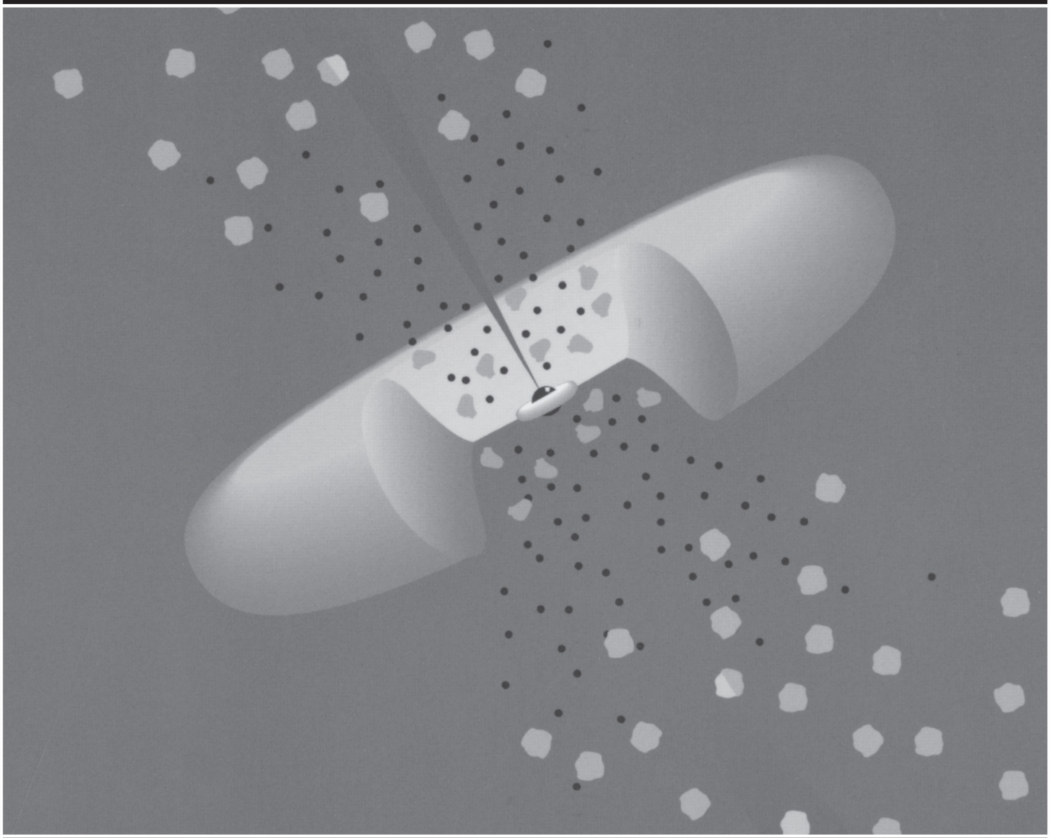
Julian H. Krolik

ACTIVE GALACTIC NUCLEI

From the Central Black Hole to the Galactic Environment



Princeton Series in Astrophysics

ACTIVE GALACTIC NUCLEI

PRINCETON SERIES IN ASTROPHYSICS

Edited by Jeremiah P. Ostriker and David N. Spergel

Theory of Rotating Stars, *by Jean-Louis Tassoul*

Theory of Stellar Pulsation, *by John P. Cox*

Galactic Dynamics, *by James Binney and Scott Tremaine*

Dynamical Evolution of Globular Clusters, *by Lyman Spitzer, Jr.*

Supernovae and Nucleosynthesis: An Investigation of the History of Matter, from the Big Bang to the Present, *by David Arnett*

Unsolved Problems in Astrophysics, *edited by John N. Bahcall and Jeremiah P. Ostriker*

Galactic Astronomy, *by James Binney and Michael R. Merrifield*

Active Galactic Nuclei: From the Central Black Hole to the Galactic Environmen *by Julian H. Krolik*

Julian H. Krolik

ACTIVE GALACTIC NUCLEI: FROM THE CENTRAL BLACK HOLE TO THE GALACTIC ENVIRONMENT

PRINCETON UNIVERSITY PRESS
Princeton, New Jersey

Published by Princeton University Press, 41 William Street,
Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press,
Chichester, West Sussex

Library of Congress Cataloging-in-Publication Data

Krolik, Julian H., 1950 –
Active galactic nuclei: from the central black hole to the galactic environment / Julian H. Krolik
p. cm.—(Princeton series in astrophysics)
Includes bibliographical references and index.
ISBN 0-691-01152-4 (cl: alk. paper)
ISBN 0-691-01151-6 (pb: alk. paper)
1. Active galactic nuclei. I. Title. II. Series
QB858.3.K76 1999
512.8—DC21 98-21857
CIP

This book has been composed in New Times Roman by TEX

The publisher would like to acknowledge Julian Krolik for providing the PostScript files from which this book was printed.

The paper used in this publication meets the minimum requirements of ANSI/NISO Z39.48-1992 (R1197) (*Permanence of Paper*)

http://pup.princeton.edu

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1
10 9 8 7 6 5 4 3 2
(Pbk.)

Contents

12. Anisotropic Appearance and Unification of Disparate AGN Varieties 424

Preface

We are both enabled and limited by the tools that we use. Because the techniques of observational astronomy change from band to band, too often we construct mental views of astronomical objects sadly limited by the unconscious donning of wavelength-sensitive intellectual blinders. This effect is particularly pernicious in the case of active galactic nuclei, whose emission is spread over a tremendous range of photon varieties. Similar criticisms can be (and often are) voiced about theoretical work—for example, because we understand atomic spectroscopy far better than plasma physics, we prefer to think about line formation regions rather than accretion dynamics. In writing this book, I have made a special effort, therefore, to be as fair to the subject as possible. That is, I have striven to treat all the different sorts of radiation on an equal footing, and I have attempted to point out where there are important questions we have been unsuccessful in answering, or have completely ignored, as well as to set out well-established results. It hardly needs be said that making the attempt is not the same thing as reaching the goal, but I hope that I have at least gone part way.

Writing any book about a lively subject of research is also an exercise in hitting a moving target. One wants to avoid ideas that may no longer be interesting a few years in the future, while at the same time one would like as much of the material as possible to be useful to readers five and ten years hence. One must, therefore, face Janus-like both past and future and make one's best guess about which old ideas will last and which new ones will thrive. My private articulation of this effort is to put in only those ideas that have at least a 50% probability of not being laughable in ten years. I await reader comments in 2008 judging to what extent I have been successful in this regard.

In line with the philosophy of the preceding paragraphs, the book is organized along what I hope is a robust train of logic. First, one cannot expect anyone to persevere without motivation. That is the job of Chapter 1. There is no subject, of course, unless one can find examples, and quantification of their properties follows immediately (Chapters 2 and 3). The most fundamental property of AGNs is their remarkable energy output; recognition of this fact leads immediately to the belief that a massive black hole lies at the center of every AGN (Chapter 4), so the student must acquire at least a modicum of knowledge about how they work (Chapter 5).

Accretion is how black holes generate power. Although it is an unrealistically simple picture, consideration of accretion in spherical symmetry leads to some powerful insights (Chapter 6). A minimally realistic view must include the effects of accretion with angular momentum; this topic leads into a discussion of how the single most important continuum component (the ultraviolet) is produced (Chapter 7). From there we move progressively outward. Chapter 8 presents what we know about X-ray and γ-ray radiation; Chapter 9 deals with radio emission and jets; and Chapter 10 tells the reader how we think the strong emission lines of AGNs are made. Somewhere outside the broad emission line region we find absorbing material that travels outward at speeds up to a few tenths of c; what little we know of this phenomenon is presented in Chapter 11. Up until this point we have pretended that all observers see the same picture, no matter what their point of view; Chapter 12 corrects this misapprehension and develops the dramatic consequences of anisotropic appearance. Chapter 13 takes up the largest scale, the relationship between an active nucleus and its host galaxy. The final substantive chapter, Chapter 14, deals with a problem that lies at the heart of the subject, but one we barely even know how to approach: what ignites AGNs, sustains them, and then extinguishes them. Finally, in Chapter 15 I give a very brief summary of where the field stands as of the date of writing.

The general level of presentation is intended to match the background possessed by a typical second-year graduate student specializing in astrophysics at an American university. That is, I assume the student has had courses in the basics of radiation transfer and elementary radiative processes, classical electromagnetism, and atomic physics. I also assume the student has a nodding acquaintance with fluid mechanics and has heard something about general relativity. In these latter areas, qualitative understanding is more important than the ability to do technical calculations. Very brief refreshers on general relativity, magneto-hydrodynamics, and fluid discontinuities (e.g., shocks) are provided in the appendices. There is also an appendix on the technicalities of estimating luminosity functions and some other, related, statistical issues, and another on the tensor virial theorem. The last appendix quickly summarizes kinematics in an expanding universe. Because the training of graduate students begins to split off into the different subfields of astronomy at about this level, I hope that the treatment will likewise be useful to professionals in astrophysics whose primary expertise is in other subject areas.

However, in recognition of the fact that there are several different communities of readers for this book, I have defined several specialized "tracks" for particular sets of readers. These are described in detail in the Guide for Readers.

Finally, no book can be complete without acknowledgments, for it is well-nigh impossible to write a book without receiving significant aid from others. Many colleagues—Marek Abramowicz, Eric Agol, Ski Antonucci, Mitch Begelman, Chris Done, Richard Green, Tim Heckman, Tim Kallman, Pawan Kumar, Chris O'Dea, Brad Peterson, Greg Shields, Mark Sincell, Mark Voit, Andrew Wilson, and Andrzej Zdziarski—did me the great service of closely reading drafts of selected chapters, helpfully pointing out—and correcting—numerous typos, oversights, and errors. Two people, Pawan Kumar and Greg Shields, used a draft of the manuscript when teaching classes and gave me valuable feedback on how it worked. Brian Boyle, Alan Bridle, Alessandro Capetti, Chris Carilli, Chris Done, Jim Dunlop, Paul Francis, Reinhard Genzel, Bob Goodrich, Bob Hartman, Anuradha Koratkar, Ari Laor, Joe Maslowski, Charlie Nelson, Paolo Padovani, Elena Pian, Paola Pietrini, Rita Sambruna, Mark Sincell, Zlatan Tsvetanov, Marie-Helene Ulrich, Meg Urry, Sylvain Veilleux, Kim Weaver, Andrzej Zdziarski, and Piotr Życki graciously produced "made-to-order" versions of figures or supplied me with proprietary data tables so that I could produce my own version of a published figure. Ron Allen insightfully (and patiently) answered my questions about the mechanics of radio observations. Jerry Kriss and the JHU FOS group gave me access to computer resources that substantially reduced the labor of figure preparation. Mark Voit provided much useful counsel about the tricky thickets of broad absorption line phenomenology and physics. John Mackenty, Brad Whitmore, and especially Tim Heckman were invaluable sounding boards and sources of carefully nuanced judgment on numerous observational questions. But most importantly, I owe a debt of gratitude to the many excellent collaborators with whom I've worked over the years, and from whom I've learned much of what I know. In rough chronological order, they have been Chris McKee, John Kwan, Tim Kallman, Mitch Begelman, Andrzej Zdziarski, Ed Pier, Mark Sincell, Chris Done, Paola Pietrini, and Tim Heckman.

Guide for Readers

While it may be an author's dream that every reader sits down and studies his book from cover to cover, it is not very likely that many readers will do so with this one. From the point of view of the graduate student, there is much more material here than can be covered in a one semester course. Practitioners already in the field, or physicists who would like to learn a little about what their astrophysicist colleagues are up to, are, for different reasons, unlikely to have the time or inclination to read it through in its entirety. To help readers such as these, I set out below a few sample tracks through the book, each catering to a specific audience.

• Track 1: A Feasible One-Semester Course

Instructors should, of course, tune their choices to the particular group of students they are teaching and the time allotted. With "rms" second-year graduate students at an American university, in 40-odd hours of lecture one might try covering:

Chap. 1 What Are Active Galactic Nuclei? And Why Does Anyone Care?

Chap. 2 How to Find AGNs

Chap. 3 Evolution

Chap. 4 Global Energetics and Black Holes, omitting §4.4.1

Chap. 5 Black Hole Physics, §5.1 only

Chap. 6 Spherical Accretion, §6.2 only

Chap. 7 Accretion Disks and the Optical/Ultraviolet Continuum

Chap. 8 X-ray and γ-ray Emission, omitting §8.6

Chap. 9 Radio Emission and Jets, omitting §§9.2.5, 9.3.1, and 9.3.2

Chap. 10 Emission Lines, §§10.1, 10.2, 10.5, and 10.7 only

Chap. 11 Intrinsic Absorption and Outflows, omitting §11.4

Chap. 12 Anisotropic Appearance and Unification of Disparate AGN Varieties

Chap. 15 Where We Stand

• Track 2: The Phenomenologist

For those people mostly interested in observations of AGN, and who wish for only a qualitative assessment of how well we understand them, the best path through the book is:

Chap. 1 What Are Active Galactic Nuclei? And Why Does Anyone Care?

