Abstract: The relationship between observed variability time and emission region geometry is explored for the case of emission by relativistic jets. The approximate formula for the jet-frame size of the emission region, $R = D_1 \Delta_1 t / \Gamma_1$, is shown to lead to large systematic errors when used together with observed luminosity and assumed or estimated Doppler factor $D$ to estimate the jet-frame photon energy density. These results have implications for AGN models in which low-energy photons are targets for interaction of high energy particles and photons, e.g. synchrotron-self Compton models and hadronic blazar models, as well as models of intraday variable sources in which the photon energy density imposes a brightness temperature limit through Compton scattering. The actual relationship between emission region geometry and observed variability is discussed for a variety of geometries including cylinders, spheroids, bent, helical and conical jet structures, and intrinsic variability models including shock excitation. The effects of time delays due to finite particle acceleration and radiation timescales are also discussed.

Keywords: acceleration of particles — BL Lacertae objects: general — galaxies: active — galaxies: jets — galaxies: quasars: general — radio continuum: galaxies — X-rays: galaxies

1 Introduction

In the standard picture of active galactic nuclei (AGN), accretion onto a supermassive black hole is via an accretion disk, and a significant fraction of the accretion power (possibly supplemented by tapping into the rotational energy of the black hole) produces twin opposing relativistic jets moving outward along the disk axis, with typical Lorentz factors $\Gamma \sim 2$–10 as inferred from very long baseline interferometry (VLBI) observations. The objects observed in high energy $\gamma$-rays are ‘blazars’. AGN in which one of the jets is closely aligned toward the observer. It is natural that in $\gamma$-rays we should see preferentially those AGN with aligned jets because the emission from the jet is Doppler boosted in energy and relativistically beamed along the jet direction (for a discussion of relativistic effects see Urry & Padovani 1995). The $\gamma$-ray emission from blazars is variable (as it is also at optical, UV, and X-ray energies). Relativistic effects also cause the observed variability time to be shorter than the timescale over which the emission changes in the jet frame.

The spectral energy distribution (SED) of blazars shows two broad peaks, the low energy peak extending from the infrared to the UV or X-ray region of the spectrum, and the high energy peak starting in the X-ray or $\gamma$-ray range. The usual interpretation is that relativistic electrons produce the low energy part by synchrotron emission, and that the same electrons produce the high energy part by Compton scattering the low energy part and/or external photons to higher energies. The 3rd EGRET catalogue of high-energy $\gamma$-ray sources (Hartman et al. 1999) contains around 70 high confidence identifications of AGN, and all appear to be blazars (von Montigny et al. 1995; Mukherjee et al. 1997). Clearly, the $\gamma$-ray emission is associated with AGN jets.

Four BL Lac objects have been detected in the TeV energy range: Mrk 421 (Punch et al. 1992), Mrk 501 (Quinn et al. 1996), 1E 5234+514 (Catanese et al. 1998), and PKS 2155-304 (Chadwick et al. 1999). Recently, the spectrum of Mrk 501 has been measured up to 24 TeV by the HEGRA telescopes (Konopelko et al. 1999). Several of the EGRET AGN show $\gamma$-ray variability with timescales of $\sim 1$ day (Kniffen et al. 1993) at GeV energies. The TeV $\gamma$-ray emission of two BL Lacs shows very rapid variability. For Mrk 421, variability on a timescale as short as $\sim 15$ min has been reported (Gaidos et al. 1996). In the case of Mrk 501, variability on a timescale of a few hours was observed during the 1997 high level of activity, and there is evidence of a 23 day periodicity (Protheroe et al. 1998; Hayashida et al. 1998) interpreted in terms of a binary black hole model for the central engine by Rieger & Mannheim (2000). These variability timescales place important constraints on the models. For example, the synchrotron self Compton (SSC) model appears to be just consistent with recent multi-wavelength observations of Mrk 421 and Mrk 501 during flaring activity (Bednarek & Protheroe 1997, 1999). However, the allowed range of physical parameters (Doppler factor and magnetic field) is rather small, and this mechanism may well be excluded by future observations. For a recent review of TeV $\gamma$-ray astronomy see Kifune (2002).

Rapid variability in intraday variability (IDV) sources is a long-standing problem as it implies apparent
brightness temperatures in the radio regime which may exceed $10^{12}$ K or relativistic beaming with extremely high Doppler factors, coherent radiation mechanisms, or special geometric effects (Wagner & Witzel 1995). The very rapid flaring observed at TeV and X-ray energies during flaring activity in blazars also presents a challenge for any model and suggests a re-examination of mechanisms which may cause very rapid variability would be worthwhile. In this paper, I concentrate on how the observed variability time is related to the geometry and motion of the emission region, and thus to the photon energy density in the emission region. The blazar emission mechanisms to be discussed include a shock excited emission region, bent jets, a shock propagating along a jet containing a helical structure and illuminating parts of the helix by enhanced interactions/emission of radiation such that the emission structure and illuminating parts of the helix by enhanced jets, a shock propagating along a jet containing a helical in the emission region. The blazar emission mechanisms to the emission region, and thus to the photon energy density variability timescale related to the emission region geometry and motion of the AGN and is simply given by $t_{\text{variability}} \propto D/L_{\text{jet}}$, where $D$ is the distance to the AGN and the observer is very much larger than the jet-frame source radius, $L_{\text{jet}} \approx D \Delta t_{\text{obs}}$. The jet-frame photon energy density is then usually assumed to be $U_{\text{ph}} \propto \pi d_{\text{jet}}^2$, where $d_{\text{jet}}$ is the jet-frame bolometric luminosity given by $L_{\text{bol}} = D^{-4} \pi d_{\text{jet}}^2 F$, $F$ being the observed bolometric flux, and $d_{\text{jet}}$ being the luminosity distance. However, this approach can lead to large systematic errors in the jet-frame photon energy density. This is important because the energy density of photons of the low energy part of the SED may determine the energy losses of electrons, and the rate of up-scattering to $\gamma$-ray energies in SSC models, and the rate of proton-photon collisions in hadronic models (see, e.g. Mannheim & Biermann 1989; Protheroe 1997; Mannheim, Protheroe, & Rachen 2001; Mücke & Protheroe 2001; Mücke et al. 2002). If the emission region is optically thin, it may also have consequences for IDV sources as I will show that it is quite possible that the photon energy density responsible for the so-called Compton catastrophe may actually be lower than usually estimated.

