major consequence of very rapid variability and high luminosity is that the source has to radiate very efficiently. This involves both the rest-mass to energy conversion efficiency as well as the intermediate steps necessary to transfer that energy to the radiating electrons and positrons (Cavaliere 1982; Guilbert, Fabian & Stepney 1982). Pair production from photon—photon collisions may cause important modifications in compact sources where a significant fraction of the primary luminosity is radiated at energies above 511 keV.
We attempt to generalize some of these arguments and show in Section 2 that electron–positron pair production processes may produce and maintain an optical depth $>1$ in $\text{e}^+ - \text{e}^-$ plasma. Any radiation emitted within the source will be ‘reprocessed’ by this plasma, limiting the spectral shape of radiation emerging from compact sources. Much of the radiated energy observed from active galactic nuclei may thus be due to mildly relativistic pairs, a conclusion which is independent of the actual energy production mechanism or even the primary spectrum (Fabian 1982). In Sections 3 and 4 we discuss the manner in which special geometry and relativistic motion can be invoked to overcome the restrictions on the rate of change of luminosity. Even isotropic sources with gas speeds as low as $0.3c$, such as may be expected from gas accreting onto to a black hole, produce significant departures from the rest-frame constraints. We show that thin-layer effects may be relevant to $\gamma$-ray bursts.

2 Spectral constraints on compact sources

In this section we are concerned with hard X-ray spectra such as are observed from many type of active galactic nuclei. The available data on Seyfert galaxies strongly indicate relatively flat spectra (energy index $\alpha \approx 0.6$) which extend to $\sim 1 \text{ MeV}$ (Mushotzky 1982; Rothschild et al. 1982; Perotti et al. 1981; Schoenfelder 1982). Most of the energy is then radiated at about $1 \text{ MeV}$. (BL Lac objects, on the other hand, appear to have steep X-ray spectra (Worrall et al. 1981) and radiate predominantly in the far UV). A spectral turnover at about $1 \text{ MeV}$ may be expected if the underlying radiation mechanism is a thermal plasma, for the cooling function increases dramatically as $kT_e$ exceeds $m_e c^2$ and pair production becomes important (see, e.g. Lightman 1982; Svensson 1982). Thermal equilibrium may not, however, be appropriate for rapidly varying compact sources. In that case a power-law spectrum of electrons may be generated by shock fronts, reconnection regions, localized strong electric fields, etc. An energy of $1 \text{ MeV}$ may not then be of any significance in the formation of such an electron spectrum and need not necessarily be expected from the observations. Photon–photon collisions, however, may be very important in a compact region, creating pairs which radiate (perhaps producing further pairs) and reduce the observed energy cut-off to $\sim 1 \text{ MeV}$.

To explore this last situation further we consider a simple idealized model in which the primary energy input is non-thermal and distributed through a region of radius $R$. Most of the volume within this region is assumed to be sufficiently empty that $\bar{n}_p \sigma_T R \ll 1$, in other words, there is not enough ordinary matter (density $n_p$) to provide significant Thomson opacity. We assume that high-energy electrons are being accelerated at a large number of separate ‘sites’. Relativistic electrons cool (via synchrotron emission if a strong magnetic field pervades the region, or via inverse Compton scattering) in $t_{\text{rad}} \ll R/c$. Thus, a ‘snapshot’ of the central region would reveal a large number of sites each of which is emitting a non-thermal spectrum. The number of electrons involved in producing the radiation at a given instant is, however, very small (contributing $\bar{n}_e \sigma_T R \ll 1$) and the ‘filling factor’ of the sites is also very small.

[The region of radius $R$ may be, for instance, the corona above an accretion disc, a black hole magnetosphere generating energy via the Blandford–Znajek (1977) process (cf. Rees et al. 1982) or the inner region of a black hole accreting in a quasi-spherical fashion.]