Chap. 2 How to Find AGNs

Chap. 3 Evolution

Chap. 4 Global Energetics and Black Holes, omitting §4.4

Chap. 6 Spherical Accretion, §6.2 only

Chap. 7 Accretion Disks and the Optical/Ultraviolet Continuum, §§7.1 and 7.6 only

Chap. 8 X-ray and γ-ray Emission, §§8.1, 8.2.2, 8.3, 8.7, and 8.8 only

Chap. 9 Radio Emission and Jets, omitting §§9.2.5, 9.3.1, and 9.3.2

Chap. 10 Emission Lines, §§10.1, 10.2, and 10.5 only

Chap. 11 Intrinsic Absorption and Outflows, omitting §11.4

Chap. 12 Anisotropic Appearance and Unification of Disparate AGN Varieties

Chap. 13 Properties of AGN Host Galaxies

Chap. 15 Where We Stand

• Track 3: The Physicist

This track is intended for the physicist who might like to get a feel for what astrophysical problems look like. It concentrates on defining the context, and on laying bare the underlying physics issues. Applications of general relativity are found almost exclusively in Chapter 5, with a few additional points discussed in §§7.3.2 and 7.5.5.

Chap. 1 What Are Active Galactic Nuclei? And Why Does Anyone Care?

Chap. 4 Global Energetics and Black Holes

Chap. 5 Black Hole Physics

Chap. 6 Spherical Accretion, §6.2 only

Chap. 7 Accretion Disks and the Optical/Ultraviolet Continuum, §§7.1, 7.2, and 7.3 (through §7.3.3)

Chap. 8 X-ray and γ-ray Emission, §§8.4, 8.5, and 8.6 only

Chap. 9 Radio Emission and Jets, omitting §§9.1 and 9.3.4

Chap. 14 Onset and Fueling

Chap. 15 Where We Stand

Webpage

As in every field of scientific research, we continue to learn new things about active galactic nuclei all the time. Moreover, some material important to the subject (e.g., images and simulations) cannot be adequately reproduced in book form. And, of course, no book is ever perfect. Internet webpages help solve the problems created by all three of these facts. To that end, the author maintains a page with a variety of information supplementing this book at `www.pha.jhu.edu/~jhk/homepage.html`.

ACTIVE GALACTIC NUCLEI

1 What Are Active Galactic Nuclei? And Why Does Anyone Care?

1.1 What Makes Them Interesting?

Active galactic nuclei (hereafter abbreviated "AGNs") are among the most spectacular objects in the sky. They produce prodigious luminosities (in some cases apparently as much as 10^4 times the luminosity of a typical galaxy) in tiny volumes (probably $\ll 1$ pc^3). This radiation can emerge over an extraordinarily broad range of frequencies: in at least one case the luminosity per logarithmic frequency interval, that is, the luminosity per "band," is roughly constant (to within factors of several) across thirteen orders of magnitude in frequency! Their line spectra are almost as remarkable as their continua: in the optical and UV, they often display emission (and occasionally absorption) lines whose total flux is several percent to tens of percent of the continuum flux, and whose widths suggest velocities ranging up to $\sim 10^4$ km s^{-1}.

Although in most cases we cannot as yet make images in which AGNs are resolved, in certain objects it is possible to do so at radio frequencies. In those cases, one often sees variable structure with apparent speeds in the sky plane of $\sim 10c$!

Active galaxies are also noteworthy in displaying very strong cosmological evolution. The most luminous active galaxies were a thousand times more numerous at redshift 2.5 than they are today. Because their high luminosities and distinctive spectra make them relatively easy to pick out, they are disproportionately represented in our tally of known high redshift sources. The fact that the luminosity of AGNs is such a strong function of redshift suggests strongly that there is something special about youthful galaxies that promotes the creation of active nuclei.

Another interesting connection between AGNs and the evolution of the Universe comes from their use as light sources for studying intervening gas clouds, galaxies, and clusters of galaxies. Our knowledge of the intergalactic medium comes almost exclusively from the study of absorption lines in the spectra of distant quasars. On the other hand, systems that are transparent, but possess moderately deep gravitational wells, like galaxies and clusters of galaxies, can form gravitational lenses. The distorted—and sometimes multiple—images of distant AGNs formed by these lenses can be used to infer the nature of the gravitational potential—and therefore the mass—of the lenses.

To the degree that we do understand AGNs physically (a rather controversial point), their basic nature involves events of considerable physical drama. Most workers in the field now believe that the power for AGNs comes from accretion onto massive black holes. If this is true, their most basic properties depend on some of the most exotic physics we know: strong-field relativistic gravity.

This whole stew—exotic physics, photons detectable with virtually every sort of astronomical instrument, and a deep connection to cosmology—has made AGNs the focus of a significant fraction of the world astronomical community's attention for the past 30 years.

1.2 What Exactly Are We Talking About? The Most Salient Properties of AGNs

In order to start any discussion of AGNs, we must first define what we mean by the term *active galactic nuclei.* Unfortunately, although the black hole model has achieved widespread acceptance, it is not yet completely confirmed; moreover, and this is part of why its confirmation remains incomplete, direct signatures of accreting massive black holes are much harder to see than a variety of more indirect signals. Consequently, the only clear way to define AGN is operationally; all we can do is to list the observable phenomena we use to find them. This procedure is, admittedly, circular, but such a "bootstrap" approach is all that is possible until we achieve a more fundamental understanding.

It turns out that AGNs can be found in many ways, but not all AGNs have every single property. Thus, there is no single defining list of qualities to look for. A better way to think about the situation is to imagine a "menu" of phenomena from which they choose, with some items more popular than others. Note that in evaluating "popularity," one's terms of definition must be carefully chosen; in samples selected on the basis of a certain property, that property will always be very popular, whereas in samples selected otherwise it may be only a specialized taste.

Table 1.1 illustrates this point by listing some of the salient observational signatures of AGN, along with brief comments on how often these properties are found, and a few concise caveats. Each of these features is hardly ever seen in normal galaxies. In the following subsections we will discuss these signatures at somewhat greater length.

1.2.1 Very small angular size

The first property is certainly the most visually striking. When the AGN is near enough that a host galaxy of reasonable surface brightness can be seen, in optical images the nucleus often appears to be a bright point whose flux can often rival, or even exceed, the flux from all the rest of the host galaxy (e.g., NGC 1566, a nearby AGN, is shown in Plate 1).

However, this simple picture is a bit misleading. Our ability to see both a bright point and its surrounding host is a function of the luminosity contrast between the nucleus and its host galaxy, a quantity that varies both from case to case and as a function of wavelength. If the luminosity of the nucleus is too small relative to the host, it will not stand out, of course. Conversely, when the luminosity of the nucleus is much greater than that of its host, light from the nucleus can overwhelm any light from the host.

This latter effect is exacerbated by the kinematics of our expanding Universe. For a start, at $z \simeq 1$ the angular size of an average galaxy is only $\sim 1''$, comparable to the seeing disk even at a good ground-based observatory. The problems don't end there. The typical luminosity of the AGNs we can find at $z \simeq 2$ is ~ 100 times greater than at the present epoch, so the luminosity contrast between the nucleus and its host is very large. In addition, although the nucleus remains effectively a point, the bolometric surface brightness of the host falls as $(1+z)^{-4}$ (see Appendix F). Finally, if the AGN has a redshift greater than about unity, the light we observe in the visible band was ultraviolet in the rest frame, and most galaxies are comparatively dim in the UV. All these effects combined make hosts of high-redshift AGNs very difficult to see (§13.1.1).

The picture we see also depends strongly on wavelength. As we will discuss at slightly greater length in the next section, many AGNs have a much greater ratio of X-ray luminosity to optical than does any normal galaxy. For this reason, their X-ray images are essentially pure points. On the other hand, radio emission, more often than not, extends over a sizable region, frequently much larger than a galaxy (§§ 1.2.7, 9.1.2).

In the long run we can hope to obtain images of even the smallest AGN structures, but the requirements are daunting. As we shall see later, there are AGN pieces spread over a very wide dynamic range of radial scale.

Table 1.1: The Menu

Property	Popularity	Comments and Exceptions
Very small angular size	Many	Wavelength-dependent
Galactic (or greater) luminosity	Many	Lower luminosity is hard to find; obscuration and beaming may mislead
Broad-band continuum	Most	Often $dL/d\log\nu \simeq const.$ from IR to X-rays; sometimes to γ-rays
Strong emission lines	Most	Sometimes very broad, sometimes not
Variable	Most	Modest amplitude; short wavelengths stronger, faster than long
Weakly polarized	Most	$\sim 1\%$ linear; a minority much stronger
Radio emission	Minority	Sometimes, but not always, extended on enormous scales
Strongly variable and polarized	Small minority	Correlated with bright radio and high-energy γ-rays; in some cases emission lines absent

Equivalent elements in different AGNs tend to have roughly the same effective temperature (i.e., luminosity emitted per unit area $F = \sigma T_{\rm eff}^4$, where σ is the Stefan-Boltzmann constant) because the effective temperature roughly characterizes the nature of the structural element. The effective temperature is also a rough guide to the typical wavelength emitted by that part of the AGN ($\lambda \sim ch/kT_{\rm eff}$) if the radiation mechanism is roughly thermal, but nonthermal mechanisms can easily make the typical wavelength far shorter.

Even at the heart of the beast, $T_{\rm eff}$ is rarely much more than $\sim 10^5$ K (§§7.1, 7.3.3); discounting radio emission regions, at the outer edge of an AGN proper, it may fall to a few hundred K (§12.4.4). We can then roughly predict that in an AGN whose observed flux is $F_{\rm obs}$, the angular size of the region whose effective temperature is $T_{\rm eff}$ is

$$\theta = 87(1+z)^2 \left(\frac{F_{\rm obs}}{10^{-11}\ {\rm erg\ cm^{-2}\ s^{-1}}}\right)^{1/2} \left(\frac{T_{\rm eff}}{10^3\ {\rm K}}\right)^{-2} \quad \text{microarcsecond,} \tag{1.1}$$

where the fiducial value of $F_{\rm obs}$ has been chosen at a level corresponding to roughly the tenth brightest AGN in the sky. Thus, the angular size of any particular AGN structure scales $\propto F_{\rm obs}^{1/2}$, so that the apparently brightest AGNs are also the largest (in angular terms)—and even they are very small indeed.

1.2.2 High luminosity

We know of AGNs with luminosities all the way from $\sim 10^{42}$ to $\sim 10^{48}$ erg s^{-1}. To put this in perspective, the characteristic luminosity of the field galaxy distribution (L_*) is $\sim 10^{44}$ erg s^{-1}. In other words, we see AGNs whose power output ranges from as little as 1% of a typical galaxy to $\sim 10^4$ times as great. However, one must be careful in interpreting these luminosities. On the one hand, we cannot easily detect active nuclei much weaker than the host galaxy, and so there may be a large population of "mini-AGNs" that are as yet unknown. On the other hand, there is also reason to think (see Chap. 12) that in many AGNs the active nucleus is obscured by extremely thick dust extinction, so that we can be grossly misled as to its true luminosity if our only measure is the power output in optical or ultraviolet light. Relativistic beaming (§§9.3.4, 12.2.1) can also substantially distort the angular distribution of light from AGNs, and, of course, there is a strong selection effect in favor of observing those whose radiating material is moving toward us. Thus, in objects where beaming is likely to be significant, it is important to distinguish between the luminosity inferred assuming isotropic radiation and the true luminosity. Obscuration, or beaming directed away from us, can, of course, so weaken the light we see from an AGN that we may not even recognize it.

1.2.3 Broad-band continuum emission

It is best to begin this discussion with a digression about an issue of notation. In most fields of astronomy, the property of light that is measured is the specific flux (F_ν or F_λ), the rate at which energy arrives per unit area per unit frequency ν or per unit wavelength λ. The standard unit for F_ν is the Jansky, 10^{-23} erg cm^{-2} s^{-1} Hz^{-1}. In high-energy (i.e., X-ray and γ-ray) astronomy, where photon-counting devices prevail, the customary measured quantity is N_ϵ $(= F_\nu/(h\epsilon))$, the rate at which photons arrive per unit area per unit energy ϵ. However, when we speak of a "band" of the electromagnetic spectrum, whether it is radio, infrared, visible, or X-ray, we generally mean a span in the logarithm of the wavelength. The term *infrared*, for example, generally refers to a range of $\sim 10^2$ in wavelength, from ~ 1 to ~ 100 μ. Therefore, for describing which band is most important in terms of energetics, the most convenient quantity is $\nu F_\nu = dF/d\log\nu = dF/d\log\lambda = \lambda F_\lambda$, the energy flux per logarithmic bandwidth. In most of this book (starting with fig. 1.1), this will be the favored form for the presentation of spectra.