Factors Determining Variability Time in Active Galactic Nucleus Jets

In this paper, I shall consider in the observer frame (unprimed coordinates) the events corresponding to: (i) the emission from centre of cylinder at $(r \equiv t_i = 0, x \equiv x_i = 0, y \equiv n_i = 0, z \equiv z_i = 0)$ with this event defining the origin of coordinates in both frames, i.e. $(x' \equiv t_i = 0, x' \equiv x_i' = 0, y' \equiv y_i' = 0, z' \equiv z_i' = 0);$ (ii) the emission from some arbitrary point in the cylinder at a later time $(t_0, x_0, y_0, z_0)$; (iii) the arrival at the telescope, located in the $x$–$y$ plane, of the photon emitted in event (i); (iv) the arrival at the telescope of the photon emitted in event (ii) (see Figure 1).

In this paper, I shall assume that the distance between the AGN and the observer is very much larger than the dimensions of the emission region ($r$ and $t$ in Figure 1), and that we are not concerned with investigating variability on timescales very much shorter than $r/c$ or $t/c$. Under this assumption, the time interval between arrival of the two photons at the telescope is independent of distance to the AGN and is simply given by

$$t_{\text{variability}} \equiv (t_0 - t_i) = \left( t_0 - \frac{x_0}{c} \cos \theta - \frac{y_0}{c} \sin \theta \right). \quad (1)$$

The error made by using this approximation is $(r/d_L)/(r/c)$ and is negligible compared with the variability timescale associated with the emission region geometry investigated in this paper which is $\sim (r/c)$.

Lorentz transformation to the observer frame gives

$$t_i = \Gamma(t_0' + \beta x_0'/c), x_0 = \Gamma(x_0' + \beta c t_0'),$$

and $y_0 = y_0'$. Noting that the Doppler factor is $D = [\Gamma(1 - \beta \cos \theta)]^{-1} = \Gamma(1 + \beta \cos \theta)$, and using the aberration formulae

$$\cos \theta = \frac{\sum D(\cos \theta' + \beta), \quad \sin \theta = \frac{1}{D} \sin \theta'}{2} \quad (2)$$

$$\cos \theta' = \Gamma D(\cos \theta - \beta), \quad \sin \theta' = D \sin \theta \quad (3)$$

Figure 1 Cylindrical emission region geometry.
equation (1) becomes
\[ t_{\text{obs}} = \frac{1}{c} (\ell^\prime c) \sin \phi^\prime - \frac{1}{Dc} \sin \phi \]  
\[ D_{\text{obs}} = \frac{\ell}{c} + \frac{\ell^\prime c}{c} \sin \phi^\prime - \frac{\ell}{c} \sin \phi. \]  

Let us suppose that the cylindrical emission region emits radiation simultaneously and uniformly throughout its volume between times \( t^\prime = -\Delta t^\prime \) and \( t^\prime = \Delta t^\prime \), as measured in the jet frame. While a simultaneous burst violates causality, it is nevertheless a useful case to consider because it enables us to determine clearly the contribution of emission region geometry to the observed variability time. For \( \phi^\prime < \frac{\pi}{2} \) the first photon to arrive would have been emitted at \( (t^\prime = -\Delta t^\prime, x^\prime = \ell^\prime/2, y^\prime = r, z^\prime = 0) \)

Similarly, the last photon to arrive would have been emitted at \( (t^\prime = \Delta t^\prime, x^\prime = -\ell^\prime/2, y^\prime = r, z^\prime = 0) \)

giving
\[ t_{\text{obs}} = -\Delta t^\prime - (\ell^\prime/2c) \sin \phi^\prime - (r/c) \sin \phi. \]  

For \( \phi^\prime > \pi/2 \) the first photon to arrive would have been emitted at \( (t^\prime = -\Delta t^\prime, x^\prime = -\ell^\prime/2, y^\prime = r, z^\prime = 0) \), and the last photon to arrive would have been emitted at \( (t^\prime = \Delta t^\prime, x^\prime = \ell^\prime/2, y^\prime = -r, z^\prime = 0) \). Hence, if we define the observer frame duration of the burst as \( 2\Delta t_{\text{obs}} \) then
\[ D\Delta t_{\text{obs}} = \Delta t^\prime + (\ell^\prime/2c) \sin \phi^\prime + (r/c) \sin \phi. \]  

and this is valid for all \( \theta \). Note that equation (8) gives the usual formula, \( \Delta t_{\text{obs}} = \Delta t^\prime / D \), if the emission region is point-like (i.e. \( \ell^\prime = r = 0 \)). If one term in equation (8) dominates, \( D\Delta t_{\text{obs}} \) gives one of \( \Delta t^\prime \), \((\ell^\prime/2c) \sin \phi^\prime \), or \((r/c) \sin \phi^\prime \).

I shall next consider the case of a cylindrical emission region in the jet being rapidly energised by a plane shock with jet-frame speed \( \beta_{\text{shock}} \) travelling along the jet, such that photons are emitted immediately after shock passage from a thin disk-like region immediately downstream of the shock. In this case, the location of the emitting disk is defined by \( x^\prime = \beta_{\text{shock}} c t^\prime \), and so the arrival times of photons at the telescope may be obtained from equation (5)

For \( (1/\beta_{\text{shock}} - \cos \phi^\prime) < 1 \) the first and last photons to be received would have been emitted at \( (x^\prime = \ell^\prime/2, y^\prime = r, z^\prime = 0) \) and \( (x^\prime = -\ell^\prime/2, y^\prime = r, z^\prime = 0) \), respectively. However, for \( (1/\beta_{\text{shock}} - \cos \phi^\prime) > 1 \) the first and last photons to be received would have been emitted at \( (x^\prime = -\ell^\prime/2, y^\prime = r, z^\prime = 0) \) and \( (x^\prime = \ell^\prime/2, y^\prime = -r, z^\prime = 0) \), respectively. Hence, the observer frame duration of the burst is
\[ D\Delta t_{\text{obs}} = \frac{1}{\beta_{\text{shock}}} \left| \frac{\ell^\prime}{2c} \sin \phi^\prime + (r/c) \sin \phi \right|. \]  

If \( \phi^\prime \) is small (\( \phi \) is very small), one finds
\[ \Delta t_{\text{obs}} \approx \frac{\ell^\prime}{2D} \left[ 1 - \frac{1}{\beta_{\text{shock}}} \right] + r\beta^\prime \]  

and for a reasonable shock speed, e.g. \( \beta_{\text{shock}} \sim \pm 0.5 \),
\[ \Delta t_{\text{obs}} \sim (2 \approx \frac{\ell^\prime}{2D} r^\prime \frac{\beta^\prime}{\beta} \]  

where \( \beta_{\text{shock}} \sim -0.5 \) corresponds to a reverse shock.