We can then discuss general constraints on the spectrum of the radiation emerging from the region, taking account of pair production, etc. Assume each ‘site’ accelerates electrons ‘impulsively’ (i.e. on a time-scale of maybe a few gyroperiods, very much shorter even than the radiative time-scale, which is itself $\ll R/c$). Suppose that these electrons generate (via,
for instance, synchrotron or Compton emission), a power-law spectrum of radiation
extending up to the \( \gamma \)-ray band.\(^*\)

Consider now what happens if a significant fraction of the 'primary' radiation is above
1 MeV (i.e. \( h\nu_{\text{max}} > 1 \text{ MeV} \); \( \alpha \neq 1 \)). We define a parametrized luminosity \( L \) related to total
luminosity \( L \)
by

\[
L = l \frac{m_e c^3}{\sigma_T} \frac{R}{R_S}.
\]

(1)

[Although our general results are model independent, we note that for this definition the
radiation from around a black hole has

\[
l = \left( \frac{m_p}{m_e} \right) \left( \frac{R}{R_S} \right)^{-1} \left( \frac{L}{L_{\text{Edd}}} \right),
\]

where \( R_S \) is the Schwarzschild radius corresponding to the Eddington limit, \( L_{\text{Edd}} \). ] If \( f_\gamma \)
denotes the fraction of the primary luminosity from a typical 'site' which is emitted as
\( \gamma \)-rays above 1 MeV, pair production effects will prevent these photons from escaping the
entire source region if

\[
f_\gamma l > 1.
\]

(2)

The pair density (and temperature) in the source region can now be estimated. Neglecting
dynamics, we suppose that an equilibrium pair density (production rate = annihilation rate)
is maintained in the source. Suppose also that the pairs are generated uniformly through the
entire volume. [This requires that neighbouring 'sites' be separated by distances \( (\ll R) \)
smaller than the mean free path of a \( \gamma \)-ray]. The Thomson depth of the source due to pairs is then

\[
\tau_{e^+ - e^-} = (xf_{\gamma} l)^{1/2} \left[ \frac{\sigma_T c}{\text{annihilation rate constant}} \right]^{1/2}.
\]

(3)

In this expression, \( x (\ll 1) \) is the mean number of pairs produced per \( 2m_e c^2 \) of energy of
photons above the threshold. If \( h\nu_{\text{max}} \) is only a few MeV, it is guaranteed to be not much
below unity. However, if the primary spectrum is very hard, and extends to such high
frequencies that most of the emission is in photons with energies \( > 1 \text{ MeV} \), \( x \) may be small
(but still corresponding to one pair per high energy photon) unless the kinetic energy of an
ultrarelativistic pair produces further pairs via some cascade process. \( \tau_{e^+ - e^-} > 1 \) whenever
the source is optically thick to photon–photon collisions at \( \sim 1 \text{ MeV} \), since the collision
cross-section \( \sim \sigma_T \).

A photon with energy \( e \geq m_e c^2 \) will have a mean free path \( \propto e^{-\alpha} \) if it moves through a
background of soft radiation with \( S(\nu) \propto \nu^{-\alpha} \). The resulting pairs will have high \( \gamma \) and will

\( \star \) We note that there is a maximum frequency of synchrotron radiation that may be produced by shock-
accelerated electrons. Even if we have relativistic shocks, so that \( \Delta \gamma / \gamma \) can be \( \sim 1 \) on a single passage
through the shock, there is a maximum attainable energy which corresponds to the value of \( \gamma_e \) for
which the synchrotron lifetime \( \sim \gamma_e^{-1} B^{-2} \) is comparable with the gyrophase \( \sim \gamma_e B^{-1} \). This gives a
maximum \( \gamma_e \) that scales as \( B^{-1/2} \), and a high-frequency cut-off to the synchrotron spectrum which is
independent of \( B \), and which corresponds to a wavelength of \( (e^2/m_e c^2) \) (classical electron radius); i.e. a
photon energy of \( \sim \gamma_e^{-1} m_e c^2 = 70 \text{ MeV} \). This means that maximally efficient shock acceleration can yield
synchrotron photons above the pair production threshold (but not by an enormous factor). Of course,
Compton scattering by the same electrons can yield much higher energy photons — all the way up to
\( \gamma_e m_e c^2 \). (If the electrons are accelerated not by shocks, but \( \text{along} \) the magnetic field direction by non-
zero \( \mathbf{E} \cdot \mathbf{B} \), then the synchrotron losses would not occur, but Compton losses would still give a comparable
limit to \( \gamma_e \). Cavaliere & Morrison 1980.)
cool quickly via synchrotron and Compton processes. If they cool mainly by emitting energy in photons with $h\nu > m_e c^2$, then each pair yields a cascade of further pairs which will tend to steepen the overall power-law spectrum towards $\alpha = 1$ (Bonometto & Rees 1971), but which could lead to a value of $x$ of order unity. If, on the other hand, the pairs cool solely by emitting synchrotron radiation with $h\nu < m_e c^2$, then no further pairs are produced but the primary spectrum at low energies is substantially augmented (but see later).