To understand what is meant by "broad-band" continuum radiation, one should first contrast the spectra of ordinary galaxies (see fig. 1.1). To a first approximation, galaxies are piles of stars. A good zeroth-order approximation to a stellar spectrum is that it is a blackbody, so the great majority of a star's luminosity comes out within a factor of three in frequency. The total span of stellar surface temperatures is only about a factor of ten, and in any particular galaxy, the stellar mix is usually such that a limited temperature range dominates the total power output. Thus, a typical galaxy emits nearly all its power within no more than one decade of frequency, and usually rather less. The only possible modification to this picture is due to interstellar dust. In many spiral galaxies the dust extinction is great enough that a significant fraction of the optical and ultraviolet light is absorbed by cool dust grains and reradiated in the far-infrared. Because there is a wide range in extinctions, the relative size of this secondary peak in the infrared varies substantially from one spiral galaxy to the next. In ellipticals, on the other hand (as shown in fig. 1.1), there is generally little dust, and hence at most weak infrared emission.

Most (but not all) AGN continuum spectra look spectacularly different from normal galaxy spectra. A particularly well observed example,

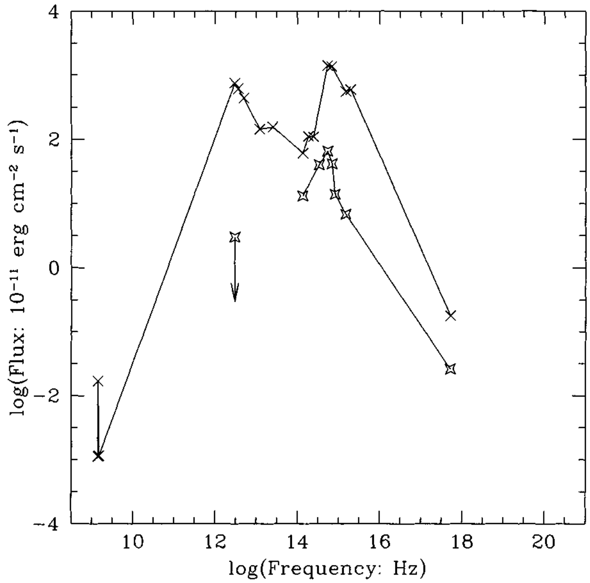


Fig. 1.1 The broad-band spectra of two typical galaxies, M 101 (Hubble type Sc: plot symbol $\times$) and NGC 4168 (Hubble type E2: plot symbol a four-pointed star). The star with an arrow is an upper limit. Connecting lines are meant only to guide the eye. In both cases, the intrinsic radiation is confined to a narrow range of optical frequencies, with only minor amounts of power emerging in the radio or X-ray bands. In the case of M 101, however, almost half the intrinsic power is absorbed by dust and reemitted at $\sim 100\ \mu$. The spike at 1.4 GHz in the spectrum of M 101 is the 21 cm HI line. At all frequencies but the near-infrared, the fluxes are integrated over the entire galaxy from maps; in that band the apertures used cover most, but not quite all, of the galaxy and therefore give fluxes that are slightly too low. Data sources are: White and Becker 1992, Rice et al. 1988, Fabbiano et al. 1992, and NED.

NGC 4151, is shown in figure 1.2. In terms of νF_ν, NGC 4151, like most AGNs, has a spectrum that is flat (to within factors of several) all the way from the mid-infrared to the hardest X-rays observed (the highest observed energy is anywhere from a few to a few hundred keV, depending on the object). Compared to normal galaxies, the fraction of the bolometric luminosity radiated in the radio band is generally an order of magnitude greater, but in some cases it is several orders of magnitude larger still; the fraction of the power that emerges in X-rays is three to four orders of magnitude larger in AGNs than in normal galaxies. NGC 4151 is so well observed because it is very nearby ($z = 0.003$), but its luminosity is relatively low,

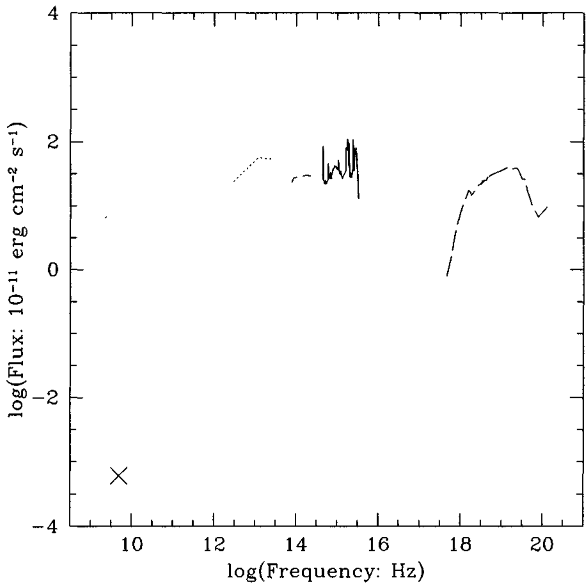


Fig. 1.2 The broad-band continuum spectrum of one of the nearest AGNs, NGC 4151, compiled from nonsimultaneous (!) data. Its flux per logarithmic bandwidth has comparable peaks in the infrared, the optical/ultraviolet, and the hard X-ray bands. Far less luminosity is radiated in radio frequencies. The dip in the soft X-ray region is due to intervening absorption, possibly in the AGN itself; the bumps in the optical/ultraviolet are smoothed versions of the strong emission lines seen in this object. The data are taken from: Ulvestad, Wilson, and Sramek (1981), the ×; Edelson and Malkan (1986), the dotted line; Kriss et al. (1995), the broken line in the optical band; the *HST* archive, the solid line; and Zdziarski, private communication, the dashed line in the X-ray band. In all cases but the four infrared points at 12 μ, 25 μ, 60 μ, and 100 μ, the apertures used were small enough that the AGN dominates the flux.

merely comparable to its host galaxy. Composite spectra illustrating what AGNs look like at high redshift and high luminosity are shown in figure 1.3.

All told, the range over which νF_ν is roughly flat is more than a factor of 10^5 in frequency, producing a spectrum far broader than any normal galaxy's. Although there are weak local maxima, it is clearly inappropriate to speak of any one frequency band dominating the output. These local maxima may, however, be signaling to us that the primary emission mechanisms change as functions of frequency (Chaps. 7, 8, 9, 12). In the case of one particular subclass of AGN, the flux in photons as hard as 1 GeV is at least as great as that in lower frequency bands, and there are a few

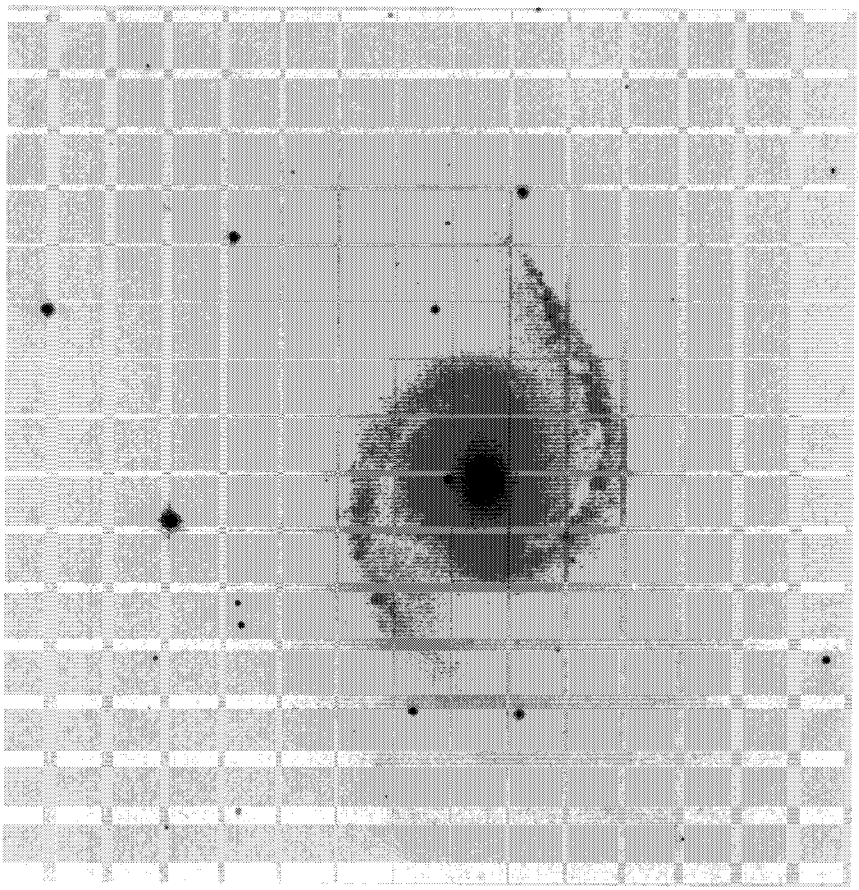


Plate 1 *NGC 1566*, a nearby active galaxy. It appears to be a fairly normal spiral galaxy, but for the extremely bright nucleus (image courtesy of Z. Tsvetanov).

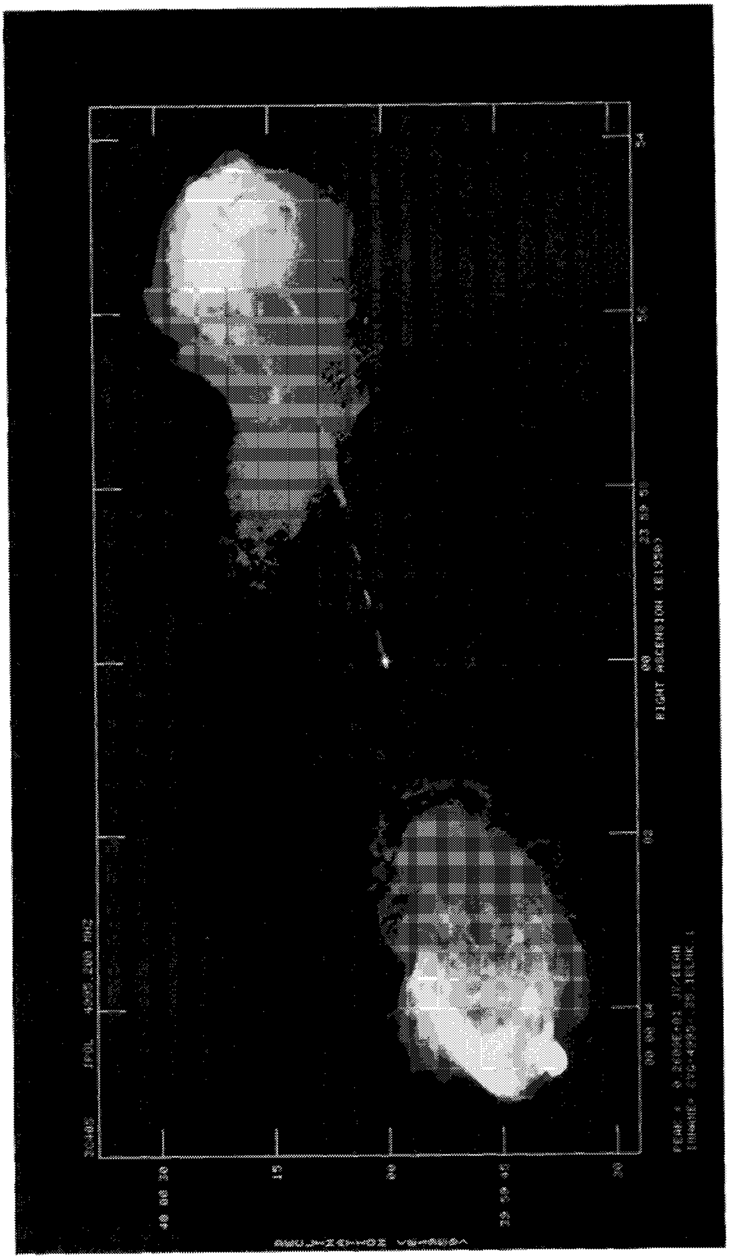


Plate 2 The original extragalactic radio source, Cygnus A, at 6 cm (image courtesy of R. Perley). The total linear extent of this source is $\simeq$ 120 kpc.

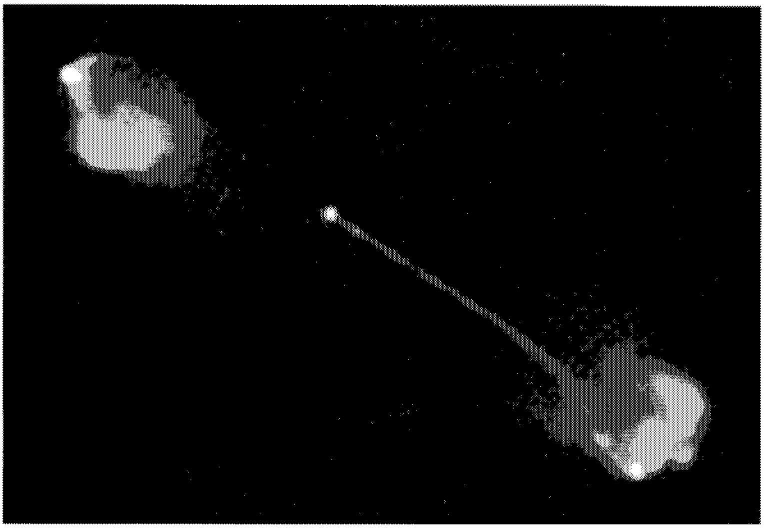


Plate 3 The FR2 radio galaxy 3C 175 at 4.9 GHz, with North to the right and East up (courtesy A. Bridle; originally published in Bridle et al. 1994). The surface brightness peaks in the outer rims of the lobes, especially in the bright hot spots. The jet, while visible, is relatively faint and can be seen on only one side of the nucleus. Even at this comparatively high frequency, the core is much dimmer than the lobes.