What we have learned from this discussion is that multiplying \( c\Delta t_{\text{obs}} \) by the Doppler factor might give the jet-frame intrinsic variability time or one of the dimensions of the emission region (possibly multiplied by some unknown factor). One is tempted to ask if \( R = D\Delta t_{\text{obs}} \) is the right dimension to put in \( U_{\text{dim}} = L/4\pi R^2 c^3 \) in order to estimate the jet-frame photon energy density. We shall discuss this point further in Section 4.

3 Monte Carlo Investigation of \( c\Delta t_{\text{obs}}, \) \( \text{Size, and D} \)

The Monte Carlo method allows the accurate calculation of expected light curves for any emission region geometry and intrinsic source variability. The emission region is modelled in the jet frame, and is represented by \( N \) particles, each of which emits precisely one photon. The number density of the particles models the geometry of the emission region, and each particle emits its photon at a time determined by the emission region geometry and variability model. The jet-frame 4-position of each photon emission event \( (r^i, x^i, y^i, z^i) \), \( i = 1, 2, \ldots, N \), is determined by the model. The emission region moves in the \( \phi \)-direction with Lorentz factor \( \Gamma \), and the observer-frame 4-position of each photon emission event \( (r^i, x^i, y^i, z^i, \phi^i, \theta^i) \), \( i = 1, 2, \ldots, N \), is obtained by Lorentz transformation. The arrival time of each photon is calculated for a given viewing angle \( \theta \) using equation (1), and is binned to the light curve.

Two emission region geometries are considered: a solid cylinder, and a spheroidal 3D Gaussian. We also consider three intrinsic jet-frame time distributions: a simultaneous burst (violates causality), a simultaneous Gaussian pulse (violates causality) and simultaneous emission at plane shock. I shall investigate effects of varying the Doppler factor \( D \), viewing angle \( \theta \), jet Lorentz factor \( \Gamma \), and shock speed \( \beta_{\text{shock}} \). Finally, the effect of smoothing due to acceleration/radiation time delays is discussed.

3.1 Solid Cylinder Emission Region Geometry

If the emission region is ‘solid’, i.e. the emissivity is constant inside the emission region and zero outside, the light...
3.2 Gaussian Emission Region Geometries

The density of ‘particles’ representing the emission region is described by

\[ \rho(x', y', z') \propto \exp \left( -\frac{x'^2}{2 \sigma_x'^2} - \frac{y'^2}{2 \sigma_y'^2} - \frac{z'^2}{2 \sigma_z'^2} \right) \]

with \( \sigma_x' = \sigma_x \) corresponding to a spherical Gaussian distribution, \( \sigma_y' < \sigma_x \) to an oblate spheroidal Gaussian distribution, and \( \sigma_y' > \sigma_x \) to a prolate spheroidal Gaussian distribution. For the case of a simultaneous Gaussian pulse, the probability of emission at jet-frame time \( t' \) to \((t' + dt')\) is \( \rho(t') dt' \), where \( \rho(t') = \exp(-t'^2/2\sigma_t'^2) \sqrt{2\pi} \sigma_t' \) is independent of position.

If the emission region has a Gaussian shape the light curve will be smooth, and in many cases will also have a Gaussian shape. This is true, for example, for the cases of a simultaneous Gaussian pulse and excitation by a plane shock. The width of the light curve will depend on viewing angle in approximately the same way as for cylindrical emission volume. For example, \( r \) and \( r' \) of the solid cylinder should be related to \( \sigma_x \) and \( \sigma_y \), respectively, for the case of the spheroidal Gaussian density.

Figure 2: Light curve from cylindrical emission region with \( \ell/2\pi = 0.1 \) moving with Lorentz factor \( \Gamma = 10 \) as observed at various viewing angles: (a) simultaneous burst for \( \theta = 0^\circ \) (solid), \( 30^\circ \) (dotted), \( 72^\circ \) (short-dashed), \( 108^\circ \) (long-dash, identical to \( 72^\circ \)), (b) excitation by a plane shock moving at speed \( \beta_{\text{shock}} = 0.5 \) for \( \theta = 0^\circ \) (solid), \( 30^\circ \) (dotted), \( 72^\circ \) (short-dash), \( 108^\circ \) (dot-dash), \( 144^\circ \) (dot-dot-dash), \( 360^\circ \) (long-dash, identical to \( 0^\circ \)).

For the case of a simultaneous Gaussian burst I find the standard deviation to be given by

\[ \sigma = \left[ (\sigma_x \sin \theta')^2 + (\sigma_y \cos \theta')^2 + \sigma_t'^2 \right]^{1/2}/D. \]

Note the similarity to equation (8) except that the terms are added in quadrature as the standard deviation is required instead of the maximum duration of the pulse which is, theoretically, infinite.

In the case of excitation by a plane shock wave with speed \( \beta_{\text{shock}}c \), the emission is from the plane where the shock cuts the spherical Gaussian density. The emission region then has a surface density of emitting particles which is a two-dimensional Gaussian, and the light curve reflects this distribution, and so is also Gaussian. The standard deviation of the Gaussian light curve depends on the viewing angle as a result of projection effects, and I find the standard deviation to be

\[ \sigma = \frac{1}{cD} \left[ (\sigma_x \sin \theta')^2 + (\sigma_y \cos \theta')^2 + \left( \frac{1}{\beta_{\text{shock}} c} \right)^2 \sigma_t'^2 \right]^{1/2}. \]

Light curves for \( \Gamma = 10, \beta_{\text{shock}} = 0.5 \) and various viewing angles are plotted in Figure 3 and are seen to lie on top of each other when plotted in units of \( \sigma \), and to be a normal distribution. Equation (15) is valid for both forward and reverse shocks. For reverse shocks, \( \beta_{\text{shock}} \) would be negative and give rise to a broader light curves (larger \( \sigma \)).

