Published in Physica Scripta, Vol. 17, 265-274, 1978.

EXTENDED AND COMPACT EXTRAGALACTIC INTERPRETATION AND THEORY

Roger D. Blandford [1] and Martin J. Rees[2]

[1] California Institute of Technology, Pasadena, California, U.S.A.
and [2] Institute of Astronomy, Cambridge, England

Abstract. The interpretation of strong double radio sources in terms of the "beam" model is reviewed. Implications of this model for source evolution and the nature of radio trails are briefly discussed. A final section is concerned with deductions about the properties of compact and variable components in quasars and galactic nuclei.

Table of Contents

1. INTRODUCTION

The strong radio sources were, historically, the first evidence for violent activity in galactic nuclei. The observations of the extended and compact sources have been reviewed at this meeting by Dr Willis and Dr Kellerman respectively. This paper is concerned with the interpretative aspects of the subject. The discussion and literature references are by no means comprehensive, and our debt to numerous authors of articles and reviews on the topic will be apparent to any reader conversant with the literature.

2. Extended sources: theoretical interpretation

2.1. General observational constraints on models

It seems likely that the simplest and most important extended sources to understand are the strong "active" doubles like Cygnus A, characterised by hot spots at their outer edges. In at least three such sources, aligned radio structure has been discovered in the nucleus. This indicates that energy is still being supplied to the extended components and that the source axis can be accurately "remembered". Further corroborative evidence for this view is provided by the "tunnels" sometimes seen in some relaxed sources, the jets in M87 and 3C 273, and the prevalence of similar aligned structure within active nuclei. It is natural to interpret the hot spots as the place where this energy is in some sense dissipated, probably by means of the acceleration of relativistic particles and the amplification of magnetic field. The tails extending back towards the parent galaxy can then be regarded as the glowing embers from earlier activity.

We shall here concentrate on describing the so-called "beam" model. When a detailed model of this general type was first proposed in 1971, the only arguments in its favour were somewhat circumstantial: for instance, it had the appealing feature that the energy content of extended sources could accumulate over their entire lifetime, which necessitated power outputs of only $\lesssim 10^{46}$ erg s^{-1} (whereas other ideas current at that time had to hypothesise, *ad hoc*, a short-lived outburst of vastly higher power); and it obviated the problem of adiabatic losses which occur if "plasmoids" generated in the galactic nucleus move bodily out into the relatively diffuse components. But three other categories of evidence have subsequently supported the beam model:

(i) There are numerous cases of "bridges" or "tails" linking the hot-spots to the central galaxy, and of continuing non-thermal activity in the galactic nuclei themselves.

(ii) Higher-resolution maps show that the electron lifetimes in the hot-spots of Cygnus A, 3C 236, etc. are shorter than the likely source lifetime (and perhaps shorter than the light travel time from the centre). This indicates the need for continuing re-plenishment or *in situ* acceleration of electrons in the hot spots.

(iii) The evidence for compact central radio components aligned with the overall source axis, indicates beams with a well-defined and persistent orientation.

Some further recent improvements in our knowledge that are consistent with beam models (or raise difficulties for alternative theories) but cannot be claimed as adding unambiguous support, include:

(iv) Better upper limits on the gas density in extended components.

(v) Continuing failure to find compact (<< 1") features within extended components away from the galactic nucleus.

(vi) Evidence that the source axis may in some cases be correlated with the rotation axis for the galaxy.

(vii) Tentative evidence for optical emission from one or two radio components.

Although the beam model seems - in broad outline - in accord with the data we would warn against adopting it too uncritically, or favouring it too strongly over some alternative ideas. Few aspects are yet worked out in adequate detail; the supporting evidence is far from overwhelming; there is in any case no reason why all radio sources should necessarily involve the same mechanism; and the history of extragalactic radio astronomy has often vindicated the sceptic.

The general rough estimates of the energy content in radio components are the same in beam models as in any theory which attributes the radio emission to a synchrotron-type process. We shall not repeat these well-known considerations here.

2.2. Outline of the "beam "model

Continuous supply of energy into radio components was first invesigated by Rees [1] who proposed that a galactic nucleus contained a source of low frequency ($\lesssim$ 1 kHZ) electromagnetic waves (e.g., a cluster of $\gtrsim 10^6$ pulsars radiating magnetic dipole radiation at their rotation frequency). The radiation would escape most easily from the nuclear region along the rotation axis (cf. [2] and [3]) and a channel would steadily be evacuated along which the low frequency waves could travel. This cavity would probably not contain much plasma, and those particles that were present would become relativistic. The plasma frequency would then be much lower than that associated with the surrounding medium and the waves would be naturally self-focussed; giving, at least qualitatively, the observed collimation.

Low frequency waves could fulfill a dual function at the hot spots, being both capable of accelerating GeV electrons and providing an electromagnetic field for them to radiate in. However, a similar model [4] was unable to account for the observations of the Crab Nebula and it was discovered that these waves would almost certainly be at the mercy of a variety of rapid parametric instabilities and resonant absorption processes. Subsequently, a more general viewpoint was taken [5 - 7] and it was suggested that the working substance behaved approximately as a "light" fluid. That is to say, the beams consist of fast moving hot plasma, probably permeated by static magnetic fields and electromagnetic wave modes. (In fact, as emphasised by Scheuer [6], the results are qualitatively unchanged if the beam consists of particles that undergo reflections at the channel walls, or relativistic particles streaming along a diverging magnetic field and conserving their adiabatic invariants.)

Although the collision mean free paths of relativistic particles in radio sources are very long, a fluid treatment can be justified because we know magnetic fields are present. The gyroradius of a relativistic particle of energy ϵ moving in a magnetic field B (Gauss) is ~ $10^{-12}(\epsilon / m_p c^2)B^{-1}$ pc. This is typically very small, compared to all relevant scales, for $B \gtrsim 10^{-6}$ G. Collective plasma effects also may reduce the effective mean free path (cf. the solar wind, which can in many contexts be regarded as a fluid, even though collisional lengths are very large).

If all the particles in the beam are ultrarelativistic, then $p = 1/3 \rho c^2$ and the sound speed is $c_s \simeq (c / \sqrt{3})$. Bernoulli's equation is then simply $\gamma_{bulk} = (1 - v^2 / c^2)^{-1/2} = (P / p_0)^{-1/4}$; and the mean random energy per relativistic particle, measured in the moving frame, varies as γ_{bulk}^{-1}. If the magnetic field contributes significantly to the total energy density and has a preferred orientation, then the pressure and magnetosonic velocities are of course anisotropic. The Debye length is also very small compared with the scale of the flow. This means that the relativistic plasma must be essentially neutral (in contrast to laboratory-scale electron beams) and the relativistic generalisation of ordinary MHD is applicable. There will be no electric field in a frame sharing the means plasma velocity ν; but in a *non*-moving frame, of course, the electric and magnetic field energies would be comparable if $\nu \simeq c$. In fact, it is not necessary for the model that the sound speed c_s in the beam actually be relativistic: the only requirement is that c_s be larger than the gravitational escape velocity, so that the direct effects of gravity on the beam material are negligible.

Suppose that a collimated beam has been established (by processes which we consider later). At a given time, the beam will have evacuated a tube or channel out to some location where it impinges on the external medium at a "working surface" which itself advances out at speed V. If the power flowing in the beam, L, is approximately conserved and stationary, then approximating the channel as a cylinder of radius r, we balance momentum fluxes at the "working surface" to obtain

$$\frac{L}{\pi r^2 \nu} \approx \rho_{\rm ext} V^2, \quad \nu \gg V \gg c_s. \tag{1}$$

where ν is the speed of the beam, ρ_{ext}, the external density, and c_s the internal sound speed. If the beam consists of relativistic plasma then $\nu \simeq c$ and relativistic fluid mechanics must be used. The pressure, p on the walls of the channel is given by

$$p \approx 0.7 \frac{L}{\pi r^2} \left(\frac{c_s}{\gamma_{\rm bulk} \nu} \right)^2 \ll \rho_{\rm ext} V^2. \tag{2}$$

This must be approximately balanced by a static external pressure if the walls are not to expand, any difference in pressure Δp leading to a transverse expansion speed $(\Delta p / \rho_{ext})^{1/2}$. However the ratio of the energy to the momentum supplied by the beam over the lifetime of the source is ~ V which exceeds the energy/momentum ratio required to sweep away the external medium by ~ $\nu / V >> 1$. The surplus (or waste) energy must then not accumulate near the "working surface" but be deposited within a "cocoon" surrounding the beam, which it is natural to identify with the low-surface-brightness tails.