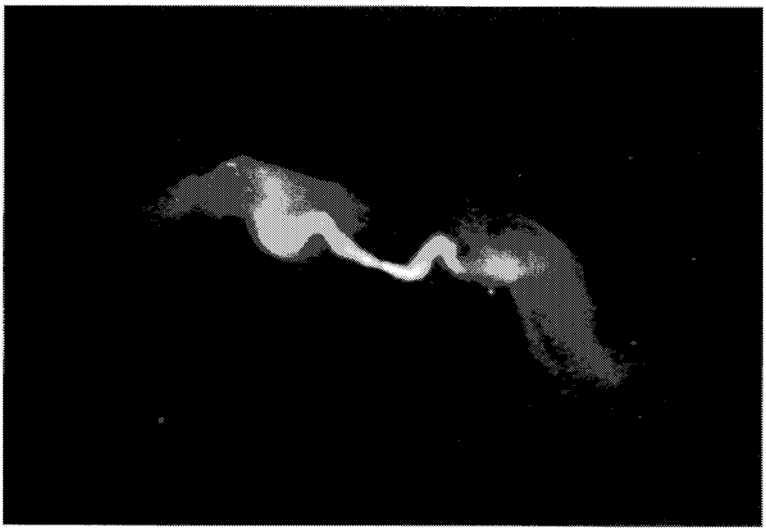


Plate 4 The FR1 radio galaxy 3C 31 at 1.4 GHz, oriented with North to the right and East to the top (courtesy A. Bridle). From one end of the radio emission to the other is $\simeq 400(h/0.75)^{-1}$ kpc. Note how the jet is clearly stronger on one side of the nucleus, and how the surface brightness declines with increasing distance from the nucleus on both sides. In addition, like many, but not all, FR1s, the jets make sharp bends after traveling out some distance from the nucleus.

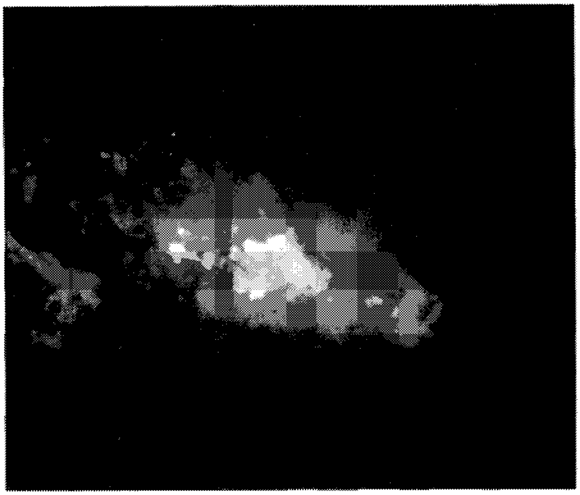


Plate 5 A 4″ × 6″ gray-scale *HST* image of [O III] 5007 emission in NGC 1068 (Macchetto et al. 1994). The nucleus is near the base of the plume, which is $\simeq 200(h/0.75)^{-1}$ pc long. Note that the vertical axis is rotated 50° from North; the [O III] plume is actually aligned fairly well with the total intensity plume shown in Plate 6.

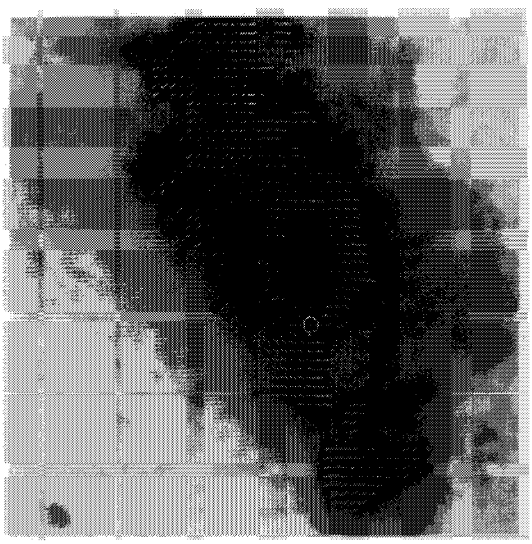


Plate 6 The innermost 3.3″ × 2.9″ of the polarizing mirror in NGC 1068 (Capetti et al. 1995). Unlike Plate 5, here North is up. Total intensity is shown in gray scale; the white lines show the local direction of the polarization $\vec{E}$-vector. Because the local polarization must always be exactly perpendicular to the direction to the source, the intersection of the $\vec{E}$-vector normals gives a very good estimate of the position of the nucleus; this is the white circle.

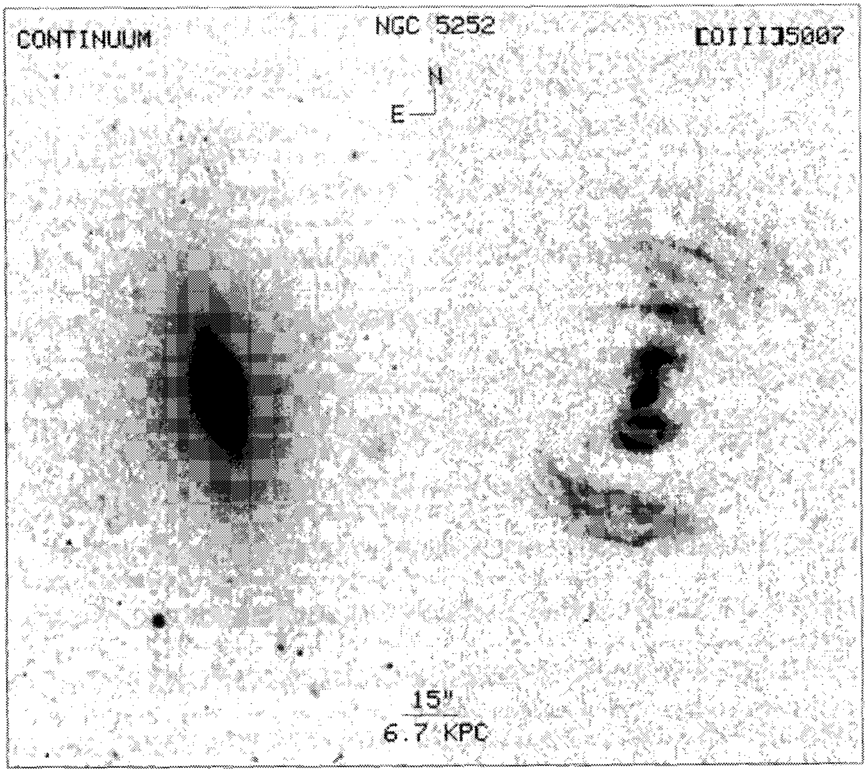


Plate 7 The "ionization cone" in the type 2 Seyfert galaxy NGC 5252 (courtesy Z. Tsvetanov). The left-hand panel shows a continuum image of the galaxy; the right-hand, an image in a filter centered on red-shifted [O III] 5007. The line emission is entirely within a sharp-edged double cone canted at an oblique angle with respect to the galactic axis.

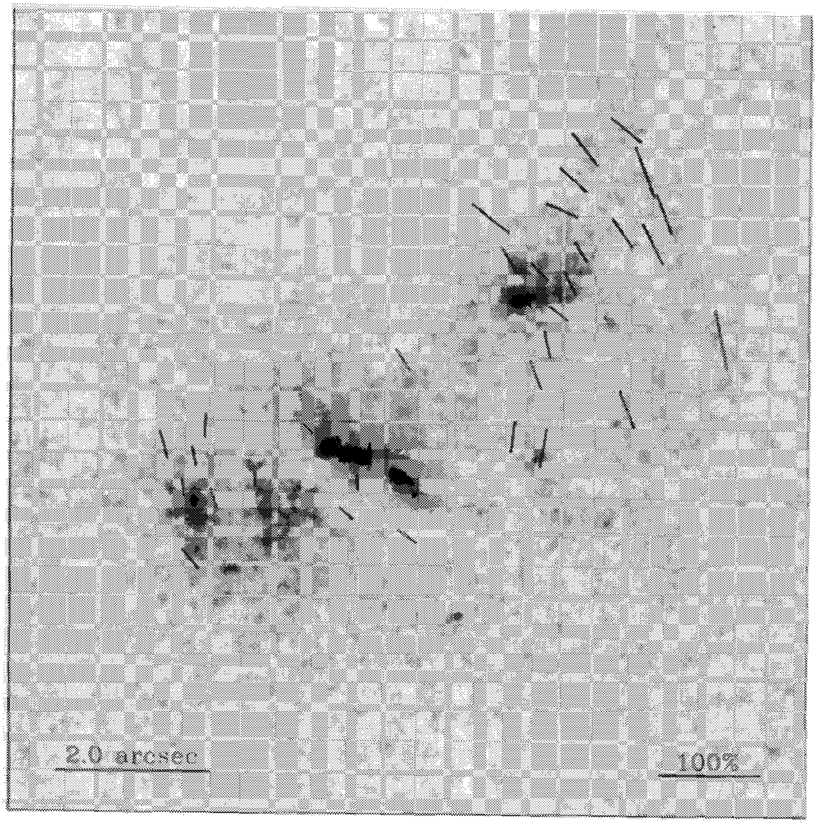


Plate 8 Near-UV (3200 Å) image of the radio galaxy 3C 321 (from Hurt et al. 1997; see also an R-band image in Plate 10). The short straight lines show the directions of the local polarization $\vec{E}$-vector; their lengths are proportional to the local polarization fraction, as shown by the scale bar. They clearly indicate that the polarization is due to scattering of light emanating from a small source, probably behind a dust lane (which is more clearly visible in Plate 10).

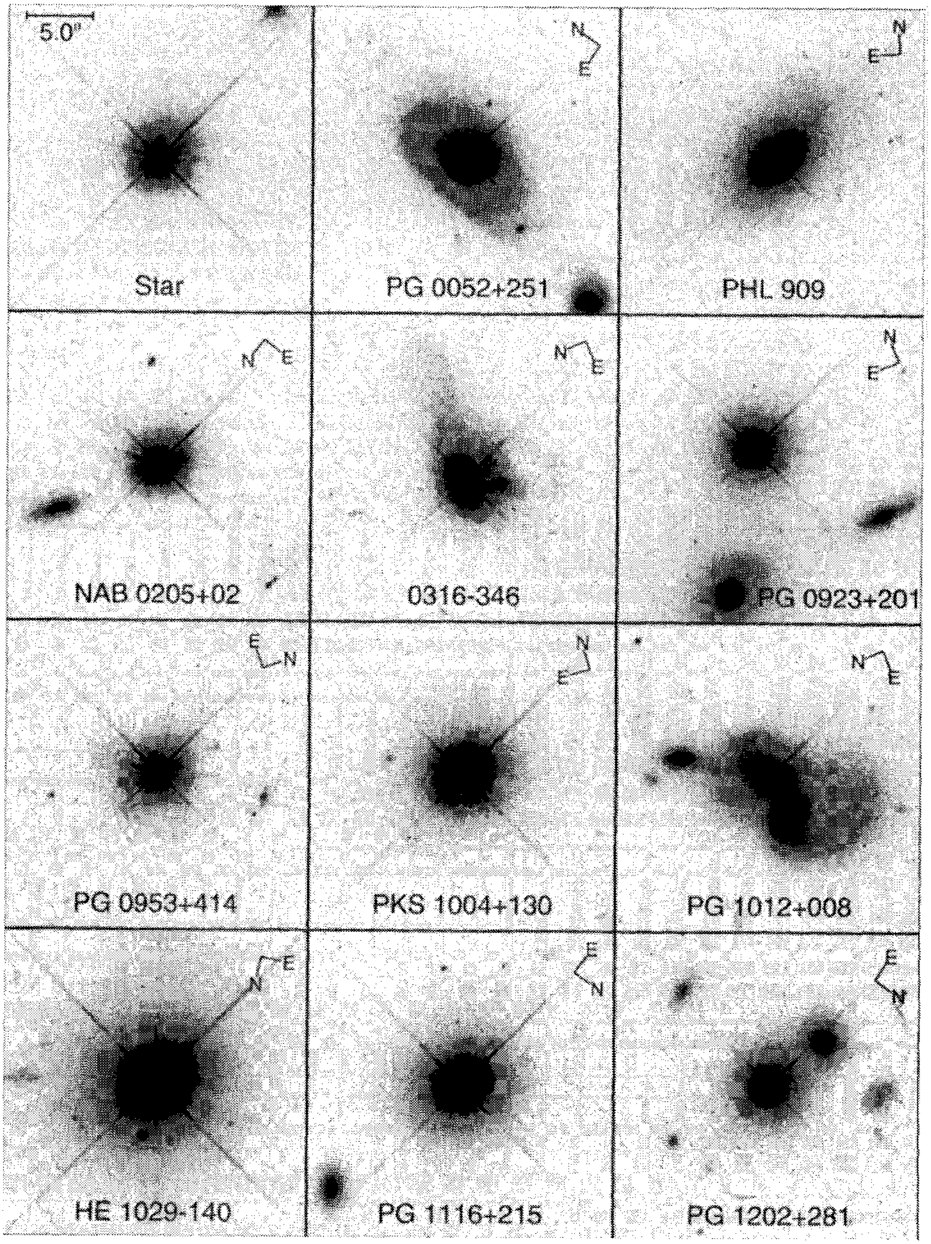


Plate 9 Eleven low-redshift ($z < 0.3$) quasars and one star (for comparison) as seen in red light by the *HST* (Bahcall et al. 1997). Clear spiral structure can be seen in some host galaxies (such as the host to PG 0052+251), but others appear to be ellipticals (e.g., PHL 909). In one case (PG 1012+008), the host is clearly undergoing a violent collision with another galaxy.