3.3 Acceleration/Radiation Delays

Instantaneous excitation by a shock wave is not a realistic approximation unless the timescales for particle acceleration and radiation (cooling) are very short compared to the transit time of the shock through the emission region. If this is not the case, the light curve would be broadened.
and smoothed to reflect the time delays associated with particle acceleration and radiation. An equation of the form \( dn'/dt' = Q'(t') - N'/\Delta \tau_{\text{gain}} \) might describe the time-evolution of the number of particles, \( N' \), radiating photons at the frequency corresponding to that observed, with \( Q'(t') \) being the source term, and \( \tau_{\text{gain}} \) representing the timescale for particle losses (or timescale over which the radiation is emitted). Then a simple smoothing function of the form

\[
\phi_{\text{del}}(\tau_{\text{del}}) = \begin{cases} 
0 & (\tau_{\text{del}} < 0) \\
\frac{1}{\Delta \tau_{\text{gain}}} \left( 1 - \exp \left( -\frac{\tau_{\text{del}}}{\tau_{\text{gain}}} \right) \right) & (0 \leq \tau_{\text{del}} < \Delta \tau_{\text{gain}}) \\
\frac{1}{\Delta \tau_{\text{gain}}} \left( 1 - \exp \left( -\frac{-\tau_{\text{del}}}{\tau_{\text{gain}}} \right) \right) \times \exp \left( -\left(\tau_{\text{del}} - \Delta \tau_{\text{gain}}\right)^2 / \Delta \tau_{\text{gain}}^2 \right) & (\Delta \tau_{\text{gain}} \leq \tau_{\text{del}}) 
\end{cases}
\]

(16)
can be used for this purpose, where \( \tau_{\text{del}} \) represents the delay between shock passage and emission by the radiating particles, and \( \Delta \tau_{\text{gain}} \) represents the duration of the acceleration following shock passage. In the case of a pre-existing population of thermal electrons shock heating is essentially instantaneous (\( \Delta \tau_{\text{gain}} = 0 \)), and \( \tau_{\text{gain}} \) represents the timescale for cooling by thermal bremsstrahlung radiation. Similarly, for a pre-existing population of relativistic electrons, the passage of the shock will essentially instantly increase the magnetic field (\( \Delta B_{\text{gain}} = 0 \)) and \( \tau_{\text{gain}} \) would represent the timescale for energy losses by synchrotron emission. If particle acceleration is required then \( \Delta \tau_{\text{gain}} > 0 \).

Figure 3 Light curve due to excitation by a plane shock of a spherical density moving with Lorentz factor \( \Gamma = 10 \) for \( \beta_{\text{shock}} = 0.5 \) and various viewing angles shown is seen to be a normal distribution.

Figure 4 Excitation of an oblate spheroidal Gaussian density \((\alpha/\gamma = 0.2)\) moving with Lorentz factor \( \Gamma = 10 \) by a plane shock with jet-frame speed \( \beta_{\text{shock}} = 0.3 \) exciting by particle acceleration and radiation described by equation (16) with \( \Delta \tau_{\text{gain}} = \tau_{\text{gain}} \) for \( t_{\text{obs}} = 0.1 \beta_{\text{jet}} / c, \alpha_{\gamma} / c, \) and \( 10 \beta_{\text{jet}} / c \) (as indicated) for viewing angle \( \theta = 5^\circ \). Solid curves give the result of convolution (equation 18).

Figure 4 shows the effect of acceleration/radiation delays on the light curve due to a plane shock \((\beta_{\text{shock}} = 0.3)\) exciting an oblate spheroidal Gaussian emission region \((\alpha/\gamma = 0.2)\) with acceleration/radiation delays described by \( \Delta \tau_{\text{gain}} = \tau_{\text{gain}} \) for \( t_{\text{obs}} = 0.1 \beta_{\text{jet}} / c, \alpha_{\gamma} / c, \) and \( 10 \beta_{\text{jet}} / c \), and viewed at angle \( \theta = 5^\circ \). In the case of no acceleration/radiation delays, the light curves would simply be normal distributions centred on \( t_{\text{obs}} = 0 \). Results are shown for the three \( \tau_{\text{gain}} \) values, and \( t_{\text{obs}} \) is plotted in terms of \( \sigma_{\text{del}} \) defined by

\[
\sigma_{\text{del}}^2 = \sigma^2 + \tau_{\text{gain}}^2 + \Delta \tau_{\text{gain}}^2 / 4 \Delta \tau_{\text{gain}0}.
\]

(17)

where \( \sigma \) is given by equation (15), such that \( \sigma_{\text{del}} \) gives a crude measure of the expected duration of the light curve. These distributions can be obtained simply by convolution of a normal distribution \( \phi_{\text{del}} \) (equation 16), taking account of the fact that jet-frame times enter in equation (16), that for any point co-moving with the jet \( \Delta t_{\text{obs}} = \Delta t / D \), and that \( t_{\text{obs}} \) is plotted in units of \( \sigma_{\text{del}} \).

\[
P \left( \frac{t_{\text{obs}}}{\sigma_{\text{del}}} \right) = \frac{\sigma}{\sqrt{2 \pi}} \int_{-\infty}^{\infty} \phi_{\text{del}}(t_{\text{obs}} - t) \exp \left( -\frac{t^2}{2 \sigma_{\text{del}}^2} \right) \, dt.
\]

(18)

The solid curves, which agree with the Monte Carlo results (histograms), are obtained from equation (18). As we see, if the timescales for the acceleration/radiation process \( (\Delta \tau_{\text{gain}} \text{ and } \tau_{\text{gain}}) \) are much less than the timescale \( \sigma \) associated with the shock passage and dimensions of the emission region then the light curve will be symmetrical, and in the case of a Gaussian spheroidal emission region will be a Gaussian distribution with standard deviation \( \sigma \) (leftmost histogram). If the timescales for the acceleration/radiation process are much larger than \( \sigma \), then the...
The light curve will reflect that of the acceleration/radiation process, i.e. equation (16) (rightmost histogram). The middle histogram shows an intermediate case.