The beam velocity is generally assumed to be supersonic and one possible way of maintaining this beam in a quasi-stationary state, and *creating* a high Mach number flow, is by means of a de Laval nozzle [7]. Suppose that there exists some continuous source of hot fluid in the nucleus, surrounded by a denser material trapped by the gravitational potential well. At first an almost spherical bubble will be inflated which can expand most rapidly along the rotation axis and eventually will be able to escape from the nuclear region. If the source is sufficiently powerful, two anti-parallel channels of hot fluid will be set up which can eventually provide the continuous energy supply for the radio components. If the flow is assumed to be stationary and isentropic then the fluid velocity will increase as the pressure decreases. When the pressure has halved the flow becomes transonic and the cross-sectional area is minimised. In this way a directed nozzle can be established. The radius of the channel at the nozzle is related to the total energy discharge, L, stagnation pressure, p_0, and sound speed, c_s by:

$$r^* \simeq 10^1 \left(\frac{L}{10^{46}\ {\rm erg\ s^{-1}}} \right)^{1/2} \left(\frac{P_0}{10^{-4}\ {\rm dyne\ cm^{-2}}} \right)^{1/2} \left(\frac{c_s}{10^5\ {\rm km\ s^{-1}}} \right)^{1/2} \ {\rm pc} \tag{3}$$

We return later to discuss further the collimation mechanism (the "nozzle" is only one of several possibilities) and the scale on which it is established. As far as the extended components are concerned, however, all that is necessary is that a collimated supersonic ($\nu > c_s$) beam be set up within a scale $R_* \lesssim 1$ kpc.

One obvious question concerns the possible seriousness of Kelvin-Helmolts instabilities in the vortex sheet between the beam and the external medium or cocoon. This has been discussed by Turland and Scheuer [8] and Blandford and Pringle [9], but until the physics of the boundary layer can be understood, or some relevant experiments can be performed, the stability of the supersonic portion of the flow must remain an open question. Nevertheless, the calculations and general physical arguments do indicate that very high Mach number flows are more likely to be stable for longer distances than mildly supersonic and subsonic jets. This is an argument in favor of a collimation mechanism in which there is no subsonic regime in the flow.

The linear sizes of the hot spots in the active sources are typically 3 kpc, and a perfectly isentropic fluid beam would in fact be focussed to a width $\lesssim$ 100 pc, and so there is some leeway for beam widening by entrainment processes and shock heating. At the hot spots, it is presumed that the bulk energy is efficiently randomised either through a strong transverse shock or by surface disruption of the beam. The general type of flow pattern expected would involve a shock where the beam energy is randomised, a contact discontinuity between shocked beam material and the external medium, and a stand-off shock moving into the extragalactic medium (whose ambient sound speed is $<< V$). This pattern is illustrated in refs. [7 - 10] (though we would expect much greater irregularities and asymmetries in any real situation).

While it is conceivable that the magnetic field in radio sources was already present in the extragalactic medium, it can readily have been transported out from the galactic nucleus along with the beam. As discussed by Blandford and Rees [7], if the central engine produces a wind in which magnetic and kinetic energy densities are comparable, this ratio can be preserved (or even enhanced) despite many orders of magnitude decrease in the plasma density as the beam moves outward: although the parallel component of B drops as r^{-2}, the perpendicular component goes as r^{-1} as the beam widens. The magnetic field would thus be predominantly transverse to the source axis in the beam and in the hot-spots; in the cocoon it would be sheared into a direction tangential to the boundary (consistent with what is observed). Note that such shearing motion can amplify a field up to equipartition, but that its dynamical effects then provide feedback which prevents it from ever becoming stronger. Although the field would have a preferred orientation, leading to high linear polarization of the synchrotron radiation, there may be many reversals. The scale of such reversals would depend on the character of the central source. If the "hot fluid" in the beam were supplied by multiple supernovae, or resulted from tidal disruption of random stars by a massive black hole, then the sign of the magnetic flux would be uncorrelated over scales containing $\gtrsim 10^{53}$ erg of energy; but a more organised field could arise if the power supply involved a single massive object, or accretion from the general interstellar medium in the central galaxy.

It is unclear what mechanism reconverts the bulk kinetic energy of the beam into relativistic electrons (with the requisite power-law spectrum) at the "working surface", but the kind of acceleration which almost certainly attains $\gtrsim$ 1 percent efficiency in, for instance, the supernova remnant Cass A (where the velocities are only $\lesssim 0.02c$) could be even more effective behind shock fronts where the velocities are much higher. Recently, some very suggestive arguments have been proposed [11, 12] according to which relativistic particles can be accelerated with roughly the observed spectrum by shock fronts.

If the flow pattern were sufficiently stable, a typical double source would evolve towards increasing size as the "hot spots" move outward, leaving a sheath or cocoon of lower surface brightness along their track; but eventually the central source would switch off, the residual relativistic plasma expanding and merging into the intergalactic medium. (The final stages of this process, and its cumulative impact on the intergalactic medium, deserve further study). The precise time-dependence of source size and radio luminosity depends on the external density, on how well the beams are collimated, etc. If this type of model applies to the "giant" double source 3C 236, the beams must have lasted, with mean power ~ 10^{45} erg s^{-1} for $\gtrsim 10^8$ yr.

Among the uncertainties and complications that bedevil any comparison with real double sources are the following:

(i) Instabilities are bound to complicate the flow pattern. (One would in fact wish to invoke some instabilities in order to explain the irregular and asymmetric structure of real sources. Also, the fact that the "hot spots" are not even smaller can best be explained by supposing that instabilities at the beam boundary have led to frictional heating and entrainment which causes r to increase faster than in the idealised isentropic case.) We are pessimistic about the prospects of firm theoretical progress in this area: it is hard enough to reproduce observed phenomena even in controlled experimental situations (water jets, etc.); and in the radio source context not only is the physics more complex (effects of compressibility, magnetic fields, etc.) but the relevant parameters (external gas density, pressures, etc.) are themselves uncertain. Perhaps wind-tunnel experiments may provide a closer analogy to real sources than the over-idealised models amenable to theoretical study.

(ii) The central power supply may have fluctuated over the source lifetime (and the distinction between an unsteady beam and a "multiple plasmoid" model [13] is only semantic as long as the interval between ejection of the plasmoids is short enough for a permanent channel through the surrounding medium to be maintained.

(iii) Asymmetry or instabilities of the collimation process may make the intensity of the two beams unequal. Conceivably "flip-flop" behaviour, where the plasma outflow squirts alternately in two opposite directions [14], could explain jets such as those in 3C 273, M87, and NGC 6251.

(iv) Inhomogeneities in the extragalactic medium may cause complex structure in the hot spots, as is seen in 3C 390.3: part of the beam may lag behind the rest, or be deflected, if it encounters a dense external cloud.