If the main loss mechanism is Compton scattering of the primary spectrum many further pairs are guaranteed if $\alpha < 1$, because the main losses arise from Compton scattering of the X-ray part of the spectrum, generating photons which are themselves above the pair production threshold. In this case $x$ may be $\sim 1$, even if most of the ‘primary’ energy goes into photons with $h\nu \gg m_e c^2$. For Compton scattering to dominate it is sufficient for the radiation energy–density $\varepsilon_{\text{rad}}$ to exceed $B^2/8\pi$, but this may not be necessary if the synchrotron emission is inhibited by self-absorption.

The net result is that, though $x$ is uncertain, there is no reason why it should be very small. Therefore all that is required in order to generate $\tau_{e^+ - e^-} > 1$ is $f_\gamma L \geq (m_e/m_\gamma)L_{\text{Edd}}(R/R_S)$.

The expression in square brackets in (3) is $\sim 1$ if the pairs are sub-relativistic (see Ramaty & Meszaros 1981) and increases if the pairs are relativistic, because the annihilation rate $\propto \gamma_{e^+ - e^-}^2$ for $\gamma_{e^+ - e^-} > 1$. We can then conclude that if $(xf_\gamma l) > 1$ we unavoidably generate and maintain $\tau_{e^+ - e^-} > 1$. Moreover, it then follows that the pairs must be sub-relativistic. This is because the non-thermal primary emission process generates predominantly soft photons (for $\alpha > 0$) and Comptonization would be catastrophic if the pairs were relativistic $(\tau_{e^+ - e^-} \propto \gamma_{e^+ - e^-})$ because of the behaviour of the annihilation rate, so the Comptonization parameter $\tau^2 \gamma_{e^+ - e^-}^2$ (already $\geq 1$ for mildly relativistic pairs if $xf_\gamma l > 1$) increases as $\alpha_{e^+ - e^-}$, implying Comptonization losses $\propto \exp[(xf_\gamma l)\gamma_{e^+ - e^-}]$.

We note that photons are trapped in the source when $\tau_{e^+ - e^-}$ exceeds unity. This means that $\tau_{e^+ - e^-}$ enters and strengthens inequality (2). This effect may be important in marginal situations where $\tau_{e^+ - e^-} > 1$ is created initially only by photons above some energy $m_e c^2 < h\nu_\gamma < h\nu_{\text{max}}$. Feedback through photon trapping may then cause $h\nu_\gamma$ to drop to $m_e c^2$ as assumed above.

The observed radiation will then have a spectrum corresponding to what is obtained by 'reprocessing' the part of the primary spectrum below 1 MeV through a Comptonizing medium with $\tau$ given by (3), and with $kT_{e^+ - e^-}$ given by the Compton equilibrium condition. Provided that $\alpha < 2$, $kT_{e^+ - e^-}$ is close to the observed high-energy cut-off, since the Compton equilibrium temperature is

$$kT_{e^+ - e^-} = \frac{h}{4} \int S'(\nu) \nu d\nu$$

where $S'(\nu)$ is the post-reprocessing spectrum).