Light curves of AGN show many different features and forms of variability. One example which appears to show light curves (solid curve) from a number of flares of Gaussian form with standard deviations $\sigma = 2.5d$ (dotted curves) plus a background flux density of 2 mCrab. One could interpret this as indicating that the energy-loss timescale of the radiating electrons is much less than 2.5 $D_d$, and that the emission region dimensions are $< 2.5D_d = 2 \times 10^{-3} D_{pc}$.

4 Dependence of Energy Density on Dimensions of the Blob

It is important to know the dimensions of the emission region for several reasons: (i) in some hadronic models the synchrotron photons are targets for photoproduction; (ii) in all models the synchrotron photons are targets for photon-photon pair production by $\gamma$-rays; (iii) in SSC models the synchrotron photon energy density determines Compton scattering; and (iv) knowing the systematic errors on photon energy density may help understand the so-called ‘Compton catastrophe’ in IDV sources which have apparent brightness temperatures well in excess of the limit $T_b < 10^{12}$ K imposed by Compton scattering (Kellermann & Pauliny-Toth 1969; Kardashev 2000) when the photon energy density in the emission region reaches the energy density in the magnetic field. One extreme example is PKS 0405-376 (Kedziora-Chudczer et al. 1997) which has $T_b > 5 \times 10^{14}$ K after correcting for interstellar scintillation (see also Walker 1998) requiring a Doppler factor of $D = 10^3$ to satisfy the brightness temperature limit.

To illustrate how critically the energy density depends on the geometry, I shall consider the case of the jet-frame emissivity following a spheroidal Gaussian density. Provided the emission is optically thin, as is almost certainly true for the optical–$X$-ray synchrotron hump in the SED of blazars, then it is straightforward to calculate the average energy density from the emission region geometry and the luminosity.

Assuming that the jet-frame luminosity, $L'$, is constant, we can estimate the average jet-frame photon energy density given the Doppler factor and emission region geometry. The simplest way of doing this, for any emission region geometry, is to use the Monte Carlo method to place $N$ points at positions $\vec{r}_i$, $i = 1, \ldots, N$, distributed according to the emission region geometry, and to give each point a luminosity $L'/N$. Then at point $\vec{r}_i$ the energy density is

$$U_{\text{phot}}(\vec{r}_i) \approx \frac{L'}{4\pi c N} \sum_{j=1}^{N} \frac{r_{ij}^2}{r_{ij}}$$

(19)

where $r_{ij} = |\vec{r}_i - \vec{r}_j|$. Averaging over the emission region distribution we obtain

$$\left\langle U_{\text{phot}} \right\rangle \approx \frac{L'}{4\pi c N} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{r_{ij}^2}{r_{ij}}$$

(20)
The result giving the jet-frame average photon density versus $\sigma_r$ is shown as the thick solid curve in Figure 7 for the case of a Gaussian spheroidal density.

In Figure 7 we also plot the jet-frame average photon density that would be inferred if we assumed that the observed variability timescale $\sigma$ and an assumed or estimated Doppler factor $D$ gave the jet-frame radius of a spherical emission region, i.e. $R = \sigma D$. This ‘observed’ jet-frame average photon energy density is simply given by

$$\langle U'_{\text{phot,obs}} \rangle = \frac{L'}{4\pi (\sigma D)^2 c}.$$  \hfill (21)

Note that $\sigma$ (given by equation 14) depends only on $\theta'$, $\sigma_c$, and $\sigma_\gamma$, so the ‘observed’ jet-frame average photon energy density depends also on $\theta'$, $\sigma_c$, and $\sigma_\gamma$ as well as $D$. The figure shows $\langle U'_{\text{phot,obs}} \rangle$ plotted against $\sigma_c$ and gives the range due to variation in $\theta'$ (shaded) for various $\sigma_c$. From equation (14) $\langle \sigma' \rangle = \sigma (\pi - \theta')$. Also, for $\theta' = \pi/2$ there is no dependence of $\langle \sigma' \rangle$ on $\sigma_c$ and this gives the horizontal lines bounding the shaded areas, whereas for $\theta' = 0$ (or $\pi$) $\sigma' \propto \sigma_c^{-2}$ for $\sigma_c \gg \sigma_\gamma$ giving the other curves bounding the shaded areas. We see that the ‘observed’ value, $\langle U'_{\text{phot,obs}} \rangle$, can be several orders of magnitude higher or lower than $\langle U'_{\text{phot}} \rangle$ if the emission region is different from a sphere, or if the intrinsic variability time $\sigma_\gamma$ is not small.

For example, take the case of $\sigma_\gamma \ll \sigma_c/\gamma$, if $\sigma_c = 10^{-2} \sigma_\gamma$ and $\theta' = 0$ then $\langle U'_{\text{phot,obs}} \rangle = 10^2 \langle U'_{\text{phot}} \rangle$, whereas if $\sigma_c = 10^2 \sigma_\gamma$ and $\theta' = 90^\circ$ ($\pi = 5.74'$ for $\Gamma = 10$) then $\langle U'_{\text{phot,obs}} \rangle = 10^3 \langle U'_{\text{phot}} \rangle$. The above result has clear implications for both leptonic and hadronic models of AGN in which photons of the low-energy peak of the SED provide target photons for inverse-Compton scattering by electrons (leptonic models) or pion photoproduction by protons (hadronic models). Using the observed variability time together with assumed or estimated Doppler factor to estimate the emission region radius $R'$ can clearly lead to large errors in $\langle U'_{\text{phot}} \rangle$.

Although the radio emission in IDV sources is usually assumed to be optically thick, if this is not the case then the above result may also have implications for IDV sources as the photon energy density responsible for causing the brightness temperature limit may actually be a few orders of magnitude lower than estimated on the basis of time-variability, and in that case much lower Doppler factors would be required to avoid the Compton catastrophe. I explore this further in a separate paper (Protheroe 2002).