(v) Transverse motions of the external medium relative to the galaxy could destroy the symmetry or linear structure, especially when the beam is so weak, or so poorly collimated, that its speed of advance is $\lesssim$ the transverse velocity. This is perhaps relevant to the interpretation of "radio trails", as we discuss more fully below. Transverse or shearing motions of the external medium would particularly affect the final stages of a source's life when the expansion is slower. Buoyancy effects whose role in radio source morphology was first emphasised by Gull and Northover [15] - may also be significant, especially for sources in clusters or associated with unusually massive galaxies. Diffusive escape of relativistic particles may be important in large "relaxed" doubles if the streaming speed of the electrons is limited to the Alfven speed. So if all the relativistic electrons in an extended source like DA 240 originate in a central hot spot, either large-scale convective motion or re-acceleration is called for.

If the beams varied on timescales 10^4 - 10^7 yr for the reasons cited under (ii) or (iii), the cocoon, which delineates the path traced out by the hot spots over the history of the source, would be non-uniform: it would be particularly conspicuous in places where the hot spots were located at times of high beam intensity. The cocoon could thus resemble a series of blobs linking each outer hot spot to the central galaxy. (This is perhaps relevant to the interpretation of so-called "double doubles".)

2.3. Interactions with the extragalactic medium

One general feature of beam models (and, indeed, of any model which invokes ram pressure confinement of extended components) is that the maximum surface brightness attainable by the hot-spots is correlated with the density of the external medium. It is hard to quantify this relation precisely: it depends on the value of V, and on the fraction of the internal pressure that is provided by the relevant relativistic electrons and magnetic fields. It would, however, be a serious embarrassment to any such model if Cygnus A (or any other powerful double with equally intense "hot spots") were found to be definitely *not* in a cluster; though there is no such difficulty with a source such as 3C 236, where the internal energy density is ~ 100 times lower and can therefore be confined (for the same V) by a proportionately more rarefied external medium.

We suggest that this consideration is relevant to two important features of the data:

(i) As described by Willis (these proceedings), Gavazzi and Perola find that the most powerful sources ($P \gtrsim 10^{25}$ Wm^{-2} Hz^{-1} ster^{-1}) have linear sizes in the range 100-250 kps (for H = 100 km s^{-1} Mpc^{-1}). The upper cut-off could be related to the scale height of the gas in a cluster of galaxies: even if the parent galaxy were sited in the centre of a cluster, and even if the central activity persisted, the beams would no longer be ploughing into dense gas (and would therefore no longer form Cygnus A-type hot-spots at the "working surface") when their distance exceeded a scale-height. (Note, however, that there are other possible interpretations of the observed correlation between P and linear size. For instance, there may be some reason why instabilities always broaden the beam before it penetrates beyond a certain distance).

(ii) The overall linear extent of double sources, and the compactness of their hot spots, may depend systematically on redshift, though the existing data are still tentative and ambiguous. Any such dependence may arise from (a) changes with epoch in the scale height of cluster gas, and (b) changes in the gas density in the source environment. As regards (b), the assumption has often been made that the relevant gas density scales just as $(1 + z)^3$; but we wish to emphasise that things may be much less straightforward. The gas density *in clusters* may not have decreased at all between $z \simeq 2$ and the present epoch. Indeed it may even have *increased*, owing to infall, ejection from galaxies, etc; and we cannot exclude the possibility that the gas density surrounding strong extended sources at $z \simeq 2$ is typically *lower* than that for Cygnus A. Data from HEAO B may illuminate this question, but at the moment we cannot say whether hot spots should be systematically *more* or *less* compact in sources at large redshifts. If, as suggested in (i), the distribution of overall linear size depends on the scale height of cluster gas, this also may depend on z in an uncertain way. It is of course already quite widely appreciated that extended sources may be influenced by inverse Compton losses on the microwave background ($\propto (1 + z)^4$) and by any dependence on cosmic epoch of the typical central power-supply and collimation mechanism. These cumulative uncertainties will certainly make it a long time before extended sources can constitute a worthwhile tool for "geometrical" cosmology.

2.4. "Jets"

Jets are indicative of some kind of asymmetry in the collimation mechanism. The optical jets in (e.g.) M87 and 3C 273 raise problems of their own, but the radio emission from the jets in NGC 6251 and NGC 315 may come from the beam itself rather than a "cocoon". There are two possibilities:

(a) The emission may come from particles accelerated by dissipation processes at the "walls" of the beam; or

(b) The electrons in the beam itself may not have been completely cooled by radiative or adiabatic losses during the out-flow from the nucleus, so that they are still able to contribute synchrotron emission.

Alternative (a) would predict limb-brightening and a magnetic field *along* the jet: but (b) would predict a magnetic field *perpendicular to* the jet.

2.5. Radio trails

Radio trails are associated with clusters of galaxies from which X-ray emission is often observed. The most probable explanation of this emission is that it is bremsstrahlung from a hot, comparatively dense thermal plasma with $n_e \simeq 10^{-3}$ cm^{-3}, $T \simeq 10^8$ K exerting a fairly large pressure. Furthermore, the random (with respect to the Hubble flow) velocities of cluster galaxies are much larger than those of field galaxies. These two facts support the natural interpretation of the radio morphology of a trail as an ordinary low-power double source whose radio-emitting regions have been swept backwards by the surrounding medium. In a sense, radio trails represent a series of experiments that have been carried out on double sources, and so may prove to be especially useful in explaining them.

The beam interpretations of radio trails is straightforward. If the channel is exposed to a transverse momentum flux, the beam will be bent, perhaps through a combination of weak oblique shocks and rarefaction waves. If the motion of the galaxy with respect to the surrounding medium is supersonic, then there will be a cylindrical bow shock in front of each channel. The flow downstream will be fairly complex and, at least in some circumstances, turbulent. Conditions are then ideal for particle acceleration and field amplification close to the galaxy. We expect that the particles will subsequently cool by synchrotron radiation and by inverse Compton scattering of the microwave background. Steepening of the radio spectrum at the ends of the trails, consistent with an equipartition field strength of a few microgauss, is what is generally observed. It is possible to estimate the momentum discharge, $\dot{\Pi}$, in the beam from the force, (~ $\rho_{ext}\ V^2_{gal} lW$), due to the intergalactic medium, of density ρ_{ext} relative velocity V_{gal}, striking the beam of length l and width W. The energy discharge $\dot{U}$ can be estimated from the rate of increase of internal energy in the beam (~ $3/2 P_{ext} A V_{gal}$) where P_{ext} is the external thermal pressure, and A, the cross-sectional area of the trails. The mass discharge $\dot{M}$ can be estimated on the basis of Faraday rotation studies. It is comforting that within the (large) uncertainties of these determinations [16] the power $\dot{U} \simeq 10^{43}$ erg s^{-1}, force $\dot{\Pi} \simeq 10^{35}$ dyne and mass loss $\dot{M}$ ~ 1 $M_{\odot}$ yr^{-1} approximately satisfy $\dot{\Pi}^2 \simeq 2\dot{U}\ \dot{M}$ (l and W are estimated from the dimensions of the radio contours). With these numbers, the velocity in the beam (~ $\dot{\Pi} / \dot{M}$) is ~ 10,000 km s^{-1}, and the Mach number (~ $(l / w)^{1/2}$) has the fairly modest value ~ 3. In particular, the beams are unlikely to be relativistic. It may be possible to perform informative gas dynamical experiments to improve these simple arguments. In particular it would be interesting to determine under what conditions the flow downstream contained a large shear, parallel to V_g, as this is probably necessary if large degrees of linear polarisation are to be achieved at the ends of the trail.

The plasmoid model of radio trails was first considered by Jaffe and Perola [17] who calculated the deceleration and subsequent motion experienced by pairs of gas clouds shot out into the surrounding medium. This yields the shape of the tail and, with further assumptions (e.g., that the internal energy of the cloud is predominantly in the form of relativistic electrons and magnetic fields), the variation of surface brightness and spectral index along it. The fit to the observations of 3C 129 is reasonably good, yielding plasmoid masses and velocities ~ $10^7\ M_{\odot}$ and ~ 2000 km s^{-1} respectively. However the radio emission falls off less rapidly along the tail than is predicted by Jaffe and Perola's model, and so Cowie and McKee [18] have modified it to include the effects of thermal pressure. If the galaxy has a substantial mass loss rate, then there will be a stand-off bow shock and the plasmoid itself will not feel the full force of the intergalactic medium until it has moved some distance (~ 20 kpc). Both of these modifications can give an improved fit to the data. There is also the possibility of particle acceleration within the tail; and Pacholczyk and Scott [19] have identified those regions where the polarisation is comparatively low with high levels of particle-accelerating magnetic turbulence, thus accounting for the apparent association of high brightness with low polarisation and vice versa.