Equate the luminosities for the unprocessed and processed spectra yields an equation for the Compton parameter $y$ [i.e. $kT_{e^+ - e^-}$ since $y = (4kT/m_e c^2)\tau^2$] which is independent of (4):

$$\left(\frac{\nu_{\text{max}}'}{\nu_e}\right)^{1-\alpha} \cdot \frac{\nu_{\text{max}}^{1+\alpha}}{\nu_e^{1-\alpha}} = \left(\frac{\nu_{\text{max}}'}{\nu_e}\right)^{1-\alpha} \cdot \frac{\alpha}{(1-\alpha)}$$

We assume here that the energies of the lower energy photons are all boosted by a factor of $e^\nu$, and that those photons between $e^\nu \nu_{\text{max}}'$ and $\nu_e$ emerge at the observed high-energy cut-off $\nu_{\text{max}}'$ (see Fig. 1). $\nu_e$ is $m_e c^2/h$. The observed spectral shape is similar to that of the
Figure 1. Schematic diagram showing how a hard primary spectrum extending to \( \nu_{\text{max}} \gg \nu_e \) is partially destroyed by photon–photon collisions and Comptonized by the resultant electron–positron pairs. All of the energy initially in band C is transferred to the photons in bands A and B. Those photons initially radiated in A experience an average upward shift of \( e^\gamma \) (where \( \gamma = 4kT\tau^2/m_e c^2 \), \( T \) is the pair temperature), whilst those in B emerge around \( \nu_{\text{max}} \sim kT/h \). These last photons are shown as a line, although they will in practice form a hump of width \( \sim \nu_{\text{max}} \). The production of further photons by pair bremsstrahlung and synchrotron radiation is likely, but not shown.

The primary spectrum provided that the power-law slope expected from a delta-function input spectrum (see, e.g. Sunyaev & Titarchuk 1980) is steeper than \( \alpha \). This probably means small \( \gamma \) and \( \nu_{\text{max}} < 100 \nu_e \). A Wien hump around \( \nu'_{\text{max}} \) (from photons initially between \( e^{-\gamma} \nu'_{\text{max}} \) and \( \nu_{\text{e}} \)) may only be noticeable if \( \alpha \) and \( \gamma \) are large (i.e. \( \nu'_{\text{max}} \ll \nu_e \)).

Equation (5) implies that \( \nu'_{\text{max}} > \nu_e/\tau^2 \) if \( \alpha \lesssim 0.5 \). However, since the total change in wavelength due to the Compton effect cannot much exceed \( \tau^2/m_e c^2 \), photons at \( \nu_e \) cannot be scattered down much below \( \nu_e/\tau^2 \) even if a smaller value of \( \nu'_{\text{max}} \) is derived from (5). Therefore, when \( \alpha \gtrsim 0.5 \) we expect that the spectrum will first steepen at \( \nu'_{\text{max}} \) and cut off at \( \nu_e/\tau^2 \). Boundary effects probably create a further power-law tail extending to \( \nu_e \) and above, and may strongly influence the overall spectral shape.

The above assumes that the pairs contribute no photons below 1 MeV (i.e. \( x \approx 1 \)). Electron–photon and electron–electron bremsstrahlung (creating a flatter spectrum) and synchrotron or cyclotron radiation (creating steeper spectra) are of course likely to generate many such photons in which case the appropriate observed value of \( \alpha \) may not be the same as the ‘primary’ value. There is then probably not a single power law. Blackbody photons from embedded dense clouds, or an accretion disc, may be an important source of infrared, optical and ultraviolet photons for Comptonization.

The sub-relativistic pair plasma may be optically thick to high cyclotron harmonics. The optical depth to cyclotron absorption is roughly determined by \( \Lambda = (3/2) \pi m_e c^2/\alpha h v_B \theta \), where \( \nu_B \) is the cyclotron frequency and \( \theta = kT e^{-e/m_e c^2} \) (see Masters et al. 1977). For synchrotron to dominate over Compton scattering \( B^2/8\pi = f_B L^2/4\pi R^2 c, f_B > 1 \), which means that \( \Lambda \approx 10^{14} (x f_\gamma R_{14})^{1/2} \theta^{-1} f_B^{1/2} \) for a source size \( R = 10^{14} R_{14} \) cm. If, for instance, \( \theta \) were \( \sim 1/5 \), and the pairs had an approximately Maxwellian distribution, then the \( e^+ – e^- \) plasma would be optically thick up to around the 100th cyclotron harmonic. Any of the primary energy in photons with \( h \nu > m_e c^2 \) which gets converted into relativistic pairs with \( \gamma \lesssim 10 \) would therefore (provided these cool by synchrotron emission) be ‘soaked up’ by the sub-relativistic pairs via cyclotron absorption.