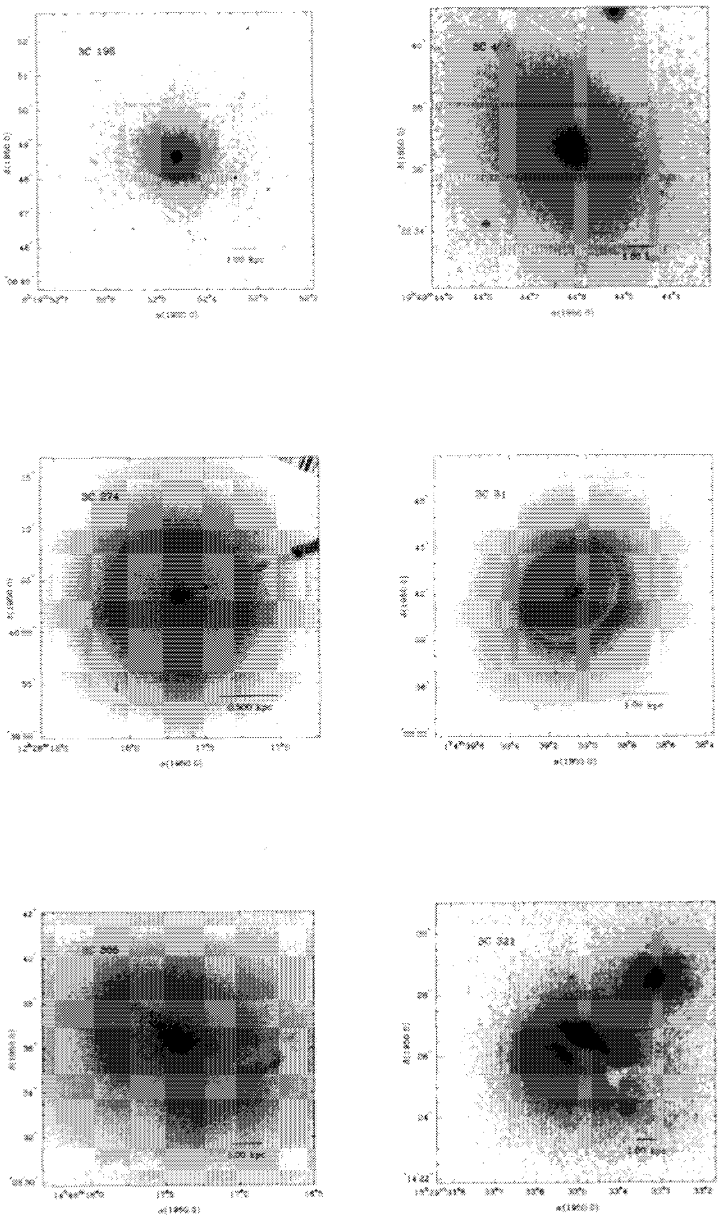


Plate 10 A sampler of R-band images of the hosts of low-redshift 3C radio galaxies, taken from Martel et al. (1997). From upper-left to lower-right, they are 3C 198 ($z = 0.082$), 3C 403 ($z = 0.059$), 3C 274 ($z = 0.004$), 3C 31 ($z = 0.017$), 3C 305 ($z = 0.041$), and 3C 321 ($z = 0.096$). The distance scale given for each is computed assuming $H_o = 75$ km s^{-1} Mpc^{-1}. Most of these are obviously elliptical galaxies, but several just as unmistakably show unusual features. 3C 274 is also known as M 87; its jet (apparent all the way from radio frequencies through X-rays) is clearly visible. 3C 305 has been disturbed in some fashion. 3C 321, for which a near-UV and polarimetric image is shown in Plate 8, has a dark dust lane across its center, and a nearby companion.

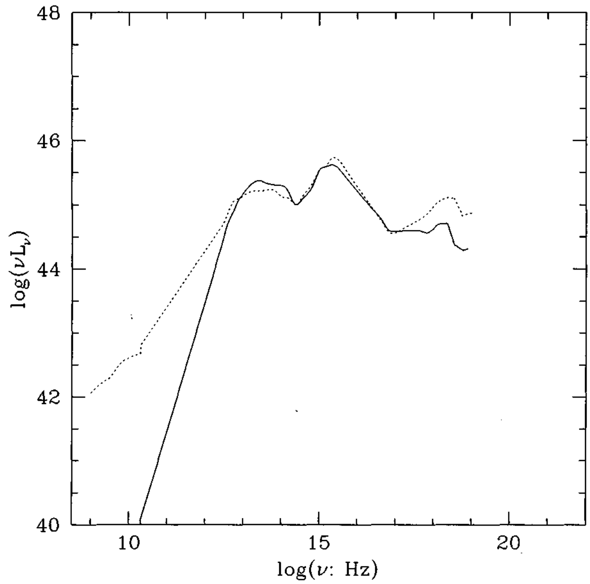


Fig. 1.3 Composite spectra for two different samples of high-redshift, high-luminosity AGNs, one (solid curve) relatively weak in the radio band, and the other relatively strong (dotted curve). The straight line segments in the far-IR/sub-mm band and the EUV are both interpolations across spectral regions where no real observations exist. Although the curves end at photon energies of a few tens of keV, this is not because AGNs are weak there (see, e.g., figs. 1.2 and 1.4).; it is because good data in the hard X-ray band and beyond exist for comparatively few objects. As these curves are composites, the luminosity scale (given in units of erg s^{-1}) should not be taken literally. Data are from Elvis et al. (1994).

examples of this subclass (e.g., Markarian 421; see fig. 1.4) with as much flux at 1 TeV as anywhere else in the electromagnetic spectrum. However, as always, there are exceptions. There are also AGNs in which the bulk of the light arrives at Earth within a frequency span of only one decade in the infrared (NGC 1068, another very nearby AGN, is a good example of this variety; see fig. 1.5).

1.2.4 Emission lines

AGN emission lines have received a great deal of attention for two reasons. First, they are often very prominent (equivalent widths are often ~ 100 Å). This makes AGN spectra stand in great contrast to the spectra of most stars and galaxies, where lines are generally relatively weak and

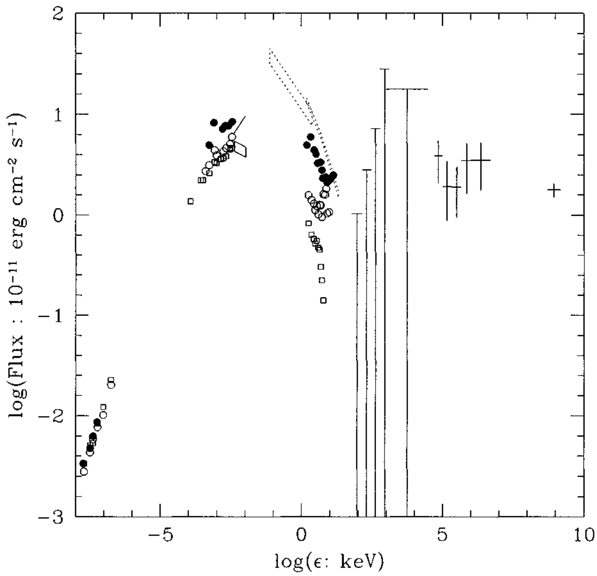


Fig. 1.4 Markarian 421 is a BL Lac object (see table 1.2). Its luminosity is spread almost equally per logarithmic frequency interval over 13 decades in frequency! Where different levels are shown at the same energy, the contrast denotes variability. The data for this plot were collected by Zdziarski and Krolik (1993).

predominantly in absorption. Figure 1.6 presents a composite spectrum that dramatically illustrates just how prominent the emission lines are. Second, because we know a great deal about atomic physics, it is easy and productive to study them (Chap. 10).

The emission lines that we see are remarkably stereotyped from one object to the next. When we have the wavelength coverage to look for them, if there are any lines at all, we almost always see Lyα, the Balmer lines, the CIV 1549 doublet, [OIII] 5007, and several others that are generally weaker. The Fe Kα X-ray line near 6.4 keV is also frequently seen.

However, there is an interesting split in the line width distribution. In some objects, many of the lines have broad wings extending out several thousand km s^{-1} from line center, whereas in others the lines are never broader than a few hundred km s^{-1}. Interestingly, the permitted and semi-forbidden lines are seen in both of these classes; in fact, when the broad wings appear, there is often a narrow core as well. The forbidden lines, on the other hand, are only seen with narrow profiles. In another

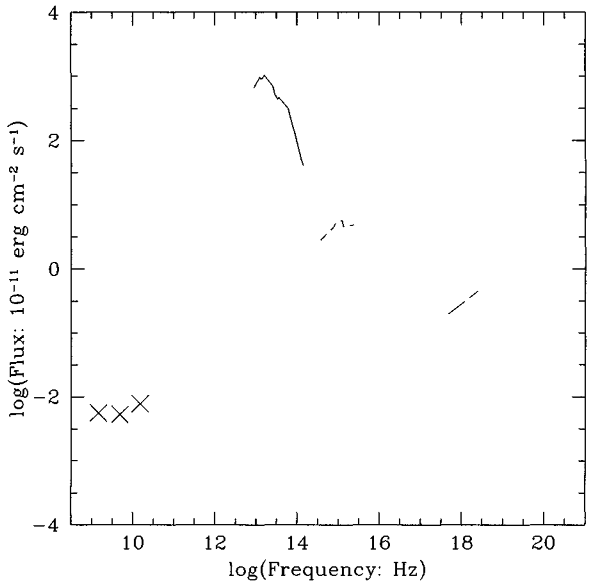


Fig. 1.5 NGC 1068 is another very nearby AGN. Unlike NGC 4151, its spectrum has a strong peak in the infrared, most likely due to very optically thick dust obscuration that reradiates the nucleus's luminosity in that band (§12.4.4). The data are from: Wilson and Ulvestad (1982), the ×'s; Rieke and Low (1975), the solid line; and Pier et al. (1994), the dashed lines. They have been analyzed in such a way as to show only nuclear radiation.

interesting correlation, those objects with only narrow lines are often quite weak from the near-infrared through the X-ray band; most of their light is generally emitted in the mid-infrared (as in the example shown in fig. 1.5).

1.2.5 Variability

It is often loosely remarked that variability is a hallmark of AGNs. This is only partially true. In the optical band, most AGNs, unlike normal galaxies, can be seen to vary, but the typical amplitude over human timescales (e.g., a few years) is often only 10% or so. Incomplete evidence suggests that the variability amplitude on the most easily observed timescales increases toward shorter wavelengths, with factors of two often seen in the X-rays. Figure 1.7 shows two cases in point, once again the frequently observed AGN NGC 4151, and the almost equally popular NGC 5548.

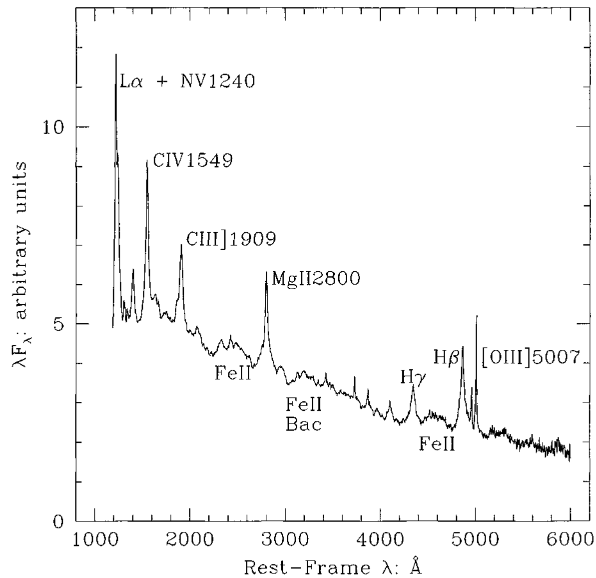


Fig. 1.6 A composite optical/ultraviolet AGN spectrum as compiled by Francis et al. (1991) from a large quasar survey. The very large contrast between emission lines and continuum, and the lines' substantial breadth, are both immediately apparent. Some of the more prominent emission lines are labeled; "FeII" is a shorthand for blends of many FeII multiplets, and "Bac" refers to the Balmer continuum.