Observational support for *in situ* acceleration within the trails is provided by observations [20] of IC 711 which show that the time required for the galaxy to produce the observed trail, whose spectrum actually *flattens* at the end, exceeds by an order of magnitude the estimated cooling time.

Jaffe and Perola [17] also proposed an alternative model drawing an analogy with the interaction of the solar wind with the Earth's magnetosphere. They suggested that clouds of relativistic particles accelerated within the nucleus stream outwards along the polar field lines of a galactic dipole, transverse to its velocity vector. Close to the nucleus the clouds are protected, but further out the field lines will be swept backwards to form a long magnetospheric tail. This field, varying as r^{-3}, will be quite large within the nuclear regions and it seems difficult to avoid large radiative and adiabatic losses. Furthermore a substantial nuclear mass is called for to provide an "anchor".

There is in fact no need for any energy to be supplied by the galactic nucleus. The total power dissipated across a bow shock by the galaxy travelling through the intergalactic medium can be up to ~ 10^{44} erg s^{-1} and only 1 per cent of this (a comparable efficiency to galactic supernova remnants) need be dissipated in the form of relativistic electrons and magnetic fields to account for the radio emission. Gisler [2l] proposes that radio trails are simply radio emitting regions associated with a bow shock seen in projection. This explanation encounters some difficulty in explaining the maps of 3C 129 and 3C 83.18 because the radio features do not appear to "stand off" from the nucleus of the galaxy, and also because the contrast between high and low brightness regions may be too large to be explicable in terms of projection without postulating large departures from axisymmetry. However this remains a viable explanation for sources that are less well resolved.

2.6. Collimation

All the suggestions that have been made for collimating the beams involve the supposition that the required buoyant fluid of hot (possibly relativistic) plasma escapes along the direction of least resistance, which is likely to be the rotation axis of some central object, disc or gas cloud. We mention three classes of mechanism.

2.6.1 *De Laval nozzles.* These were first discussed in detail by Blandford and Rees [7]. Aside from uncertainty about the stability of the configuration to Kelvin-Helmholtz *and* Rayleigh-Taylor instabilities, there are observational constraints on the properties of the gas cloud, which has to be in pressure balance with the outflowing fluid. Blandford and Rees, who considered the specific example of Cygnus A, envisaged a gas cloud with scale height $R_* \simeq 100$ pc supported by gas pressure. This cloud would be a strong X-ray emitter, and the X-ray emission would have been enacceptably high if the scale height R_* were decreased. This question was discussed by Wiita [22], who showed that a somewhat smaller scale was possible. (The required density $\propto R_*^{-2}$, so the emission, which depends on $\rho^2 R_*^3$, varies as R_*^{-1}). The assumption that the external pressure is provided by the *ram* pressure of infalling matter, rather than the thermal pressure, eases the X-ray constraint, but free-free absorption ($\propto T^{-3/2}$) then becomes important. These values of R_* are larger than (or comparable with) the sizes of compact sources, and the parameters are insensitive to the character of the central energy source, provided that its size does not exceed a few parsecs.

As argued by one of us elsewhere in these proceedings, there now seem good reasons for identifying the energy source with a massive accreting black hole. One can then envisage the nozzle as being established on a much smaller scale (maybe $\lesssim 100$ Schwarzschild radii). The confinement would then be provided by an optically-thick cloud of infalling material supported primarily by radiation pressure, and the X-ray constraint is then irrelevant. In this model the timescale for variations in the collimation, and for changes in the external cloud, may be as short as a few days.

2.6.2. *Flaring disc* If the relativistic plasma were generated by magnetic flares in the inner region of an accretion disc [23], or by some exotic process such as $e^+ - e^-$ pair production near a Kerr black hole [24], (and perhaps channelled by large-scale poloidal magnetic fields), then it may already be sufficiently collimated to make its mean speed along the axis $> c_s$. There is then no need for a nozzle, though external pressure would refine the collimation as the beam penetrated further from the nucleus.

2.6.3. *Radiation-pressure-driven outflow* A third possibility, not necessarily completely distinct from (i), is that the beam consists of material accelerated out along the axis solely by the pressure of thermal radiation. If the accretion flow near the hole resembles a "settling solution", then the effective sound speed is $\sim c(r / r_s)^{-1/2}$. If angular momentum causes rotational flattening, then material blown off from near the hole will attain terminal velocities comparable with the escape velocity from its starting-point.

There is an appealing advantage to models in which the collimation is established very close to the black hole. The data on extended sources requires that the beam orientation should remain steady for $\gtrsim 10^8$ yr, apart from the slow wandering or precession indicated by the apparent "mirror symmetry" in some extended source components (see maps of Cygnus A, 3C 192 and NGC 315 in Dr Willis's contribution). If the confining gas comes from disrupted stars, random infall of gaseous debris, etc, there is no guarantee (and not necessarily any likelihood) that the orientation would achieve this long-term constancy. But near a rotating (Kerr) black hole, Lense-Thirring precession would cause the accreted material to acquire symmetry with respect to the hole's rotation axis (which can change only on the timescale $\sim 10^8$ yr on which its mass doubles by accretion). The details of this essentially relativistic effect have been discussed by Bardeen and Petterson [25], with application to compact X-ray sources; but in the present context, the effect would ensure constant beam orientation in any model where collimation occurred in the relativistic domain.

Note that, if the collimation is set up on scales $\lesssim 100 r_s$, even the most compact radio structure accessible to VLBI observations is very much a "secondary" phenomena, being on scales several orders of magnitude larger. The compact radio components may nevertheless provide important clues to the nature of the (much smaller) object that energises them, and we now turn to consider their properties in more detail.

3. COMPACT RADIO COMPONENTS

3.1. Angular size constraints

As with the extended sources, the primary radio emission is usually attributed to synchrotron radiation by relativistic electrons in a partially ordered field, although alternatives have been investigated. The evidence for this is perhaps not so good as in the extended sources, as the observed degrees of polarisation are much lower; however the majority of the sources can be satisfactorily interpreted using this hypothesis [26].

The source energetics are critically dependent on the angular size θ, which can be estimated in several different ways. As described by Dr Kellerman, VLBI can provide angular information on milliarcsecond structure, which can (with the aid of a cosmological model) be converted into a linear size and shape; and an expansion speed that apparently often exceeds *c*. Searches for interstellar scintillation, as yet undetected, provide lower limits on θ in the range ~ 1 - 10 *micro*arcsecond.

Secondly, and especially for sources that are too weak for VLBI, an upper limit on the angular size can be set from the variability timescale $\theta_{\rm var}$ (ignoring kinematic effects discussed below).

$$\theta_{\rm var} \lesssim \frac{c\tau_{\rm var}(1+z)}{D} \tag{4}$$

where D is the *luminosity distance* to the source ($D = H_0^{-1} q_0^{-2}[z q_0 + (1 - q_0)(1 - \sqrt{1 + 2 q_0 z})]$ in a Friedmann cosmology).