Further work on the detailed emergent spectrum requires that the geometry, magnetic field, electron spectrum and other parameters be specified. We have shown though that
where (2) is satisfied (which is quite easily done) then the generation of an optically thick cloud of sub-relativistic, or mildly relativistic, pairs is inevitable. This leads to a cut-off in the emerging hard X-ray spectrum. The power law observed in the X-rays from such sources may be basically determined by some non-thermal mechanism, although ‘thermal’ Comptonization may have provided much of the energy.

3 Variability and the efficiency limit

The outgoing radiation generally interacts with the matter responsible for the radiation. The mass $M$ of radiating matter is given by

$$\Delta L \Delta t = \eta M c^2$$

(6)

where $\eta$ is the matter to energy conversion efficiency. In a spherical mass the competition between light crossing time and photon diffusion leads to the limit.

$$\Delta L < \frac{\eta c^4 m_p \Delta t}{\sigma}$$

(7)

which for the Thomson cross-section and an efficiency $0.1 \eta_{0.1}$ is

$$\Delta L < 2 \times 10^{41} \eta_{0.1} \Delta t \text{erg s}^{-1}$$

(8)

(Fabian 1979; Fabian & Rees 1979). Substantial luminosity variations are unlikely to occur on time-scales shorter than that given by this limit, which appears to be fairly well obeyed by most sources, despite its simplicity.

The value of $\Delta L$ at the efficiency limit is $\eta$ times the Eddington limiting luminosity for a mass with a Schwarzschild radius of $c\Delta t$. It is thus related to the more specific black hole and accretion arguments of Elliott & Shapiro (1974) and Lightman et al. (1978). Note that in our case $c\Delta t$ does not have to correspond to the Schwarzschild radius but could be some more localized region.

3.1 ENERGY TRANSFER AND RADIATION

The above limit assumes that the radiated energy was previously stored in the form of matter for a time $\gtrsim \Delta t$. It is, of course, possible that magnetic flares are responsible and that the individual electrons (or positrons) that radiate are re-used many times (see Blandford & Rees 1978; Cavaliere & Morrison 1980). The electrons must then spend very little time at sub-relativistic energies in order that $\tau < R/c$. $\eta$ may effectively be $> 1$. We require that the radiation mechanism is rapid enough to release the stored energy within $\Delta t$, i.e.

$$t_{\text{rad}}/\Delta t < (kT_e/\eta_p m_p c^2).$$

(9)

This means that the radiative cross-section must exceed $(\eta m_p c^2/kT_e)(m_e c^2/kT_e)^{1/2} \sigma_T$. The bremsstrahlung cross-section of $\sim \alpha \sigma_T$ is not large enough to yield high $\eta$ (see Lightman et al. 1978). It is also necessary that the process coupling the radiating electrons (or positrons) to the source of energy be rapid. In the case of two-body electron–ion coupling in a thermal plasma we then require that $kT_e \lesssim 35 \text{ keV}$ (Guilbert et al. 1982). Spectra extending to several hundred keV from rapidly varying sources imply some faster process than two-body relaxation. There is then no means yet of estimating the coupling time-scale or the shape of the electron velocity distribution (Gould 1982).

3.2 GEOMETRY

We have assumed above that the radiating region is roughly spherical. The limit is relieved for a thin sheet-like region and for a special viewing direction (perpendicular to the sheet).
A more general sheet-like geometry is a thin spherical shell. The efficiency limit may then be increased by a factor of $\pi \Delta R/R$, where $\Delta R$ is the thickness of the shell at $R (\Delta R \ll R)$. We require that $\pi \Delta R < R$ and that the whole region is ‘switched on’ in a time $\ll \tau \Delta R/c$. The optical depth $\tau = M (4 \pi R^2 m_p c^{-1})$ where $M$ is the mass in the shell. Radiation pressure will cause the shell to expand unless a strong anchored magnetic field is present,

$$B \geq \frac{\Delta L}{R^2} \left( \frac{\sigma \Delta t}{2 \pi m_p c^3 \eta} \right)^{1/2},$$

(the luminosity is well over the Eddington limit if $\pi \Delta R/R \gg 1$). Consequently this situation is unlikely to be relevant to active galactic nuclei, but may apply to the 1979 March 5 $\gamma$-ray burst for which we determine $B \geq 10^{13}$ G (implying the presence of a neutron star).