### 5 Variability due to Photon Pile-up in Observation Time

The simplest example of photon pile-up in observation time is a bent jet. Jets may bend if they pass through a stratified cold and high density region (Mendoza & Longair 2001). We approximate the trajectory of an emission region moving with speed $\beta c = (1 - 1/\Gamma^2)^{-1/2}$ along a bent section of jet by motion around a section of a circle of radius $r$ in the $x-y$ plane: $(x = r \cos \omega t, y = r \sin \omega t, z = 0)$, where $\omega t = \beta c \sin \eta$. For an observer in the $x-z$ plane at angle $\eta$ to the $x$-axis, the observation time is then given by

$$t_{\text{obs}} = t - (r/c \cos \omega t) \cos \eta$$ \hfill (22)

and the Doppler factor is

$$D = [\Gamma(1 + \beta \cos \eta \sin \omega t)]^{-1}.$$ \hfill (23)

The Doppler factor raised to the 4th power is plotted against observation time for $\Gamma = 10$ and various observation angles in Figure 8. I have plotted $D^4$ as it is appropriate...
for bolometric flux from a moving isotropic source; it also applies to the specific flux $F$, for $F \propto v^{-1}$. We see that even for modest observation angles the light curve is strongly peaked, essentially a delta function when the emission region direction is closest to the line of sight. Of course the finite size of any emission region will broaden the distribution. For example, for a laboratory frame emission region length along the jet $\ell$, the burst would have duration $\Delta t_{\text{obs}} \sim \ell/c$.

### 5.1 Helical Jet Structures

VLBI observations show that helical jets or helical structures in jets may be fairly common in AGN (Raniakyo et al. 1998), and theoretical studies have shown that wave-like helical structures can occur as a result of jet precession (Hardee 2000). Several papers discuss helical jet models or the application of helical models to specific sources (Camenzind 1986; Rosen 1990; Tateyama et al. 1998; Qian et al. 1992; Schramm et al. 1993; Steffen et al. 1995; Villata & Raiteri 1999). Certainly helical jets or structures would be important in determining the light curve of $\gamma$-ray and neutrino emission from blazars, and various suggestions have been made about the mechanisms involved (Despringre & Fraix-Burnet 1997; Marcowith, Henri, & Pelletier 1995).

I consider a filamentary helical structure embedded in the jet with Lorentz factor $\Gamma$ whose axis coincides with the jet axis and is excited by a plane shock travelling along the jet with jet-frame speed $\beta_{\text{shock}}c$. The helical structure could be, for example, a flux tube containing a relatively high magnetic field, or a tube of high plasma density arising from a density perturbation in the plasma entering the jet. Helical magnetic fields with an Archimedean spiral topology similar to the ‘Parker spiral’ field of the heliosphere may well be expected in AGN jets.

The excitation of a filamentary helical structure is described by the jet-frame 4-vector $(t', x', y', z') = (\lambda, \phi / \beta_{\text{shock}}c, \lambda \phi, r \cos \phi, r \sin \phi)$, where $r$ is the radius of the cylinder containing the helix, $\lambda$ is the helix wavelength, $\phi = \beta_{\text{shock}}c \lambda / c$, and the jet is pointing in the $z$-direction. Lorentz transformation to the galaxy frame gives $(t, x, y, z) = (\Gamma [-1 + \beta_{\text{shock}}], \Gamma [1 + \beta_{\text{shock}}], r \cos (\beta_{\text{shock}}ct / \lambda), r \sin (\beta_{\text{shock}}ct / \lambda))$. The galaxy-frame speed of the location of the excited part of the helix can exceed $c$, but this does not violate causality as no particles or information propagates at this pattern speed which is

$$v_{\text{pattern}} = \frac{\beta_{\text{shock}} + \beta}{1 + \beta_{\text{shock}} \beta} \left[ 1 + \left( \frac{r_{\text{shock}}}{\lambda \Gamma (\beta_{\text{shock}} + \beta)} \right) c \right].$$

(24)

Observation in the $x-z$ plane at angle $\theta$ to the jet axis ($y$-axis) gives

$$t_{\text{obs}} = \Gamma [1 + \beta_{\text{shock}} - (\beta_{\text{shock}} + \beta) \cos \theta] + r \sin \theta \sin (\beta_{\text{shock}}ct / \lambda)/c.$$   

(25)

$$\frac{dt_{\text{obs}}}{dt'} = \Gamma [1 + \beta_{\text{shock}} - \cos \theta (\beta_{\text{shock}} + \beta)] + r \sin \theta (\beta_{\text{shock}} / \lambda) \cos (\beta_{\text{shock}}ct / \lambda).$$

(26)

Whenever $dt_{\text{obs}}/dt' = 0$ the light curve will have a cusp. If cusps are possible, they will occur at times corresponding to

$$\cos \phi = \frac{\Gamma [1 + \beta_{\text{shock}} - \cos \theta (\beta_{\text{shock}} + \beta)]}{r \sin \theta (\beta_{\text{shock}} / \lambda)}.$$  

(27)

Provided the model parameters give $\cos \phi$ in the range $-1 \leq \cos \phi \leq 1$ (cos $\phi$ depends on the helix geometry, shock speed, jet Lorentz factor, and viewing angle). If $\cos \phi = \pm 1$ one cusp per helix wavelength will occur, and if $-1 < \cos \phi < 1$ multiple cusps occur, otherwise no cusps are present in the light curve. However, if $\cos \phi$ is close to $\pm 1$ the light curve will be peaked. This is illustrated in Figure 9 for $\Gamma = 10, \beta_{\text{shock}} = 0.5$, and $\lambda = r$. Figure 9(a) shows $t'$ plotted against $t_{\text{obs}}$, for five viewing angles, and Figure 9(b)–(d) shows the resulting light curve for each of the five viewing angles (the corresponding Doppler factors are also given).

Apart from the periodicity, the light curves shown in Figure 9 are reminiscent of those of blazars. All of these light curves correspond to instantaneous emission from the point on the helical filament at the time of excitation. The cusps and peaks would in reality be smoothed to some extent by the finite width of any helical structure as well as by any delays associated with acceleration and radiation timescales. Furthermore, as the shock weakens while propagating down the jet, the successive peaks/cusps in the light curve would decrease in height. One possible scenario could be that there is a succession of shocks at random intervals which propagate down the jet, and because individual shocks weaken as they propagate, each shock would cause only one or two peaks (due to only one or two cycles of the helix). In this case, it may not be possible to distinguish between a simple bent structure in the jet and a helical structure — if appropriately aligned, a simple bent structure would cause a cusp in the light curve in exactly the same way as a helical structure. An alternative to a helical filamentary structure within the jet is a helical jet which would produce a qualitatively similar light curve, but would have the Doppler factor varying with position along the jet as the viewing angle relative to the local jet direction changes.