Thirdly, in some sources a low-frequency turnover, interpretable as synchrotron self-absorption, yields an estimate of the brightness temperature and hence of the angular size. (If a turn-over is not observed, the synchrotron hypothesis leads to a lower limit on the angular size.) Essentially the limit arises because when the source becomes self-absorbed the radiation brightness temperature $\propto S\,\theta^{-2}$ is comparable with the kinetic temperature of the electrons ($\propto (\nu_n / B)^{1/2}$) emitting at the frequency ν_n where the flux turns over. Numerically,

$$\theta_{\rm SA} \simeq \left(\frac{B}{10^{-4}\ {\rm G}}\right)^{1/4} \left(\frac{S_n}{1\ {\rm Jy}}\right)^{1/2} \left(\frac{\nu_n}{1\ {\rm GHz}}\right)^{-5/4} \times 10^{-3}\ {\rm arcsecond} \tag{5}$$

where $S_n = S(\nu_n)$. More detailed expressions correcting for cosmological and spectral effects are given by Jones, O'Dell and Stein [27]. More serious modifications are required with alternative radiation processes. For example with proton synchrotron radiation the brightness temperature in a source of fixed size and field strength is increased by ~ $(m_p / m_e)^{1/2}$ but only at the expense of imposing severe constraints on the relativistic electron density [28]. With inverse Compton scattering, the brightness temperature is decreased. Coherent mechanisms of course permit still higher temperatures.

A fourth limit on the angular size comes from considering the consequences of inverse Compton scattering of the radio photons by the same electrons that produced them. If the radiation energy density $U_{\rm rad}$ exceeds the magnetic field energy density $B^2 / 8\pi$, then a relativistic electron will lose more energy producing Compton photons, typically in the infrared or optical range, than radio. These Compton photons can themselves be scattered, and so on until their unscattered frequency satisfies $\gamma\, h\, \nu \gtrsim m_e c^2$ with γ the Lorentz factor of the electron that produced the original radio photon. At these frequencies, electron recoil is important and the cross-section is diminished, thus suppressing further scattering. Large Compton fluxes are undesirable for two reasons: they greatly increase the energy requirements of the source and they may come into conflict with existing observational upper limits, particularly at X-ray energies. This therefore implies that $U_{\rm rad} \lesssim B^2 / 8\pi$.

If we assume that the radiation is produced in a homogeneous spherical source of angular diameter θ, and that $\nu\ S(\nu)$ is maximised at some frequency ν_u with flux $S(\nu_u)$, typically in the mm range, then this condition becomes

$$\theta \gtrsim \theta_{\rm ic} \simeq \left(\frac{S(\nu_u)}{1\ {\rm Jy}}\right)^{1/2} \left(\frac{\nu_u}{1\ {\rm GHz}}\right)^{1/2} \left(\frac{B}{10^{-4}\ {\rm G}}\right)^{-1} \times 10^{-3}\ {\rm arcsecond.} \tag{6}$$

Strictly, we have no guarantee that the flux S_u arises from the same location and at the same time as S_n and so we can obtain a definite lower bound on the angular size by equating ν_n with ν_u

$$\theta \gtrsim \theta_c \simeq \left(\frac{S_u}{1\ \mathrm{Jy}}\right)^{1/2} \left(\frac{\nu_u}{1\ \mathrm{GHz}}\right)^{-9/10} \times 10^{-3}\ \mathrm{arcsecond} \tag{7}$$

or equivalently in terms of the observed brightness temperature, $T_n \lesssim 10^{12}$ K [29]. (It is a curious coincidence that this is about the maximum brightness temperature that can be detected from a ~ 1 Jy source using VLBI with an intercontinental baseline.)

As discussed in detail by Burbidge, Jones and O'Dell [26], and by O'Dell [30], the angles θ_{VLB}, θ_{var} and θ_c are in genera comparable, although the upper limit θ_{var} tends to be smaller than θ_{VLBI} and the lower limit θ_c tends to be larger. Thi encourages confidence in the electron synchrotron hypothesis. A second conclusion is that generally $B_{eq} > B_c$, and so the particle pressure exceeds the magnetic stresses. It is in general impossible to derive precise estimates of magnetic field strengths, energy contents, etc because all such quantities depend on very high powers of θ.

The estimated energy content of compact components is sometimes as large as ~ 10^{56} erg, assuming isotropic emission. These estimates (which are larger than could be obtained from a single supernova-style explosion) could be reduced by adopting some special geometry. But they are in any case much smaller than the energies involved in extended sources, which is consistent with the view that the latter are built up by nuclear activity over periods $\gtrsim 10^6$ yr.

3.2. Radiative transfer

So far we have assumed implicitly that the observed radiation is unchanged in spectrum and polarisation from what was emitted. In fact there are a variety of propagation effects which can substantially modify the emergent radiation and indeed in some sources we can obtain important physical constraints on the conditions within and around the radio components simply from our failure to detect these changes.

Free-free absorption will effectively suppress the emission below a frequency

$$\nu_{\mathrm{ff}} \simeq (EM)^{1/2} T_4^{-3/4}\ \mathrm{MHz} \tag{8}$$

where $EM = \int (n_e / 1\ \mathrm{cm}^{-3})^2 (dl / 1\ \mathrm{pc})$ is the emission measure of the thermal plasma, of temperature $10^4 T_4$ K, along the line of sight. In those sources, e.g., associated with quasars displaying prominent emission line spectra, typical values are $EM \sim 10^{12}$, $T_4 \simeq 1$ and so the emitting filaments cannot completely occult the radio source.

A second (fairly *un*important) limit on the thermal electron density can be set from the manifest absence of Razin suppression of the synchrotron emission. This gives

$$n_e \lesssim 6 \times 10^4 \left(\frac{B}{1\ \mathrm{G}}\right) \left(\frac{\nu}{1\ \mathrm{MHz}}\right)\ \mathrm{cm}^{-3}. \tag{9}$$

Of more interest are the limits derived from polarisation observations. Most sources display 1-10 per cent linear polarisation and $\lesssim$ 0.1 per cent circular polarisation. The fact that this much linear polarisation emerges and that its plane of polarisation does not rotate during the evolution of a source places an upper bound on the internal Faraday rotation of the source. Both thermal and relativistic electrons can contribute to the rotation and the expected depolarisation. The equation of radiative transfer has been solved in a variety of source models.

For a homogeneous source the Faraday rotation due to thermal plasma is $\Delta\phi \propto n_e BR / \nu^2$, where n_e is the thermal electron density; and the brightness temperature where the source is just optically thin is $T \propto n_{rel} Bl / \nu^2$ where n_{rel} is the density of relativistic electrons emitting at this frequency. We can therefore obtain a limit on $n_e / n_{rel} \lesssim \Delta\phi (3 \times 10^9 / T)$ *independent* of B, as most sources have $T \simeq 10^{12}$ K and $\Delta\phi \lesssim 1$ (otherwise they would be strongly depolarized). This limit is typically ~ 10^{-3} - 10^{-4} and is so small that the Faraday rotation due to relativistic electrons must be investigated. The rotation due to the relativistic plasma is $\Delta\phi \propto n_{rel} BR / \nu^2 \langle \ln\gamma / \gamma^2 \rangle$; and for the brightest sources, not showing changes in $\Delta\phi$, the electron distribution function must cut off below energies ~ 10-100 MeV [31, 32].

As radio components have a typical degree of polarisation $\lesssim$ 10 per cent of the theoretical maximum for a synchrotron source (~ 70 per cent) they must exhibit some structural de-polarisation, which will increase the limits on n_e and n_{rel}, as the field will reverse several times along the line of sight, reducing the rotation measure. It is in principle possible to reduce $\Delta \phi$ to a negligible value without destroying the polarisation simply by shearing a source with small scale magnetic turbulence, but in practice this is probably difficult to maintain for any length of time since the timescale for isotropising the field is only that taken by an Alfven wave to traverse the scale of an irregularity. A second type of inhomogeneity occurs if a lot of the mass is contained in filaments that do not completely occult the source. In this case the radio-emitting plasma is probably dynamically decoupled from the filaments. The natural inference is then that the radio emitting regions contain predominantly relativistic plasma.