The spectral arguments leading to (1)–(3) depend on assuming a specific model in which the primary non-thermal photons with the power-law spectrum are injected uniformly throughout the volume. If we took instead the case when the energy production occurred in a thin shell (or in a flat disc) of radius $R$, we would find that the ‘false photosphere’ has a thickness of only $(xf/\gamma)^{-1} R$ and an optical depth $\tau_{ee}$ that is always of order unity. The annihilation line itself may then be a prominent feature. (Contrariwise, in the case when the energy production is distributed through a spherical volume, and $\tau_{ee} \gg 1$, only the annihilation photons from near the surface escape before being degraded by multiple scattering.)

### 3.3 Dynamics

We have so far ignored dynamics. However, if we incorporate our spherical region in a more realistic context, e.g. an accretion disc corona, then relativistic velocities can be generated. Within the sphere, (energy density of photons)/(rest mass density of pairs) $\approx \gamma$. This suggests that high-$\gamma$ bulk motions can be generated hydrodynamically, since the photons would be trapped if $\tau_{ee} > 1$. However, $\tau_{ee}$ always drops to $\sim 1$ in a dynamical time-scale if the internal energy source is switched off. It is therefore not obvious that high bulk $\gamma$s can arise in this way unless internal dissipation (and high-energy photon generation) continues as the material flows out. However, an optically thin ‘false photosphere’ of $e^+ - e^-$ can be accelerated by radiation pressure to a modest $\gamma$ (Phinney 1982).

If the pairs flow away at velocity $v$ from the primary source region of Thomson depth $\tau_p$, we can determine the optical depth of the wind $\tau_w$ by integrating the pair density which is given by

$$\frac{dn}{dr} = \frac{2n}{r} - f_a \frac{n^2 \sigma c}{v}.$$

We assume that the pairs are relatively cold ($kT < m_e c^2$) and that pair annihilation then produces photons in a Gaussian centred on $\sim m_e c^2$. A net loss of pairs into photons then occurs due to the collision of blue-wing ($h\nu > m_e c^2$) photons preferably with other blue-wing photons, resulting in an accumulation of red-wing photons. We include this effect with the factor

$$f_a = \frac{\int_0^\infty \exp \left[ -\frac{(E-1)^2}{2 \theta^2} \right] (eE-1)^{3/2} dE}{\int_1^\infty \exp \left[ -\frac{(E-1)^2}{2 \theta^2} \right] (eE-1)^{3/2} dE} \cdot \frac{\int_1^\infty \exp \left[ -\frac{(E-1)^2}{2 \theta^2} \right] (eE-1)^{3/2} dE}{\int_0^\infty \exp \left[ -\frac{(E-1)^2}{2 \theta^2} \right] (eE-1)^{3/2} dE},$$

$$[\theta = (kT/m_e c^2) \ll 1],$$
which is an integration over the relative blue-wing photon–photon rate (see Gould & Schréder 1967). Numerical integration gives \( f_a \approx 0.45 \), which means that

\[
\tau_w \approx 2.2 \frac{v}{c} \log \left( \frac{\tau_c}{2.2v} \right). \tag{13}
\]

### 3.4 Implications of the Pair Opacity

At first sight, one would think that when (2) holds the minimum variation time-scale would exceed \( R/c \) by a factor \( \tau_{e^+e^-} \) (given by 3). If this were so it would mean that where pair production was important, we would have a more stringent constraint on the variability in all wavebands (X-ray, optical, etc.) and not merely the fireball constraints (Cavallo & Rees 1978) on variations in the \( \gamma \)-rays themselves. However, this is not so, because the time-scale for establishing or changing the equilibrium pair density is less than \( R/c \) (it is in fact \( R/c \tau_{e^+e^-}^{-1} \)). Therefore if the power output of the primary engine were to change (which it could do on a time-scale \( R/c \)), then \( \tau_{e^+e^-} \) would itself respond on this same time-scale. So relations (7) and (8) still apply even if \( \eta \) and the fraction of primary luminosity in \( \gamma \)-rays are high enough to generate \( \tau_{e^+e^-}^{-1} > 1 \). There is no significant upper limit on the efficiency, provided that the non-thermal particles generating the primary spectrum can be re-used many times in each light travel time as discussed in Section 3.2, so \( \Delta L/\Delta t \) can still be arbitrarily large.