It cannot be emphasized too strongly (fig. 1.7 underlines this point) that any statement about variability is strongly dependent on the timescale in question. A well known theorem in Fourier analysis states that the variance of a function of time $F(t)$ sampled at intervals Δt over a duration $t_{\rm tot}$ is equal to the integral of the power spectrum over the range of frequencies to which such measurements are sensitive:

$$\mathrm{Var}(F) = \int_0^{t_{\rm tot}} dt\, \left[F(t) - \langle F\rangle\right]^2 = \int_{1/t_{\rm tot}}^{1/(2\Delta t)} df\, |\hat{F}(f)|^2. \qquad (1.2)$$

The form of this relation makes it very clear that the measured variance depends on the sampling unless it is good enough that essentially all the variability power is contained within the observed frequency range. Unlike stars, whose variability is often dominated by periodic components (consider eclipsing binaries or Cepheid variables), AGNs for the most part vary with no special timescales; that is, their Fourier spectra are broad-band, just as their photon spectra are. In consequence, the amplitude of variability for AGNs is a slippery thing to measure.

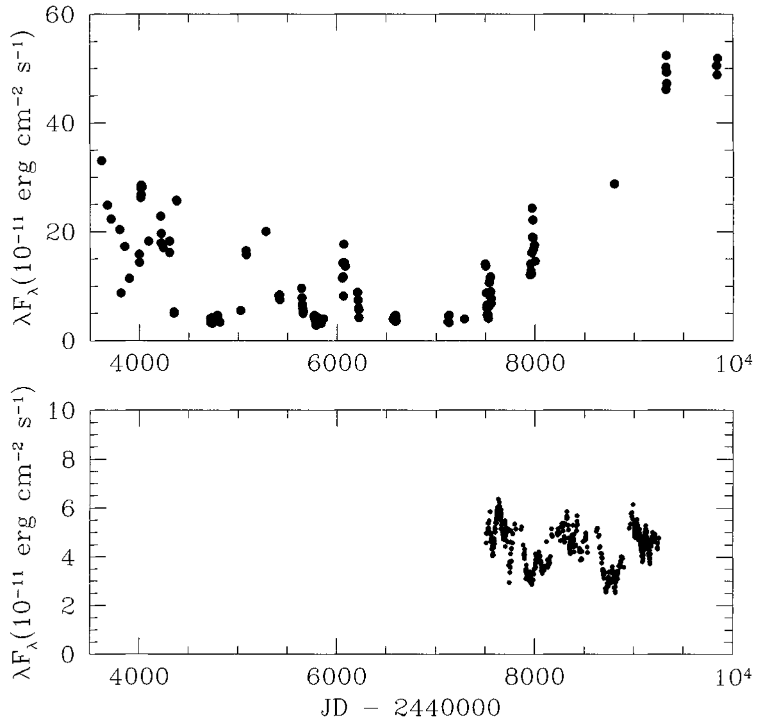


Fig. 1.7 (Top) A long-term UV (1455 Å) lightcurve for NGC 4151 (from Ulrich et al. 1997); (bottom) a somewhat shorter optical (5100 Å) lightcurve for NGC 5548 (data from Peterson et al. 1991, 1992, and 1994). In the UV, fluctuations of factors of several are common and can occur on timescales ranging from weeks to years. In the optical band, the fluctuations tend to be rather smaller.

A small subset of AGNs vary much more strongly, even in the optical band. In some cases, fluctuations of a factor of two have been seen from night to night, and cumulative changes of factors of 100 have occurred over year timescales (see, e.g., fig. 1.8). Strong variability is also very strongly correlated with three other properties: strong polarization, compact radio structure, and strong high-energy γ-ray emission.

1.2.6 Polarization

Most stars are intrinsically unpolarized, but the light we receive is generally linearly polarized by $\sim 0.5\%$ due to interstellar dust transmission polarization. The same is true for most galaxies. Most AGNs are also weakly polarized, but just enough more strongly for their polarization distribution to be statistically distinguishable from that of stars: typically they are lin-

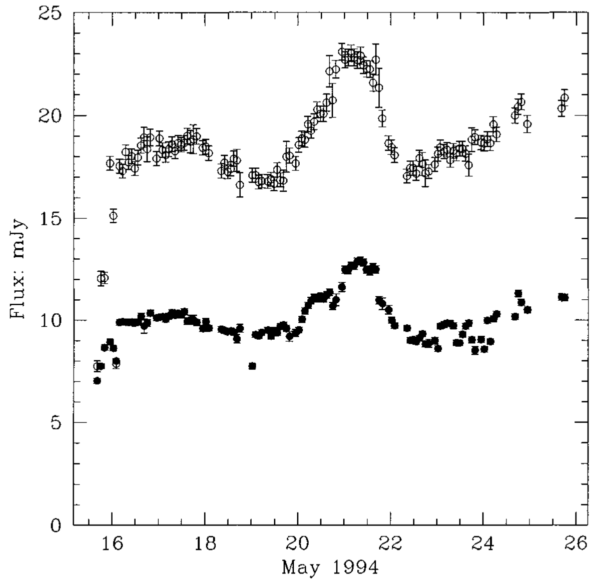


Fig. 1.8 Two ultraviolet lightcurves for PKS 2155-304. The open circles represent continuum flux at 2800 Å; the filled circles show the continuum flux at 1400 Å. In this case, changes of several tens of percent occurred within 1 day. Data courtesy of E. Pian, originally published in Pian et al. (1997).

early polarized, with fractional polarization $\simeq$ 0.5–2%. However, again as for variability, a minority—including the same minority that have strong optical variability, as well as some of the AGNs with only narrow emission lines—are much more strongly polarized, often $\sim$ 10% in linear polarization. Those objects that are strongly polarized and strongly variable in total flux are also highly variable in both the magnitude and the direction of the polarization. This variability has its bounds, however; circular polarization has never been detected, and the limits are reasonably tight.

Technical caveats are in order for discussing polarization as well. When the luminosity of the nucleus does not completely dominate the luminosity of the host galaxy, the polarization we measure depends on the aperture used, for starlight can substantially dilute the polarization of the nucleus. This effect is most noticeable in the AGNs with only narrow emission lines. In those cases, the (linear) polarization of the true nucleus can be as high as tens of percent but can only be seen after virtually all starlight has been eliminated (Chap. 12).

In addition, whether or not we detect polarization depends on the wavelength observed. Radio emission is usually linearly polarized at the few to few tens of percent level, but the observed value can be artificially suppressed by insufficient angular resolution, for the angle of polarization often varies from place to place (§9.1.4). Some AGNs whose spectra show exceedingly broad ultraviolet absorption troughs (see Chap. 11) are strongly polarized in the absorption features and more weakly polarized in the continuum. Others become strongly linearly polarized at wavelengths just shortward of the Lyman edge. Whether the X-ray emission is polarized is a question that awaits more sensitive instruments than those available in the 1990s.

1.2.7 Radio emission

Strong radio emission is the last of the distinguishing marks of AGNs. Historically, however, it was effectively the first. Some of the earliest radio astronomical observations discovered that many bright radio sources come in the form of double lobes with a galaxy located halfway between them (e.g., Cygnus A, shown in Plate 2). These radio lobes were the first evidence recognized as indicating nonstellar activity in external galaxies. After the compilation of the 3C catalog (see Chap. 2) in the late 1950s, efforts to identify its members led to Maarten Schmidt's realization in 1963 that the bright optical point source associated with one of these radio sources, 3C 273, possessed the (then) shockingly large redshift of 0.158. This realization cracked open the field of active galaxies; suddenly the unidentifiable optical point sources associated with numerous radio sources became objects with interpretable optical spectra, and known distances and luminosities.

Because radio astronomical techniques are extremely powerful, and a large fraction of all bright radio sources are AGNs, many of the *known* AGNs are strong radio emitters, and a great deal is known about the phenomenology of that emission. For example, only in the radio band is milliarcsecond imaging (by Very Long Baseline Interferometry, or VLBI) currently available, and AGN radio emission is mostly resolved on this angular scale.

However, this depth of knowledge is a bit misleading. Even in AGNs where the radio band is relatively strong, it never accounts for more than $\sim 1\%$ of the bolometric luminosity; in addition, less biased surveys have shown that the great majority of AGNs emit a much smaller fraction of their total power in the radio.

1.3 AGN Nomenclature

Because subclasses of AGNs exist that all share the same choices from this menu, numerous subvarieties have been named. Whether the taxonomical effort that created this "zoo" has helped clarify our view of this subject, or led to further confusion, is a matter of some debate. Nonetheless, to understand the conversation, one must learn the language.

Table 1.2 presents the lineaments of AGN zoology. The names themselves reveal how roundabout scientific progress can be. Some are descriptive: *radio-loud* and *radio-quiet* are fairly self-explanatory terms, and, as we shall discuss in Chapter 9, the distribution of the fraction of the bolometric luminosity that is radiated in the radio band is quite bimodal. *OVV* is an acronym for "Optically Violently Variable," a term that is equally direct: this class is marked by exceptionally rapid and large amplitude variability in the optical band. Other designations refer to the names of the first people to identify the class: Carl Seyfert pointed out the first six *Seyfert galaxies* (thirty years later they were subdivided into two principal types, according to whether their emission lines did or did not have broad wings: see §10.1.3); the "FR" in *FR1* and *FR2* stands for Fanaroff and Riley, who pointed out an interesting distinction in both luminosity and morphology among the radio galaxies; this will be discussed in detail in Chapter 9. Some class titles are deliberate coinings. *Quasar* was originally the pronounced form of "QSRS," an acronym for "quasi-stellar radio source," but over the years its usage evolved so that it now denotes nothing about the radio luminosity of an object. Today it is often used as a synonym for "generic AGN"; very low luminosity AGNs are sometimes called "micro-quasars," for example. And some have truly quirky histories. Variable stars are given names having two letters and an abbreviated form for the constellation in which they are found. The prototype of the AGN variety now called *BL Lac objects* was originally thought to be a variable star in the constellation Lacerta and was therefore given the name "BL Lac." Quirkier still is the origin of the term *blazar* (not found in the table because, 20 years after its introduction, its use is still somewhat nonstandard). This coinage is meant to unite the OVV and BL Lac classes and was invented as a joke by Ed Spiegel, the after-dinner speaker at the first conference organized on BL Lac objects. The name stuck both because the two varieties do resemble each other (compare their two lines in the table, deliberately put immediately adjacent to each other) and because its connotation of "blazing" is very appropriate to these objects whose power output varies so dramatically.

Other classes have been defined in the past, but are now rarely used. *N galaxies* are elliptical galaxies with bright optical nuclei, generally first noticed because of their large radio power. Most would now be called *broad line radio galaxies.* "QSO," a term still occasionally used, is an acronym for "quasi-stellar object" that has now been largely replaced by *radio-quiet quasar.* Osterbrock and his associates defined several Seyfert types intermediate between 1 and 2 (see §10.1.3 for a more detailed discussion); these classes can still sometimes be seen in the literature.

The column headings are abbreviations for items in the menu: "Point-like" refers to whether an optical point source can be seen. "Broad-band" means that there is comparable luminosity in the infrared, optical, and X-ray bands. "Broad lines" and "narrow lines" indicate the existence in the optical and ultraviolet spectra of lines several thousand km s^{-1}, or several hundred km s^{-1} in width, respectively. "Radio" means that the fraction of the luminosity emitted in the radio is relatively large, perhaps $\sim 10^{-3}$ of the bolometric luminosity (defining radio as $\nu \sim 10$ GHz). To be considered significantly variable, the members of the class should vary by an order of magnitude or more in the optical band over a timescale comparable to a human lifespan. To receive a "yes" in the "Polarization" column the optical light should be at least a few percent linearly polarized.

The astute reader will notice that these classifications divide up into groups: radio-loud versus radio-quiet, strongly variable versus all others, and narrow emission lines only versus broad emission lines as well. These groupings suggest an arrangement in a three-dimensional parameter space, illustrated in figure 1.9. The existence of these groupings has led to a great deal of effort to explain these multiple categories in terms of a smaller number of variables. These efforts are discussed in detail in Chapter 12. For now, it is best simply to take note of the connections.

This same hypothetical reader will notice that type 1 Seyfert galaxies have identical table entries to those of radio-quiet quasars. He or she might then ask why they are listed separately. That is a good question. In fact, there are a number of objects that have been classified "type 1 Seyfert" by some observers, and "radio-quiet quasar" by others. In practice, the only distinction is whether a host galaxy is visible. When it is, the AGN is called a Seyfert galaxy, whereas when none is visible, it is called a quasar. Since the distinction depends on the resolution and background level of the instrument, and does not depend primarily on the object itself, it is clearly not a very useful one. Its only objective correlate is with luminosity: Seyfert galaxies are on average two orders of magnitude less powerful than quasars. It is for this reason, of course, that the host galaxy is visible, for we can only see host galaxies when the luminosity from the AGN does not overwhelm the starlight.