A circular polarisation can also be produced as a consequence of propagation through a relativistic plasma [33, 34]. This arises because the normal modes of propagation through a plasma are not purely circularly polarised; and so, by a straightforward modification of the mechanism responsible for Faraday rotation, a linearly polarised wave emitted by the synchrotron process will be partially converted into a circularly polarised wave. If the relativistic electron distribution function is a power-law with *s* $\gtrsim$ 2 and extends down to some energy $\gamma_{min}\, mc^2$, then the expected degree of circular polarisation π_c produced by radiative transfer near turnover is approximately related to the Faraday rotation by

$$\frac{\pi_c}{\Delta\phi} \simeq \frac{\gamma_{\rm min}^3}{\gamma_n^2 \ln \gamma_{\rm min}} \qquad (10)$$

where $\gamma_n\, mc^2$ is the energy of electrons radiating at the turnover frequency. The sign of the circular polarisation generally reverses near turnover.

Circular polarisation is also produced intrinsically in the synchrotron process [35 - 37] and it too reverses near turnover. The expected magnitude is

$$\pi_c \simeq \frac{3}{\gamma} \simeq 3\left(\frac{B}{1\ {\rm G}}\right)^{1/2}\left(\frac{\nu}{1\ {\rm MHz}}\right)^{-1/2} \quad (\nu \gtrsim v_n) \qquad (11)$$

In an inhomogeneous source these estimates must be reduced like $\Delta \phi$. π_c can in principle be reduced to a negligible value without seriously diminishing the linear polarisation.

Reported values of π_c are in the range 0.1-1 per cent. If reliable upper limits < 0.1 per cent turn out to be very common, then they are probably too small to be understood in terms of a relativistic electron synchrotron source, and one intriguing possibility is that the source comprise equal numbers of electrons and positrons. If observations could be pushed to the point where this seemed the most attractive solution, then it would have important consequences for isolating the particle acceleration mechanism. The sources where one might expect detectable circular polarisation are those such as the low-power compact component of 3C 84 where the field may be $\gtrsim$ 0.1 G.

Another propagation effect that can profoundly influence the character of the emergent radiation is *induced* Compton scattering. If we consider scattering off free electrons, then induced scatterings [38, 39] are more important than spontaneous scatterings if the brightness temperature T exceeds $m_e c^2 / k \simeq 10^{10}$ K. (The rate of induced scattering out of a photon state exceeds the spontaneous rate by a factor of order the photon occupation number ~ $kT / h\nu$. However there are also induced scatterings into this state and so the net induced scattering rate is reduced by the relative Compton shift, $\Delta \nu / \nu \simeq (h\nu / m_e c^2)$). Typical brightness temperatures are $T \simeq 10^{12}$ K and so if the optical depth to spontaneous Thomson scattering, τ_T traversed by the emergent radiation lies in the range $10^{-2} \lesssim \tau_T \lesssim 1$, then induced effects can influence the spectrum without increasing the source size by spontaneous scattering. (It is necessary that the thermal plasma be close to the source as the induced scattering rate $\propto \theta^4$ where θ is the angle subtended by the source [40, 41]. This is because induced scatterings only occur between states that contain a high number of photons already. As the photons are "hotter" than the thermal plasma they will fall in frequency whilst conserving their total number, and if $(kT / m_e c^2)\, \tau_T^2 \gtrsim 1$ the spectrum will be severely distorted. Linear polarisation can also be created and destroyed by this mechanism [41]. However, comparison with eq. (8) shows that free-free absorption at ~ 1 GHz in a source of size ~ 1 pc is likely to occur before induced scattering unless the electron temperature is $\gtrsim 10^5$ K. (If free-free absorption became important it would in any case raise the temperature [42].) A second constraint on the physical conditions in such a scattering region is that the magnetic field strength be much less than in the source in order that the Faraday rotation not be greater than observed.

3.3. Kinematics of superluminal expansion

As described in the contribution by Kellermann, there is now good evidence for superluminal expansion in at least three compact sources, if we adopt the cosmological interpretation of large redshifts. A variety of physical models have been proposed to account for this phenomenon, which we now describe using a loose kinematic classification developed in the review by Blandford, McKee and Rees [43] (to which the reader is referred for further details and references to the literature).

When a radio source of size R changes in a time ~ R/c, then light travel time effects in the emergent radiation must be taken properly into account. The main requirements on a physical model, if the sources observed so far are typical, is to produce a double or aligned triple expanding with a speed $2c$ - $8c$, successive expansions occurring in the same direction. Although "acausal" models have been proposed in which there are two or more independent outbursts, the observed predominance of expansion over contraction raises problems for this interpretation. That superluminal expansion is indeed possible can be simply demonstrated by considering a "ballistic" model in which a shell of radio emitting plasma moves radially outwards from an origin `o' with speed ν. At a fixed *observed* time, the locus of the shell is a prolate ellipsoid with one focus at 0 and eccentricity ν/c. The observer therefore sees a circle of radius equal to the semiminor axis of the ellipsoid expanding with speed $2V(1 - V^2 / c^2)^{-1/2}$, which exceeds $2c$ if $\nu > 2^{-1/2}\ c$. In a straight-forward variation designed to produce an expanding double, two radio-emitting plasmoids or beams are ejected along antiparallel directions with the nearer making an angle ~ $(1 - \nu^2 / c^2)^{1/2}$ with the observer direction. An expansion speed $\nu_0 \simeq \nu\ (1 - \nu^2 / c^2)^{1/2}$ will be observed. However the probability of the observer having this orientation is only ~ $2c^2 / \nu_0{}^2$, which is typically very small when $\nu_0 >> c$. Extreme superluminal expansion would then have to be a comparatively rare phenomenon amongst a complete sample of compact radio sources.

In a simple example of a "screen" model a relativistic signal (e.g., a shock) illuminates a thin circular ring, (e.g., the inner edge of a disc of accreting matter), radius R. An observer in the plane containing the ring will see a double, of separation x, separating with a speed ν_0 ~ $(R / x)c; x << R$. Screen models are usually characterised by deceleration of the observed expansion, although more complex observations can be interpreted using suitably tailored screens. The signal motion need not be rectilinear. One ingenious suggestion [44] is that the emitting electrons are channelled at relativistic speeds along the north and south polar field lines of a magnetic dipole. They would then only be observed when the field line pointed towards the observer, and an outburst originating near the centre of the dipole would create an expanding double. (For an equatorial observer $\nu_0 = 4.4c$.)

In most simple types of screen model superluminal effects can usually be achieved without the observer having a privileged orientation (in contrast to ballistic models). Apparent velocities $\nu_0 >> c$ will however only be observed for a fraction ~ $(c / \nu_0)^2$ of the total emission time, and the actual size of the screen will exceed the observed size by a factor typically ~ (ν_0 / c).

There are many examples in physics of cases where the space velocity of some feature (although not the velocity of energy transport) can exceed c. For example the line of intersection of two almost parallel surfaces (e.g. shock waves), rapidly approaching one another, can be arbitrarily large. In "phase" models (e.g., Epstein and Geller [45]) the observed expansion is caused directly by such a mechanism rather than light travel time effects. Proper allowance for these effects will however in general result in a subrelativistic apparent expansion, unless the observer lies in a direction roughly perpendicular to the phase velocity, again with a rather low probability ~ (c/v_0).