### 3.5 Polarization

An important subclass of rapidly varying sources, the OVVs (Angel & Stockman 1981), also display large polarization which may also be variable. Here \( \tau_{e^+e^-} \) must be less than unity, otherwise the only polarization that would be observed would be that due to scattering in non-spherical geometries. If the source dimensions were \( < c\Delta t \), this would preclude the possibility of any significant fraction of the primary power being in \( \gamma \)-rays. If an OVV such as 3C 446, which exhibits high (up to 40 per cent) and variable (\(~1\) day) polarization, an X-ray spectrum extending to an MeV or more, then a strong case could be made for relativistic beaming along a preferred direction. The pairs in a source do not, of course, produce any Faraday rotation.

### 4 Source variability due to relativistic motion

If the source is moving, the energy emitted per unit energy per unit solid angle per unit time is given by

\[
F(k, n, t) = [\gamma(1 - \beta \cdot n)]^{-3} F_r(k_r, n_r, t_r) \tag{14}
\]

where \( k \) is the photon energy; \( n \) is the direction of the observer from the source; \( t \) is the time; \( \beta c \) is the source velocity and subscript \( 'r' \) denotes the source rest frame. Integrating (1) over the photon energy we find

\[
F(n, t) = [\gamma(1 - \beta \cdot n)]^{-4} F_r(n_r, t_r). \tag{15}
\]

Assuming that the source radiates isotropically in its rest frame we obtain:

\[
F_{r\max} = \frac{M_p c^4}{3\sigma_T} \eta \Delta t_r
\]

\[
= [\gamma(1 - \beta \cdot n)]^{-1} \frac{M_p c^2}{3\sigma_T} \eta \Delta t. \tag{16}
\]
Using (2) and (3) we obtain

$$L_{\text{max}} = 4\pi F_{\text{max}} = [\gamma(1 - \beta \cdot n)]^{-5} L_0^{\text{max}}$$

(17)

where $L_{\text{max}}$ is the maximum observed luminosity of the source and $L_0^{\text{max}}$ is the maximum luminosity as determined above (equation 7) for a source at rest.

For a source moving directly towards the observer we have:

$$L_{\text{max}} = \frac{L_0^{\text{max}}}{[\gamma(1 - \beta)]^5} = [\gamma(1 + \beta)]^5 L_0^{\text{max}}.$$  

(18)

These Doppler corrections are plainly very large if relativistic beaming occurs. However, we wish to emphasize that these corrections can be important even when there is no systematic outflow. If the energy is supplied by accretion on to a black hole then most of the emission comes from small radii: for instance if 10 per cent of the rest mass is to be radiated then most of the power comes from distances less than about 10 Schwarzschild radii implying that $\beta \geq 0.3$. For $0.3 < \beta < 0.5$ then $4.7 < [\gamma(1 + \beta)]^5 < 15.6$ and so the maximum observed luminosity for a source moving towards the observer can be increased by an order of magnitude if $\beta < 0.5$ ($\gamma < 1.15$).

Consider a source which consists of a number of identical components which have identical speeds but an isotropic velocity distribution. The average source luminosity is then:

$$\langle L_s \rangle = n \gamma^2 [1 + (\beta^2/3)] L_r$$

(19)

where $n$ is the number of components and $L_r$ is their individual rest frame luminosity. For $\beta = 0.5$, $\langle L_s \rangle = 1.4 n L_r$ and so the average luminosity is much less sensitive to such mildly relativistic speeds. In order to see significant variability, $\Delta L \sim L$, due to a single component:

$$[\gamma (1 + \beta)]^4 < n \gamma^2 [1 + (\beta^2/3)] < \gamma^2 (1 + \beta)^4 [1 + (\beta^2/3)]^{-1}.$$  

(20)

For $0.3 < \beta < 0.5$ then $3 < n < 6$. In terms of $L_0^{\text{max}}$ we find:

$$\langle L_s \rangle^{\text{max}} = \gamma^5 (1 + \beta)^5 L_0^{\text{max}}.$$  

$\langle L_s \rangle^{\text{max}}$ is the maximum average source luminosity for which it is possible to observe variations $\Delta L = \langle L_s \rangle$ in the lifetime of a single component of the source.