Table 1.2: The AGN Bestiary

Beast	Pointlike	Broad-band	Broad Lines	Narrow Lines	Radio	Variable	Polarized
Radio-loud quasars	Yes	Yes	Yes	Yes	Yes	Some	Some
Radio-quiet quasars	Yes	Yes	Yes	Yes	Weak	Weak	Weak
Broad line radio galaxies (FR2 only)	Yes	Yes	Yes	Yes	Yes	Weak	Weak
Narrow line radio galaxies (FR1 and FR2)	No	No	No	Yes	Yes	No	No
OVV quasars	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BL Lac objects	Yes	Yes	No	No	Yes	Yes	Yes
Seyferts type 1	Yes	Yes	Yes	Yes	Weak	Some	Weak
Seyferts type 2	No	Yes	No	Yes	Weak	No	Some
LINERs	No	No	No	Yes	No	No	No

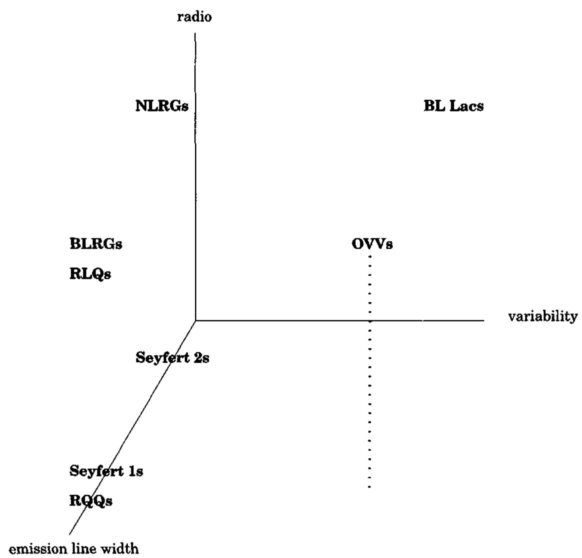


Fig. 1.9 The principal subvarieties of AGNs schematically arranged according to relative power in the radio band, emission line width, and variability. All combinations are possible except that there are no highly variable radio-quiet objects.

The last entry in the table, *LINERs*, denotes a category at the very margin of activity. As of this writing, it is still debatable whether these galaxies are truly active or not, or whether it is even a well defined category in any physical sense. Dubbed "LINER" for "Low-Ionization Nuclear Emission Region" by Tim Heckman in the early 1980s, these galaxies show strong emission lines, like Seyfert galaxies, but the relative strengths of those lines from low-ionization stages are rather greater than in Seyfert galaxies (see §10.1.3 for a quantitative definition).

It is also important to recognize that these categories may not exhaust the true list. It is entirely possible that there are varieties of AGNs that we do not yet recognize. Some people believe, for example, that the ultraluminous infrared galaxies discovered by the *Infrared Astronomical Satellite* (*IRAS*) may house AGNs (§13.6). Similarly, others suspect that there may be a class of *type 2 quasars*, by analogy with the division of the Seyfert class (§12.6.5). At present, limits on the existence of such objects are very weak. Surprises are always very much a possibility in the field of AGNs.

2 How to Find AGNs

As we have just seen, the definition of what constitutes an active galactic nucleus is a purely observational construct, and a somewhat waffly one at that. Therefore, what different people mean by the term tends to be based on their favorite means of finding AGNs. This anarchy does not lead to catastrophe because, fortunately for us, there is substantial overlap between the different search techniques in use. However, it does mean that one must be very careful when working at the quantitative level to make explicit one's usually tacit prejudices of definition. Because much hard work is necessary to accumulate AGN samples, the people responsible often develop very strong emotional attachments to them and make strong claims for their unique benefits. In evaluating these claims, one should always remember that *all surveys are biased.* This is not a pejorative remark; it simply reflects the fact that no single definition of AGNs has yet been found to be all-inclusive.

On the other hand, one should not lapse into complete relativism. There are certain qualities that distinguish good surveys from poor. Perhaps the most important of these is the ability of the observers to actually follow their ground rules. In that case, one can at least be confident that the actual selection criteria correspond to the ones stated. It also helps for these criteria to be framed simply, so that their implications can be (comparatively) easily recognized. Surveys also differ in their efficiency (the rate at which confirmed candidates can be found relative to the expenditure of human time, telescope time, and other resources), or in their degree of interest (do they explore new properties or reveal new classes of objects?).

Before describing the principal means different people have used to find, count, and describe AGNs, it is worthwhile to make another comment. If one is to compile a statistical sample, the large number of objects required to make the statistics meaningful generally means that one is resource-limited, that is, the information collected on each object will be the bare minimum required to fill its table entries in the sample. Because more detailed studies inevitably require more time both at the telescope and in analysis, in any given category there are generally only a few favorite objects about which a great deal is known. Consequently, the several brightest cases of each type are observed over and over again, with the hope that the detailed properties discovered in them will prove to be representative of the general class. Because they have little choice, many people proceed on this basis, but you should be warned that this is a dangerous expectation.

2.1 Optical Color

Most searches for AGNs are based on their extremely broad-band continua (see §1.2.3). In the optical band, that fact has the consequence of creating color distinctions between AGNs and ordinary stars or galaxies. These distinctions are the basis for the single most popular AGN survey technique. More surveys for AGNs have been conducted using optical color criteria than by any other method. This is partly because so many people are trained in the techniques of filter photometry, and partly because these methods are comparatively efficient—a few images suffice to supply a very large number of candidate objects.

To conduct such a survey, one first takes an image of the sky in some region through at least two different optical filters. In any given field, most of the objects in the image will not be AGNs. Stars and galaxies always outnumber AGNs, no matter where the field is or what its magnitude limit is. To cull the AGNs from the background, one must apply a filter to the object list.

The essence of the color selection method is that one can find AGNs because their optical spectra have different shapes from those of stars and galaxies, and hence the ratios of their fluxes in different bands (their "colors") can be distinguished. The distinction comes from several sources. First, the continua of AGNs tend to be broad and smooth, while stellar (and hence galactic) spectra are thermal and therefore strongly curved (see figs. 1.1 and 1.2). Thus, the color method is based directly on searching for objects with the sort of broad-band continua remarked on in §1.2.3. The breadth of typical AGN spectra means that they have more flux in both the UV and IR relative to the V band than do stars. Second, when the redshift is great enough to bring Lyα into the visible band, this emission line can be strong enough to substantially increase the flux in the band into which it falls. Third, when the redshift is somewhat greater yet, intervening Lyα and Lyman continuum absorption can remove enough flux to depress the band(s) in which this absorption falls.

Of course, if the colors of the AGN plus its host are to be different enough from normal galactic colors to be distinguished, it immediately follows that the optical luminosity of the AGN (in our frame) must be at least comparable to, or preferably much greater than, the luminosity of the host. Consequently, color surveys are automatically biased against low-luminosity AGNs. They also often carry a second, implicit bias: surveys hoping to find nearby AGNs often require all candidates to be extended in order to eliminate stars quickly, while surveys hoping to find more distant AGNs often require all candidates to be points in order to eliminate faint galaxies. The former criterion misses AGNs in which the nucleus is much more luminous than its host; the latter is biased in exactly the opposite way.

To see how surveys of this sort work in practice, consider first a simple cut in the single-color space of $U - B$ (the peak in the transmission coefficient for a standard "U" filter comes at about 3600 Å, while the greatest transmission for a "B" filter is near 4300 Å). The hottest (therefore UV-brightest) main sequence stars have $U - B \simeq 0.4$–0.5 mag. In contrast, typical (low-redshift) AGN spectra have $U - B \simeq -1$ mag. Thus, nearly all stars and galaxies immediately fall out if one takes only the "UV-excess" objects into the survey. The major contaminants in an UV-excess survey are hot subdwarfs (sdB stars) and white dwarfs, with both (and especially the subdwarfs) becoming relatively less of a problem as the flux limit is pushed to lower levels. Of course, stars of all varieties are fewer when the survey is done at high Galactic latitude. Conversely, the density of stars is so great in the plane of the Galaxy proper that it is frustratingly inefficient to look for AGNs within $\simeq 10°$ of the Galactic plane.

One of the earliest color surveys was done by Markarian (1967; the final list is in Markarian, Lipovetsky, and Stepanian 1981). In this case, in order to eliminate subdwarfs, he required his objects to be extended. The survey covers a large part of the northern sky ($\simeq 10,000$ square degrees) and extends down to about $m_B = 15.5$ mag. Roughly 1500 galaxies are on the Markarian list, about 10% of them Seyfert galaxies. Although the sample definition has never been well calibrated, this survey was an important pioneering effort, for it turned up many interesting objects.

The most thoroughly studied color survey is the Palomar Bright Quasar Survey of Schmidt and Green (1983). This survey is known by two alternate acronyms, BQS and PG (for Palomar Green). Selecting only the pointlike objects with $U - B < -0.44$ mag, they found $\simeq 1700$ candidates brighter than $m_B \simeq 16$ mag in 11,000 square degrees of sky. Follow-up spectroscopy confirmed that 114 are AGNs, for an AGN efficiency of $\simeq 7\%$.*

To find fainter AGNs by the same technique, one takes longer exposures on a smaller solid angle. Boyle et al. (1990), for example, were able to find quasars down to a limit of $m_B = 20.9$ mag, but over a field of only 245 square degrees. Their color cut was very similar to that of the BQS, $U - B < -0.35$ mag.

Unfortunately, we do not know whether *all* quasars have such smooth "blue" optical spectra. The only way to find out whether we are missing any by selecting only on the basis of UV excess is to search using other criteria. To this end, there have also been a number of searches using multiple colors. For example, Warren, Hewett, and Osmer (1991) took images in five filters: U, B, V, R, and I (the V band is centered on 5400 Å; R on 6400 Å; and I on 8500 Å). Stars occupy a well defined "sausage" in four-dimensional color space (five flux measurements yield four colors, as it is only flux ratios that count), so that Warren et al. could define their candidates to be any object with a stellar image that lay a certain number of standard deviations outside this sausage in any direction. Indeed, by generalizing the color criterion, Warren et al. were able to find many AGNs that would not have been found by a pure UV excess criterion. A still larger survey of this variety is planned as part of the Sloan Digital Sky Survey (SDSS). Again there will be five color images, but the limiting magnitude for the images will be $m_B \simeq 22$ mag, and one-quarter of the sky will be searched. With this size survey, $\sim 10^6$ pointlike quasar candidates should be found, and $\sim 10^5$ (most of the brightest $\sim 10\%$) confirmed by spectroscopy.

One final difficulty must be noted that applies to all optical surveys: there can be substantial intervening absorption between us and the AGNs we seek. Some of this absorption is local—dust extinction in the plane of the Milky Way is quite substantial—but very distant intervening absorption becomes increasingly important the farther out we look. On average, gas clouds optically thick at the Lyman edge are found on any given line of sight with a density per unit redshift roughly $0.25(1+z)^{1.5}$ (Stengler-Larrea et al. 1995); if one with $z \geq 4$ happens to lie between us and an AGN, it would blank out the spectrum in the B band and blueward. Large numbers of neutral H clouds with smaller column densities can also be found. Their numbers increase rapidly with redshift, leading to a significant depression of AGN continua at observed wavelengths shorter than $1216(1+z)$ Å when

*Viewed from a different perspective, the BQS survey had a much greater efficiency for finding either hot subdwarfs ($\simeq 40\%$) or white dwarfs ($\simeq 20\%$). Wisely not letting good data go to waste, Schmidt and Green became experts in the statistics of subdwarfs and white dwarfs as a by-product of performing an AGN survey.

$z \geq 2$. Thus, this intervening absorption artificially diminishes B-band flux when the quasar redshift is greater than about 2.5. In addition, the density of galaxies on the sky is great enough that, if they contain significant dust, the light from many very high redshift quasars may suffer considerable extinction. Taken together, all these mechanisms can have the effect of making it increasingly difficult for us to find AGNs by their (observed frame) optical light as their redshifts increase.

2.2 Optical Emission Lines

Another distinguishing characteristic of (some classes of) AGNs is strong optical and ultraviolet emission lines (as discussed in §1.2.4). These can carry enough flux that they are apparent in even crude, low-dispersion spectra. Because very few stars have emission lines, this contrast in properties creates another opportunity for a potentially efficient AGN search technique. The main limit is that conventional spectroscopic methods yield only one spectrum per observation. Given the density of faint objects on the sky ($\sim 10^4$ per square degree at a flux limit of $m_B \simeq 22$ mag), one clearly requires a much more rapid device to make this approach practical.

The solution adopted by many different groups is to place a dispersing element across the telescope's light path before the light hits an imaging detector. These devices are variously called "objective prisms," "grisms," and "grenses," but their basic principles are all the same. The image obtained when any one of them is used consists of a large number of strips, each one a low-dispersion spectrum of the object at that location. It is relatively easy to search these strips for emission bumps that identify AGN candidates (in fact, grisms were used in the Markarian survey to make it easier to find continuum color anomalies). Higher quality spectra are then taken for all the AGN candidates to confirm their identifications and more accurately measure such properties as redshift, emission line fluxes, and continuum shape. Searches of this sort are especially useful for finding high redshift AGNs ($z > 2$) because the ultraviolet rest-frame emission lines are very prominent, whereas, as we have just discussed in the previous section, the colors of high redshift AGNs are hard to define a priori.