As well as the observed rate of expansion, we must also consider the relative surface brightness of different parts of the source. Of course this depends in detail on a variety of astrophysical effects particle acceleration, magnetic field strength, energy and isotropy of the central outbursts etc.- of which our knowledge must be limited. One important effect that can be considered is the Doppler shift of the emitting plasma parametrised by the factor $D = (1 - \nu^2 / c^2)^{1/2}(1 - \nu_{em} \cos\theta / c)^{-1}$ where θ is the angle between the velocity ν_{em} of the emitting plasma (i.e., of the frame in which the emission is more or less isotropic) and the line of sight. ν_{em} need not be the same as the velocity of the signal. It can be argued that ν_{em} must be neither too small nor too large, and is optimally mildly relativlstic. If $\nu_{em} << c$, $D \simeq 1$ for all θ and so all regions accessible to the signal should radiate towards the observer, not just those displaying superluminal expansion. Unless the observer has a special orientation, the flux from the most rapidly moving components would be dominated by that from the slower moving background, contrary to what is observed. If ν_{em} is ultrarelativistic then only material moving with an angle $\theta \lesssim (1 - \nu^2 / c^2)^{1/2}$ will be seen and then it is difficult to explain how an observer in an arbitrary direction will see doubles expanding along a preferred axis.

A type of model that seems able to explain the existing data was proposed by Blandford et al. [43]. An outburst associated with the primary energy source initiates a relativistic signal (for instance, a burst of electromagnetic energy or a strong shock) which propagates outwards illuminating a suitably shaped screen (such as the walls of a channel) similar to that invoked in Section 2 to be responsible for the supply of energy to extended double radio sources. The burst gives the radio emitting plasma a mildly relativistic velocity, thus ensuring that only the front of the screen is highlighted. (Note that this mildly relativistic expansion velocity is possibly required to avoid excessive Compton losses, and is perhaps implied by the inferred dominance of particle pressure.) With this sort of geometry, separation speeds ~ 5*c* are achievable without the observer having a particularly privileged orientation. It remains to be seen if this type of explanation is adequate to account for observed sources.

These "intrinsic" kinematical models are not the only ones capable of reproducing superluminal expansion. Radio waves produced in a single spherically symmetric outburst can also be refracted or scattered so as to produce an apparently separating double. If the radiation is scattered through a small angle ϕ whilst propagating towards the observer, then the unscattered photons will arrive before the scattered photons and a roughly circular source of expansion speed ~ $2c / \phi$ will be produced. Unfortunately, in a model like this, the primary source size must be smaller than the observed size (implying coherent emission), in contrast to most intrinsic models, and this increases the ratio of the Compton scattered flux to the synchrotron flux. One such model, proposed by Wilson [41], involves a very compact variable source at the centre of a *non*-variable line source (which may be regarded as the inner part of the "beams" of an extended source) embedded in a cloud of thermal plasma. If *induced* scattering dominates spontaneous scattering in the cloud, then an outburst in the central component will be followed by an apparent "expanding double" of enhanced surface brightness moving out along the line source. This type of model has the virtue of being able to explain superliminous expansion along a fixed axis which is *not* close to the line of sight; but has the possible disadvantage that the apparent brightness of the expanding components would be strongly frequency-dependent.

The VLBI and variability data can generally be explained in terms of relativistic effects where the (phase or physical) velocities corresponding to Lorentz factors $\lesssim 5$. The main exceptions to this are spectacular variables such as AO 02354+164 and those sources - e.g., 3C 454.3 - which appear to vary at frequencies below 1 GHz.

3.4. Dynamical considerations

Although there are models which obviate the need for relativistic bulk motion, it is important to emphasise that such motions are by no means unreasonable - and might even be expected. The brightness temperature limits tell us that the magnetic field is too weak to confine the relativistic plasma. So, unless this is bound by a much higher density of thermal plasma (and this can probably be excluded on other grounds), relativistic expansion would inevitably ensue. So these arguments, which are independent of the observed angular structure (and were even given prior to the VLB observations), at least support models involving intrinsic relativistic expansion.

The dynamics of a relativistic plasma have been discussed in a few idealized cases:

3.4.1. *Free expansion of an initially uniform spherical cloud of relativistic gas* [47]. If an initially uniform cloud of ultra-relativistic plasma starts expanding from radius r_0, then the Lorentz factor of the bulk expansion when the sphere has attained a radius $r >> r_0$, will be ~ r / r_0, the material then being mostly in a thin shell whose thickness remains ~ r_0. The internal energy per particle, measured in a comoving frame, would of course decrease by the same factor ~ r / r_0.

3.4.2. *Steady relativistic wind* Bernoulli's equation for a spherically symmetric relativistic outflow requires that the Lorentz factor for the bulk flow be proportional to *r*. Thus each comoving shell in the wind behaves rather like the expanding sphere in (i). Note that, if such a wind were to make a transition from subsonic to supersonic flow on passage through a "sonic radius" where the *gravitational field* plays the role of the nozzle (by analogy with the solar wind) then this would need to occur at a place where the escape velocity were $c / \sqrt{3}$ - i.e., near the Schwarzschild radius of a relativistically deep potential well [47].

3.4.3. *Relativistic blast waves* If a "point explosion" causes relativistic expansion of a blast wave into a uniform surrounding medium, then $\gamma_{\text{blast}} \propto r^{-3/2}$. The swept-up particles acquire random energy, measured in a frame moving with the shock corresponding to $\gamma \simeq \gamma_{\text{blast}}$. These particles are concentrated in a shell of thickness ~ $ct\, \gamma_{\text{blast}}^{-2}$ ($\propto r^4$). When *r* gets so large that the shock stops being relativistic, there is a transition towards the ordinary Sedov solution. One can consider more general cases when the blast wave moves into a medium of non-uniform density or into a wind which itself has a relativistic outward velocity. This problem has been discussed in detail by Blandford and McKee [48].

Any of the above could be elaborated into a possible model or variable radio sources. Note, however, that one of the few things VLBI studies tell us unambiguously is that actual compact sources are *not* spherically symmetrical! Note also that the bulk expansion rate in models (i) and (ii) would in fact "saturate" at a finite γ equal to the mean initial Lorentz factor of the thermal motions of the ions. (If the plasma consists of electron-position pairs, the terminal γ could be much higher.) These models are applicable only if the external medium is rarified enough to be ignore. For a relativistic outburst of initial total energy ϵ, expanding with Lorentz factor γ, this requires

$$\frac{4}{3}\pi \rho_{\rm ext} R^3 c^2 \lesssim \epsilon / \gamma^2 \tag{12}$$

i.e., the particle density in the external medium must satisfy

$$n \lesssim 10^{-1} \left(\frac{\epsilon}{10^{54}\ {\rm erg}}\right) \gamma^{-2} \left(\frac{R}{10^{19}\ {\rm cm}}\right)^{-3}\ {\rm cm}^{-3}. \tag{13}$$

Otherwise the relativistic ejecta would be braked at too small a radius to account adequately for the radio structure.

3.5. Some observational tests

None of the current models aiming to account for superlight expansion seems entirely convincing, and the fact that few can yet be decisively excluded is solely a consequence of the limited data available. We list here some questions which, if answered, would help to narrow down the field.

3.5.1. *Does the position angle and/or the expansion speed vary during a single outburst?* Models involving relativistic bulk motions can produce either accelerating or decelerating expansions. Indeed for those sources where the component separation is observed to expand by a factor ~ 2, it would be surprising if some change in the observed velocity were not seen. Blast wave models are generally decelerating, whereas free expansion of relativistic plasma (which is perhaps unlikely because adiabatic cooling would cause a rapid decline in flux density at frequencies where the source is optically thin) may accelerate. The expansion rates for "screen" models are determined mainly by the shape and orientation of the screen, but if the outburst or screen are not axisymmetric, the position angle may change. One interesting possibility is that the velocity may be accurately constant during an outburst. In this case the expansion would probably have to be attributed to one or more coherent objects moving inertially, although under certain conditions blast wave models can exhibit constant velocity phases. The present data seem consistent with uniform expansion during an outburst.

3.5.2. *Are the position angle and expansion speed "remembered" from one outburst to the next?* Successive outbursts should display the same alignment if they represent expansion along a preferred direction (e.g., the rotation axis of a central massive object or rotating gas cloud). A reproducible expansion speed would be a strong argument in favour of a screen model, but the absence of such would not necessarily rule it out because the signal velocity could vary between outbursts, which would alter the kinematics. In fact the two outbursts in 3C 120 seem to have had different speeds.