### 5.2 Conical Shocks

Lind & Blandford (1985) have considered the possibility that hotspots seen in VLBI images of radio jets may actually be relativistically moving conical shocks. They applied the relativistic shock jump conditions to a plane parallel flow entering a forward conical shock with cone angle $\eta$ to find the angle $\xi < \eta$ at which the flow initially diverges with respect to the jet axis. Defining the down-stream region between the cones with angles $\xi$ and $\eta$ to be the emission region, and taking account of Doppler
The generation of conical shock structures is often seen to occur in simulations of relativistic jets after the introduction of a fast perturbation (Bowman 1994; Gomez et al. 1997). These structures, which are typically an alternating sequence of forward and reverse conical shocks, are stationary or slowly moving in the galaxy frame, and have cone angles $\eta \sim 1 / \Gamma$. Salvati, Spada, & Pacini (1998) discuss emission by a conical shock. They consider the case of a density perturbation, confined to thin flat disk, travelling relativistically along the jet causing particle acceleration and emission where the disk cuts a stationary forward conical shock such as that illustrated in Figure 10.

Salvati et al. (1998) assume that the timescales for acceleration and emission are negligible. They show that boosting, they model the brightness distribution to simulate VLBI images.
their model can lead to highly peaked light curves whose shape depends on the viewing angle, and used these light curves to fit Mrk 421 TeV flare data. For the same input, and assuming there is no Doppler boosting of the emitted radiation, I am able to reproduce exactly their Figure 2 which shows the observed flux for various viewing angles. However, bearing in mind that the shock is stationary and that the pile-up in observing time is already included, the bolometric flux emitted by part of the downstream flow will be Doppler boosted by $D_{\text{local}}^{-1}$, where $D_{\text{local}} = \left(\Gamma d_1(1 - \beta_1 \cos \theta_{\text{local}})\right)^{-1}$. $\Gamma_2$ is the Lorentz factor of the downstream flow, and $\theta_{\text{local}}$ is the viewing angle with respect to the line of sight and the local downstream flow direction which varies around the shock as indicated in Figure 10 ($\theta_{\text{local}}$ ranges between $\theta_2$ and $\theta_1$). To obtain the Lorentz factor of the downstream flow, it is easiest to Lorentz transform in a direction parallel to the shock plane to a frame in which the flow is normal to the shock.

For the case of cone angle $\eta = \sin^{-1}(1/\Gamma)$ assumed by Salvati et al. (1998), i.e. an oblique shock at angle $\eta$ to the upstream flow, and using the relativistic equation of state, I find that for $\Gamma = 10$, $\eta \approx 7.25^\circ$ and $\eta \approx 1.89^\circ$ and that the Lorentz factor of the downstream flow is related to that of the upstream flow by $\Gamma_2 \approx 0.8017841 \Gamma_1$. My result for the observed flux for the same input as Salvati et al. (1998), but including the Doppler boosting taking into account the local downstream flow directions, is given in Figure 11(a) and shows that the inclusion of Doppler boosting causes the peak at $\theta = 0$ to be higher than that at $\theta = 0.9\theta_2$, the opposite to that found by Salvati et al. (1998). However, since they used one viewing angle in their fits to Mrk 421 TeV flare data (their Figure 3), and because the divergence of the downstream flow is rather small, their fits are still valid and their model remains an interesting mechanism for flare production. The same authors (Spada, Salvati, & Pacini 1999) have applied their model to IDV sources and are able to explain brightness temperatures up to $3 \times 10^{17}$ K.

I wish to extend the work of Salvati et al. (1998) by including the reverse shocks, and ultimately a sequence of stationary reverse and forward shocks. In Figure 11(b) I show the light curve for a reverse conical shock with respect to the jet axis, the angle between the downstream direction and the line of sight varies between $\theta_2$ and $\theta_1$.
jet flow is always parallel to the jet axis. The light curve due to a thin, initially flat, disk travelling relativistically along the jet is then calculated by the Monte Carlo method, in which 'particles' are placed over the surfaces of all of the cones. Each cone has the same number of particles, and they would appear uniformly distributed across the cross section of the jet when viewed along the jet axis. A side view showing the location of these particles is given in Figure 12. An initially flat thin disk is launched along the jet with initial Lorentz factor $\Gamma_1 = 10$, and each time a part of the disk crosses one of the shocks the Lorentz factor of that part of the disk drops appropriately, distorting the disk. The time at which each 'particle' emits its photon is determined by the time at which the (distorted) disk reaches the 'particle'. The resulting light curve for various viewing angles is shown in Figure 13(a) which gives the light curve from Figure 13(b) which gives the light curve from Figure 13(a) for $\theta = 9^\circ$ on a linear scale.

As can be seen, the light curve is quasi-periodic. It would be strictly periodic if the downstream flow velocity were identical to the upstream flow velocity. This might be the case if the re-collimation of the diverging flow from the conical shocks results in restoration of the flow Lorentz factor to near the upstream value, and in this case the flux from successive cycles would be at about the same level, as observed in Mrk 501 flares. The two flow velocities would also be roughly the same if the initial jet Lorentz factor were much higher than $\Gamma_1 = 10$ used in Figure 13(a) such that the intervals between flares due to a pair of conical shocks was roughly the same. However, in this case the flux of successive flares would diminish as the Doppler factor decreases. Turning this argument around, we may have a method of determining the minimum jet Lorentz factor from observational data, e.g. the $\sim 23$ day periodicity in the 1997 Mrk 501 data (Protheroe et al. 1998; Hayashida et al. 1998) may be used to put a lower limit on $\Gamma_1$.