A number of surveys of this sort have been very important historically. The very first AGNs to be identified as interesting objects, although they were not thought of as "AGNs" until much later, were the six emission line galaxies singled out by Seyfert in 1943 because they had strong, high-ionization emission lines that were concentrated in their nuclei. More recently, emission line surveys have done the most to expand our knowledge of high-redshift quasars; the first survey able to discover large numbers of quasars with $z > 2$ was done by Hoag and Smith (1977). In the earliest such surveys, though, there were problems with establishing a quantitative definition for the selection criteria. The search for emission bumps was done by eye, and it was found post hoc that different people had substantially different subjective selection criteria.

In modern versions of this survey technique, as much as possible is automated to avoid subjectivity and nonuniformity. To construct the "Large Bright Quasar Sample," for example, 800 square degrees of sky were searched down to $m_B = 18.8$ mag (Foltz et al. 1987, 1989; Hewett et al. 1991; Chaffee et al. 1991; Morris et al. 1991; Hewett, Foltz, and Chaffee 1995). Each spectrum produced on the image by the objective prism was measured, and candidates were selected on the basis of extreme blue color, strong emission or absorption features, or strong continuum breaks. This survey found $\simeq 1000$ quasars. Similar techniques were employed in the survey of Schmidt, Schneider, and Gunn (1986a,b; Schneider, Schmidt, and Gunn 1994). In their work, the candidates were selected on the basis of a threshold in emission line equivalent width and signal/noise ratio. This survey proved especially successful in discovering very high redshift quasars. Despite the great improvement in this method provided by automation, it still rises or falls, of course, on the validity of its fundamental assumption: that quasars can be distinguished on the basis of sharp spectral features such as emission lines.

There is another version of this search technique that is, in some ways, the least biased method of finding AGNs yet invented. We began this section by noting that because of the rarity of AGNs, any practical search technique must employ an efficient discriminator. But this requirement evaporates if the main point of the survey is to study normal stars or galaxies (imagine, for example, that the BQS survey had been motivated primarily by searching for hot dwarf stars). In that case, effort expended on the non-AGNs isn't wasted, and any AGNs found come "for free." The only special provisos are that the survey be large enough to uncover a statistically interesting number of AGNs, and that the selection criteria should not in some way eliminate AGNs. Moreover, because surveys focused on normal stars or galaxies will be keyed to their properties, they are (almost) independent of the properties of any AGNs they turn up and are therefore largely unbiased with respect to AGN variety or characteristics. The principal bias such surveys impose is between AGNs with and without visible hosts. That is, galaxy surveys ignore AGNs with such great contrast relative to their hosts that they appear pointlike, whereas stellar surveys ignore AGNs in which the hosts are visible. A secondary bias that arises in galaxy surveys is that they are also likely to miss very low luminosity AGNs because such weak AGNs do not contribute enough light to whole-galaxy spectra to be recognized.

Two such surveys have been especially useful. One is the CfA redshift survey. Primarily designed to study large-scale structure in the Universe, this sample contains all galaxies within selected regions of the sky brighter than a certain visual magnitude limit (originally the limit was $m_V = 14$ mag, but an extended survey stretched the limit to $m_V = 15.5$ mag, and further stretching is possible). Because the essential quantity desired for mapping is the redshift, spectra were taken of all the galaxies in the sample. Measuring a redshift demands a higher quality spectrum than merely looking for large equivalent width emission lines, but not such high quality that the job consumes prohibitive amounts of telescope time. Spectra of this quality are, however, quite good enough to enable a quantitative check for large equivalent width emission lines. On this basis Huchra and Burg (1992) found roughly 50 Seyfert galaxies, divided approximately equally between types 1 and 2. As the total sample contained $\simeq$ 2000 galaxies, they found that the local Seyfert fraction is $\simeq$ 2.5%. There are, however, two biases in this survey, which act in opposite directions. On the one hand, an especially bright type 1 nucleus could increase the total flux from a galaxy sufficiently to bring it into the sample even though the host flux by itself would not be great enough. This effect overestimates the Seyfert fraction. On the other hand, low-luminosity Seyfert nuclei of both types are missed because their lines do not have great enough equivalent width relative to the stellar continuum of the host galaxy.

The second such sample is the Revised Shapley-Ames catalog of bright galaxies (Sandage and Tammann 1981). This sample is based on photographic plates and so has a fuzzy magnitude limit around $m_{\rm pg} = 13$ mag. Of its $\simeq$ 1300 galaxies, $\simeq$ 50 are Seyfert galaxies of luminosity comparable to those found in the CfA survey, and another $\simeq$ 20 are radio galaxies. However, using very small apertures centered on the nuclei of a large (486 galaxies) complete subsample of this catalog, Ho, Filippenko, and Sargent (1997a) showed that very low luminosity AGN activity is extremely common: roughly 40% of the RSA galaxies contain LINERs, and 10% have Seyfert nuclei.

The Sloan Digital Sky Survey, mentioned above in the color context, also falls into this category. In this survey, galaxies will be selected for spectroscopy on a basis wholly independent of their nuclear properties, and therefore AGNs in hosts bright enough to select, but not so bright as to swamp the AGN, should be very fairly sampled.

2.3 Radio Flux

2.3.1 The 3C catalog and the confusion limit

Historically, the first method by which AGNs as such were found was the identification of radio sources (radio emission is the distinguishing mark of AGNs discussed in §1.2.7). In the early 1950s, radio telescopes with large enough total apertures to detect first dozens, and then hundreds, of sources were built by groups at both Cambridge University (U.K.) and Sydney University (Australia). However, it took almost a decade to assemble the first useful large catalog, the 3C (for 3rd Cambridge) catalog, a collection of almost 500 sources brighter than 9 Jy at 178 MHz. A few years after the initial publication of the 3C catalog, a revised version (called 3CR) was compiled with more accurate fluxes (Bennett 1962). This list contains 328 sources with $\delta > -5°$ and the same flux limit as in the original version of the 3C catalog.

The name of this catalog calls for a digression: why is the first radio source catalog called the third? The reason is instructive, and its lessons are by no means limited to radio surveys. It is solely a matter of historical accident that these lessons were first learned as a result of doing surveys in that band; they apply with equal force all the way across the electromagnetic spectrum.

Suppose that you use a telescope whose primary beam has an area of 1 square degree, and that in the entire sky there are 1000 sources bright enough for you to detect. Then in each telescope beam, the expectation value of finding a source is $\simeq 0.024$, and the probability of finding two in any particular beam is 3×10^{-4}. Because there are 41,000 beam areas over the sky, one expects there will be $\simeq 12$ beams in which two objects are found. Thus, you will measure the correct flux for $\simeq 990$ of the 1000 objects you find.

The situation changes dramatically, however, if the number of sources per beam bright enough to detect increases much beyond this level. If instead there are 10,000 sources bright enough to detect, the expectation value per beam is now 0.24, and one expects multiple sources in 2.5% of the beams, or 1020 cases. Now 10% of all the beams with sources have misleading flux levels, for wherever several sources can all be found within a single beam, their fluxes have been mistakenly combined. Moreover, the list which you are attempting to create is missing all but one of the sources at each location where several sources are merged into one.

In fact, the situation is even worse than this. Typically, the number of sources per unit solid angle increases fairly rapidly with decreasing flux (measuring this rate of increase is a matter of considerable interest; see Chap. 3). That means that if the probability of finding two individually detectable sources in a single beam is interestingly large, the probability of finding two sources, each slightly weaker than the sample's flux limit, is even greater. If their mean flux is greater than half the flux limit, their total flux exceeds the threshold for inclusion in the catalog. Thus, by this means inadequate angular resolution can create large numbers of spurious sources.

Defining the angular resolution becomes a bit tricky when the telescope beam is synthesized interferometrically. The beam shape of a single round dish is quite simple and concentrated: outside the central maximum, the Airy rings decline rapidly in amplitude over an angular scale of order the ratio of the wavelength to the telescope diameter (see §13.1.1). However, depending on its configuration, the beam shape of an interferometer can have very complicated side lobes, sometimes spreading all across the field of view. When this is the case, a single very bright source can create numerous spurious weak sources if these side lobes are not carefully removed. Thus, although one measure of the angular resolution is the width of the central peak in the beam pattern, if there is a large dynamic range between the brightest and faintest sources, the character of the side lobes also figures into what one means by the angular resolution. For this reason, it is wise to place surveys looking for faint sources in regions of the sky where there are no bright sources.

This whole set of problems is known as the *confusion limit* and was not well understood at the time of the first and second Cambridge surveys. Unfortunately for the astronomers compiling them, their surveys were, in fact, badly confusion limited. This fact was brought pointedly to their attention in a classic paper by members of the Sydney group (Mills and Slee 1957). Ultimately, the problem was solved by an upgrade of the Cambridge radio telescope to twice its former operating frequency, and the resulting fourfold decrease in the beam size permitted the measurement of reliable fluxes and positions on the third go-around.

One might think that the lesson of the confusion limit is to build survey telescopes with as few sources per beam as possible. Although there is some truth in this, it cannot be pushed to an extreme. For one thing, it is obvious that such an approach is very inefficient: why waste time observing all those empty fields? However, there is another, subtler problem. All flux measurements are subject to noise. If a large positive noise fluctuation occurs in a location on the sky where no real source exists, one might mistakenly identify it as a real object. Such large fluctuations are rare, but if nearly all antenna beams are empty, the chances are good that spurious sources will appear in at least some of them.

To illustrate this point quantitatively, let us return to the numerical example posed for our discussion of the confusion limit. When there are 1000 real sources distributed over 41,000 beams (mean number of sources per beam $= 0.024$), there are 40,000 empty beams. If the system noise is Gaussian, and we set the detection threshold at 3σ, the probability per beam of finding a spurious source is 0.00135, so the expectation value for the total number of spurious sources is $\simeq 540$; that is, the catalog is inflated by 50% over its real size! Increasing the detection threshold to, say, 5σ, makes a big improvement: the probability per beam of finding a spurious source falls to $\simeq 3 \times 10^{-7}$. Thus, the optimum balance between angular resolution and limiting flux is set by a trade-off between the conflicting goals of avoiding source confusion, maximizing observational efficiency, and rejecting spurious sources (a particularly complete discussion of these issues can be found in Murdoch, Crawford, and Jauncey 1973).

2.3.2 Optical identifications and the discovery of quasars

Listing radio sources and their fluxes has much intrinsic interest, but astronomical objects almost always yield the most detailed information when studied in the visible band. The familiar complexities of atomic spectroscopy give us a tremendous boost in diagnosing the circumstances inside distant objects. Consequently, as soon as the first sources were discovered, a few perceptive astronomers launched optical observational programs to find what optical sources might be identified with them and to study their properties.

The first step in this effort was to achieve enough angular resolution with radio observations that the optical objects in the radio error boxes were few enough to make it possible to choose which of them was responsible for the radio emission. The Sydney group was the first to achieve high enough angular resolution to permit identification of a few “radio stars” with optical counterparts and found (somewhat shakily) that both M 87 and NGC 5128 were the homes of powerful radio sources (Centaurus A is centered on the latter galaxy). Just a few years later, positions for a few more bright radio sources were refined by an order of magnitude, allowing Baade and Minkowski (1954) to measure the redshift of the radio galaxy Cygnus A, a (then) astonishing 0.056. By 1960, Minkowski was able to identify the radio source 3C 295 with a galaxy at $z = 0.46$ (Minkowski 1960).

Other 3C radio sources (e.g., 3C 48 and 3C 273) proved to be more difficult to identify. In these cases, unlike, for example, 3C 295, the optical image of the most likely candidate in the radio error box was pointlike, and so didn't appear to be a galaxy at all. Moreover, their optical spectra were dominated by emission lines, a fact that greatly puzzled astronomers used to the absorption line spectra of normal galaxies. Without guidance about what sort of object they might be, optical astronomers had few clues about what sorts of lines to look for in the spectra, and the measured wavelengths were not even close to anything familiar. The mystery was only cleared up in 1963 when Maarten Schmidt realized that the spectra could be understood if they were adjusted for what was still considered (despite the example of 3C 295) an extraordinarily large redshift—0.158 for 3C 273 (Schmidt 1963).

The fact that the first quasars were discovered through the identification of radio sources led to another instructive mistake. The very term *quasar*, now used for any high-luminosity AGN with broad optical and ultraviolet emission lines (and sometimes for AGNs of any variety), was coined as a contraction of "quasi-stellar radio source," and for almost a decade it was believed that the typical high-luminosity AGN (although to use this terminology is an anachronism) was "radio-loud." Only after about 1970 was it realized that only about 10% of all quasars radiate as much as 0.1% of their total luminosity in the radio range. This historical misunderstanding reemphasizes that surveys find what they are looking for, and that there is no single best way to find all examples of a phenomenon as diverse as AGNs.

Although the effort consumed nearly 30 years, the 3CR catalog is now virtually completely identified (one or two objects still remain outstanding). Its full redshift range extends past $z = 3$. Roughly three-fourths of its members are radio galaxies, with most of the rest quasars, but there also a few BL Lac objects and Seyfert galaxies (and also a small number of galactic supernova remnants, like the Crab Nebula).