It is therefore possible for the average luminosity of an isotropic source to be $[\gamma(1 + \beta)]^5$ times greater than that given by (8), even though the source is stationary. The same result applies to disc geometry if we interpret $\beta c$ as the maximum magnitude of the velocity vector projected into the plane containing the axis of the disc and the line-of-sight. We note that what has been calculated is the variability due to the internal relativistic kinematics of the source. If we allow the structure of the source to vary as well and regard $n$ and the geometry to be statistical averages then it is possible to increase $\langle L_s \rangle$ even more.

The probability that a single component is directed towards the observer sufficiently accurately to give a luminosity boost within a factor $f$ of the maximum value is given by

$$P(f) = \frac{1 - \beta}{2\beta} \frac{1 - f^{1/4}}{f^{1/4}} \gamma(1 + \beta)^{-8} < f < 1.$$  

(21)

For $\beta = f = 0.5$, $P \approx 0.1$ and the probability that $\Delta L \geq L/2$ in the lifetime of a single component when there are altogether six is about 35 per cent. Relativistic boosts are thus likely to be observed.
5 Discussion

We have shown that rapidly varying sources are unlikely to display spectra that extend beyond a few MeV, unless there are relativistic bulk motions. Pair production via photon–photon collisions produces a significant opacity in a compact source, whenever the power generated above 1 MeV exceeds $\sim 10^{-3}$ of the Eddington limit or equivalently $L_0^{\max}/\eta$. Compton scattering and secondary photon production are expected to ensure that the electrons (and positrons) that produce the photons which we observe are at most mildly relativistic. This conclusion would also naturally follow from any thermal radiation process.

We expect that more luminous sources with hard primary spectra should have lower observed cut-offs. Source evolution such as envisaged for quasars may then imply that this cut-off increases as the redshift decreases. This process may be relevant to any standard evolving spectrum required to produce the observed shape of the X-ray background spectrum (cf. De Zotti et al. 1982; Piccinotti et al. 1982; Leiter & Boldt 1982).

Light curves from sources in which pair production occurs are dependent upon the residual luminosity and scattering opacity within the source. Complete annihilation occurring within an initially pair-thick source on a time-scale $< R/c$ could lead to a burst of originally trapped soft photons (naively implying $\eta \gg 1$), together with annihilation radiation. Flux increases may occur on a slower time-scale than flux decreases because of pair production and annihilation. Sources in which $\tau_{e^+e^-}$ continually exceeds unity may display less variability than might otherwise be expected from the primary radiation and energy release processes. This may be relevant to the lack of short time-scale ($< 1$ day) variation in most active nuclei reported by Tennant & Mushotzky (1983). Observations of variability of both the total luminosity and high energy cut-off are important.

Relativistic motions involving $\beta < 1$ in isotropic sources can strongly influence the observed rate of change of luminosity. $\eta \sim 1$ as inferred from application of equation (8) is possible even if the actual value (as defined in 6) is $\leq 0.2$.

Geometric factors may be large, especially for those sources in which there are strong magnetic fields capable of anchoring a thin shell of matter. It seems inevitable that much of the radiation observed from compact sources with hard spectra has been reprocessed by electron–positron pairs. This may limit our ability to understand the underlying energy release mechanisms in such sources. (Large flux variations, particularly when, or if, the source temporarily switches off may prove to be most valuable for observation.) However, there is the encouraging prospect that the emergent spectrum may be amenable to detailed modelling and may, through its dependence on $\tau_{e^+e^-}$, offer clues (independent of variability) to the dimensions of the source, and to the nature of the primary radiation at high photon energies.

Acknowledgments

We are grateful to Sterl Phinney for helpful discussions. MJR acknowledges the hospitality of the Institute for Advanced Study, Princeton, where part of this work was done, and support through NSF grant PHY 79-19884. ACF thanks the Royal Society for support.
References