Figure 14 shows an example light curve due to multiple jet perturbations. In this case, I have used an exponential distribution of times between the injection into the jet of a density perturbation with a mean interval of $r/c$. I have also used an exponential distribution for the strength of each perturbation. The resulting light curve looks, at least qualitatively, as good as any model for the variability of fluxes from blazars. Note that even though there is no strict periodicity resulting from the excitation of a sequence of shocks in this case, each flare episode has two strong peaks due to (for this viewing angle) the apex of the first reverse
Factors Determining Variability Time in Active Galactic Nucleus Jets

6 Conclusion

Many factors can influence observed variability time. The connection between $\Delta t_{\text{obs}}$, Doppler factor, and emission region geometry is non-trivial, and so measuring $\Delta t_{\text{obs}}$ may give, at best, one of the dimensions of the emission region. Using $R' = c \Delta t_{\text{obs}} D$ and $U' = U'/4 \pi R'^2 c$, the jet-frame photon energy density may then lead to overestimation or underestimation of the jet-frame photon energy density by orders of magnitude. This is clearly of importance in any AGN model in which the low energy photons produced in the jet are targets for interaction of high energy particles or radiation, such as in SSC models and hadronic blazar models. Although not discussed in detail in the present paper, the escape of $\gamma$-rays from the emission region depends on the optical depth to photon–photon pair production interactions. This optical depth can be uncertain by orders of magnitude in the same way as the photon energy density, and will also depend on viewing angle. One must therefore be careful when using the observation of apparently unattenuated $\gamma$-rays, and an observed variability time, to place limits on the Doppler factor.

The uncertainty in the jet-frame photon energy density discussed in this paper may also have implications for the high brightness temperature/Compton catastrophe problem of IDV sources. In this case, it is the energy density of target photons which limits the brightness temperature through the competition of inverse-Compton scattering with synchrotron radiation, and the target photon energy density may actually be lower than estimated if the emission region is non-spherical.

If the jet is bent or helical, or has some other favoured geometry (e.g. conical shocks), cusps in $t'$ vs. $t_{\text{obs}}$ and/or a varying Doppler factor may cause narrow peaks in the observed light curve irrespective of other factors. Distinguishing between these cases from the observed light curve alone is likely to be difficult. One way of distinguishing whether a flare is due to (i) an emission region moving around a bent or helical path, or (ii) a shock exciting a curved, conical, or helical structure within a jet, is that in the first case the flare is caused by a change in viewing angle with respect to the motion of the emission region leading to a change in Doppler factor, whereas in the second case the flare is due to a pile-up in observation times with no change in Doppler factor. Hence, in case (i) not only will the observed flux increase during a flare, but the photon energies will also increase — in $(\nu F_{\nu})_{\text{peak}}$ by factor $x^2$ accompanied by a shift in $\nu_{\text{peak}}$ by factor $x$ ($x$ is ratio of final to initial Doppler factor). In case (ii), however, since there is no change in Doppler factor there should be no shift in $\nu_{\text{peak}}$ accompanying an increase in $(\nu F_{\nu})_{\text{peak}}$.

Distinguishing between the excitation of a conical and a helical structure by a plane shock would be almost impossible — note the qualitative similarities between light curves depicted in Figures 9(d) and 13(b). Relativistic jet simulations (Bowman 1994; Gomez et al. 1997) do show the presence of conical shocks, and these shocks appear after a large perturbation (e.g. from 4 to 10 in the simulations of Gomez et al. 1997) in Lorentz factor of the matter entering the jet through the nozzle. This can result in a sequence of quasi-stationary to superluminal reverse and forward conical shocks extending from the nozzle to the perturbation as it moves along the jet (Agudo et al. 2001). For the conical shock model (Salvati et al. 1998) discussed here to work, a subsequent perturbation would need to result in a plane shock, or plane thin perturbation of some kind, which could travel along the jet and excite the pre-existing conical shocks. As far as I am aware, whether or not this could occur has not been demonstrated. A similar uncertainty hangs over whether or not helical jet structures, which may themselves be shocks with a twisted ribbon topology, resulting from a perturbation entering the jet through the nozzle, could subsequently be excited by the passage of a plane shock or perturbation. Nevertheless, in both cases, if the viewing angle is favourable pile-ups in $t_{\text{obs}}$, and hence flares, could occur simply as a result of the motion of the conical or helical patterns which may themselves be sites of enhanced emission. Note that in recent 3D relativistic jet simulations, the introduction of a 1 per cent helical velocity perturbation at the nozzle results in a helical pattern propagating along the jet nearly the beam speed (Aloy et al. 1999), and that the conical shocks resulting from a perturbation jet Lorentz factor can range from being quasi-stationary to superluminal (Agudo et al. 2001).

Figure 14 Light curve due to repeated excitation of the sequence of stationary conical shocks discussed in the text for an initial jet Lorentz factor $\Gamma = 10$ as seen at viewing angle $\theta = 9^\circ$. The time interval between perturbations entering the jet is sampled from an exponential distribution with mean $1/\gamma$, and the strength of the perturbation is also sampled from an exponential distribution.
In conclusion, in models for flaring in AGN in which the emission comes from a localised region (blob) co-moving with the jet, time variability is non-trivial to interpret in terms of emission region geometry and Doppler factor. A further complication is that flaring may arise instead due to curved or helical motion of a blob, even if the emission is constant in the instantaneous rest frame of the blob. In this case, apparent flaring is due to the change in viewing angle, and hence Doppler factor. Similarly, if the viewing angle is favourable, relativistic motion of curved or helical filaments or surfaces can lead to observation of flares. Excitation of curved or helical jet structures by shocks or perturbations can also lead to pile-ups in $t_{\text{obs}}$, and hence large apparent increases in flux. Observations of time variability in AGN are therefore non-trivial to interpret and may lead to large systematic errors in estimated jet-frame photon energy density, Doppler factor, and the physical parameters of the emission region.

Acknowledgments

I thank Peter Biermann for helpful discussions and Alina Donea for a careful reading of the manuscript. I would like to thank the anonymous referees for their suggestions which have led to improvements to the paper. My research is supported by a grant from the Australian Research Council and a grant from the University of Adelaide.

References

Kardashev, N. S. 2000, ARep, 44, 719
Kifune, T. 2002, PASA, 19, 1
Konopelko, A., et al. 1999, APh, 11, 135
Mannheim, K., Protheroe, R. J., & Rachen, J. 2001, PhRvD, 63, 023001
Mücke, A., & Protheroe, R. J. 2001, APh, 15, 121
Mücke, A., Protheroe, R. J., Engel, R., Rachen, J., & Stanev, T. 2002, APh, in press