3.5.3. *What are the absolute positions of the components and which ones are moving fastest?* Clearly the development of phase-stable VLBI would require a considerable technical advance. This could, however, tell us which parts of the source are moving if the nucleus can be resolved into two or more components, one of which is unchanging. Alternatively, for those sources for which there is a sufficiently nearby unresolved radio source (e.g., 3C 345/NRA 0512), one might obtain some absolute positional information.

3.5.4. *How do the components of expanding doubles evolve?* If the emission is due to synchrotron radiation, the component size must always be large enough to bring the brightness temperature below the plausible upper limit set by the inverse Compton restriction. It would be important to know how the source size or surface brightness distribution depends on wavelength and also whether or not a particular component faded out earlier. Some models (particularly those involving propagation effects) make specific predictions about wavelength dependence or the distribution of surface brightness. In simple screen models, the faster component would have lower brightness, whereas in models where the Doppler effect is more important, the reverse would be true. In sources showing successive outbursts, higher sensitivity observations would allow an old and fading component to be followed even when superposed on a young and bright source. This could considerably improve the interpretation of the kinematics. Some "phase-effect" models have a long recovery time and cannot readily account for repeated outbursts.

3.5.5. *What are the polarisation properties of the individual components?* If the emitting electrons are channelled along an effectively rigid magnetic field, then the observed radiation should be strongly polarised normal to the source axis it if is synchrotron radiation, and parallel if it is curvature radiation. If the field direction makes a small angle to the line of sight then perhaps a substantial circular polarisation could be observed (unless electrons and positrons are created in equal quantities). Induced Compton and plasmon scattering models also predict characteristic polarisation patterns related to the source geometry.

3.5.6. *Can we identify the time when a given outburst was initiated?* Brightness temperature limitations generally preclude the possibility of seeing the start of an outburst at radio wave-lengths. However, an unusual change in the optical or X-ray output may serve to identify the time origin (e.g., in 3C 273 which is particularly important in the case of screen models as this in principle allows a unique inversion of the geometry of the screen if the signal velocity is taken to be *c*.

3.5.7. *Is the position angle of compact double structure aligned with extended radio structure?* So far the evidence for this is mostly positive (although there are some known exceptions): this favours models in which the extended sources are being fuelled continuously by their central components. It is thus tempting to suggest that both sources share a common axis defined by the angular momentum of the central regions. This would imply that the directionality (if not complete collimation) was imprinted on a scale $\lesssim$ 1 pc. As already mentioned, accretion onto a black hole - whether disk-like or quasi-spherical - could establish a "twin exhaust" flow pattern even on a scale $\lesssim$ 100 Schwarzschild radii. 3C 273 is the only extended source known to display superliminal expansion, and it is very important to discover more examples where extended structure is also present. One can be fairly sure that the axis of a typical double sources does not make a particularly small angle with line of sight. So if rapid expansions with $v_0 >> c$ are found to be fairly common, it would rule out some models (e.g., "ballistic" models, which yield superluminous effects only when the motion is directed almost along the line of sight). Cygnus A is particularly important in this context, because there is fairly convincing evidence that the overall source axis lies almost in the plane of the sky.

REFERENCES

1. Rees, M.J., Nature 229, 312 (1971) (errata, p.510).
2. Shklovski, I.S., Sov. Astron. 6, 465 (1962).
3. Piddington, J.H., Monthly Notices Roy. Astron. Soc. 133, 163 (1966).
4. Rees, M.J., Nature Phys. Sci. 230, 55 (1971).
5. Longair, M.S., Ryle, M. and Scheuer, P.A.G., Monthly Notices Roy. Astron. Soc. 164, 243 (1973).
6. Scheuer, P.A.G., Monthly Notices Roy. Astron. Soc. 166, 513 (1974).
7. Blandford, R.D. and Rees, M.J., Monthly Notices Roy. Astron. Soc. 169, 395 (1974).
8. Turland, B.D. and Scheuer, P.A.G., Monthly Notices Roy. Astron. Soc. 176, 421 (1976).
9. Blandford, R.D. and Pringle, J.E. Monthly Notices Roy. Astron. Soc. 176, 433 (1976).
10. Rees, M.J., Comm. Astrophys. 6, 112 (1976).
11. Bell, A., Monthly Notices Roy. Astron. Soc. (in press) (1977).
12. Blandford, R.D. and Ostriker, J.P. Phys. Rev. Lett (in press).
13. Christiansen, W.A., Monthly Notices Roy. Astron. Soc. 164, 211 (1973).
14. Gull, S.F., Cambridge Ph.D. thesis, 1975.
15. Gull, S.F. and Northover, K.J., Nature 244, 80 (1973).
16. Miley, G.K., Wellington, K.J. and van der Laan, H., Astron. Astrophys. 38, 381 (1975).
17. Jaffe, W.and Perola, G.C., Astron Astrophys. 26, 423 (1973).
18. Cowie, L. and McKee, C.F., Astron Astrophys. 43, 337 (1975).
19. Pacholczyk, A.G. and Scott, J.S., Astrophys. J. 203, 313 (1976).
20. Wilson, A.S. and Vallee, J.P., Astron. Astrophys. 58, 79 (1977).
21. Gisler, G.R., Cambridge Ph.D. thesis, 1976.
22. Wiita, P., Astrophys. J. (in press) (1977).
23. Shields, G.A. and Wheeler, J.C., Astrophys. Lett. 17, 69 (1976).
24. Blandford, R.D. and Znajek, R.L., Monthly Notices Roy. Astron. Soc. 179, 433 (1977).
25. Bardeen, J.M. and Petterson, J., Astrophys. J. (Lett.) 195, L.65 (1974).
26. Burbidge, G.R., Jones, T.W. and O'Dell, S.L., Astrophys. J. 193, 43 (1974), and references cited therein.
27. Jones, T.W., O'Dell, S.L. and Stein, W.A., Astrophys. J. 192, 261 (1974).
28. Rees, M.J., Astrophys. Lett. 2, 1 (1968).
29. Kellerman, K.I. and Pauliny-Toth, I.I.K., Astrophys. J. 155, L.71 (1969).
30. O'Dell, S.L., Proc. Cambridge Quasar Conference (in press).
31. Jones, T.W. and O'Dell, S., Astron. Astrophys. 61, 291 (1977).
32. Wardle, J.F.C., Nature 269, 563 (1977).
33. Pacholczyk, A.G., Monthly Notices Roy. Astron. Soc. 163, 29p (1973).
34. Jones, T.W. and O'Dell, S., Astrophys. J. 215, 226, (1977).
35. Rees, M.J. and Sciama, D.W., Nature 216, 147 (1967).
36. Legg, M.P.C. and Westfold, K.C., Astrophys. J. 154, 499 (1968).
37. Pacholczyk, A.G. and Swihart, T.L., Monthly Notices Roy. Astron. Soc. 153, 3P (1971).
38. Kompaneets, A.S., Sov. Phys. JETP 4, 730 (1956).
39. Sunyaev, R.A., Astrophys. Lett. 7, 19 (1970).
40. Levich, E.V., Sunyaev, R.A. and Zeldovich, Y.B., Sov. Phys. JETP 35, 733 (1972).
41. Wilson, D.B. (in preparation).
42. Krolik, J. and McKee, C. (in preparation).
43. Blandford, R.D., McKee, C.F. and Rees, M.J., Nature 267, 211 (1977).
44. Sanders, R.H., Nature 248, 390 (1974).

45. Epstein, R.I. and Geller, M.J., Nature 265, 219 (1976).
46. Vitello, P. and Salvati, M., Phys. Fluids 19, 523 (1976).
47. Michel, F.C., Astrophys. Space. Sci. 15, 153 (1972).
48. Blandford, R.D. and McKee, C.F., Monthly Notices Roy. Astron. Soc. 180, 343 (1977).