THE RISE AND FALL OF QUASARS

Maarten Schmidt
Palomar Observatory
California Institute of Technology
Pasadena, CA 91125, USA

It is a great honor for me to speak to you tonight on the subject of quasars. In the past twenty-five years, an enormous amount has been learned about these objects and the field has become very technical. I recognize that there are many people in the audience who are not professional astronomers and I hope not to disappoint them. At the same time my colleagues may recognize a few things that I touch upon. They will also hopefully forgive me for leaving out many attributions: many hundreds of astronomers have contributed to the field.

1. DISCOVERY OF QUASARS 1960-63

The discovery of quasars happened during a concentrated effort in the 1950s and 1960s to obtain optical identifications of extragalactic radio sources. Many of the strong radio sources were identified as giant elliptical galaxies. Minkowski's success in confirming the identification of 3C 295 with a galaxy of redshift 0.46 was a major achievement in 1960.

It was arround that time that Tom Matthews, at Caltech, obtained an accurate position of the radio source 3C 48. Allan Sandage took a plate of the field in September 1960, which showed a stellar object with faint nebulosity at the radio position. He also obtained the first spectra of the stellar object which contained strong broad emission lines that could not be identified. Photometry showed that the object had an ultraviolet excess and that it varied in brightness over several months. In the announcement in the March 1961 issue of Sky and Telescope, the possibility that it was a galaxy was dismissed in favor of an interpretation in terms of a peculiar local star.

In 1961 and 1962 Tom Matthews concentrated on the radio positions of radio sources of small angular diameter in the hope of detecting galaxies in very distant clusters. Without knowing it in advance, he attempted in those two years optical identifications of as many as seven radio sources that eventually turned out to be quasars. Some of the optical objects were misidentifications, including 3C 273 which was identified with a galaxy. In the other cases, the optical objects were

D. McNally (ed.), Highlights of Astronomy, Vol. 8, 33–44.

https://doi.org/10.1017/S1539299600007498 Published online by Cambridge University Press

stellar, but the spectra which I obtained defied interpretation.

The source 3C 273 provided the clue to the mystery of the radio stars. Hazard et al. had been observing lunar occultations of the source in 1962, showing that the source was double. One component coincided with a bright star, of magnitude 13, and the other with a jet-like feature. I suspected that the bright star was a foreground object unrelated to the radio source. Essentially with the idea of eliminating the star from consideration, I took some short-exposure spectra of it in December 1962. Surprisingly, the star showed a number of broad emission lines.

It took another six weeks for the mystery to be resolved. Hazard had written up the occultation results for publication and suggested that I write a companion article about the optical spectrum. While writing the manuscript, I took another look at the spectra. I noticed that four of the emission lines showed a pattern of increasing strength and increasing separation from blue to red. For some reason, I decided to construct an energy level diagram based on these lines. The results seemed to contradict the regular spacing of the lines. To check on that, I proceeded to take the ratio of the wavelength of each of the lines to that of the nearest Balmer lines, which do have a regular spacing pattern. The first ratio, of the line at 5630 Å to H-beta, was 1.16. The second ratio was also 1.16. When the third ratio was 1.16 again, it came as a flash that I was looking at a Balmer spectrum redshifted by 16 per cent.

A little while later, Jesse Greenstein and I were able to identify the emission lines observed in 3C 48 at a redshift of 37 per cent. The impact of the realization that fairly bright star-like objects could have large redshifts was stunning. We soon had intense discussions about what caused these large redshifts and within a week came to the seemingly reckless conclusion that these "stars" were at distances of thousands of millions of the light years. Considering the controversy about quasar distances that was to develop later, I believe in hindsight that we did all right in the time span of a week. I will come back to the interpretation of the redshifts presently.

2. OPTICAL PROPERTIES

In the first few years following the discovery of their redshifts, there was little understanding of the quasar phenomenon, as was well expressed in George Gamov's exhortation in the May 25, 1964 issue of Newsweek:

> Twinkle, twinkle, quasi-star
> Biggest puzzle from afar
> How unlike the other ones
> Brighter than a billion suns
> Twinkle, twinkle, quasi-star
> How I wonder what you are.

As a consequence, their definition had to be established empirically from observed properties listed below.

https://doi.org/10.1017/S1539299600007498 Published online by Cambridge University Press

2.1 Star-Like Image

Quasars were originally required to have a dominant star-like component. This came from the observation that quasars look at first sight just like stars. In practice, that meant that much of the light came from a diameter of less than one arcsecond. However, nebulosity such as seen around the nucleus of 3C 48 is probably the rule.

2.2 Emission Lines

The emission lines shown by quasars are Lyman-alpha (1216 Å), C IV (1549 Å), C III (1909 Å), and Mg II (2798 Å). Quasar spectra also show the Balmer lines of hydrogen and sometimes forbidden lines as seen in planetary nebulae. The lines are typically 50-100 Å wide, corresponding to velocities around 5000 km s^{-1}.

2.3 Large Redshifts

Recognition of the line pattern is usually unambiguous, so there is rarely any major uncertainty about a quasar's redshift. The redshifts range from a few per cent to several hundred per cent. The largest redshift is 4.43 for a quasar found by Warren et al. in Cambridge. This redshift is as large as a bullet shot with a velocity that is 93.4 per cent of the speed of light. The total number of quasars with known redshift is around 4000. Their total number in the universe is of the order of millions.

2.4 Variability

Variability, such as first seen in 3C 48, turned out to be common in quasars. The observations of 3C 273 go back one hundred years--they show at times rapid variations, with a time scale of about a month. This means that the size of the light emitting region cannot be larger than a light month. In some quasars, and recently in 3C 273, even faster variations have been observed, producing upper limits to the size of the objects of the order of light weeks or light days.

3. INTERPRETATION OF REDSHIFTS

The interpretation of the redshifts of quasars has given rise to much controversy. I believe that this was partly caused because quasars were discovered too early. In the early days, I used to discuss four hypotheses for the redshift. We will recall these briefly. I will discuss the cosmological hypothesis, which is accepted by most astronomers today, last.

3.1 Local Explosion

This early proposal by Terrell posed that quasars were all ejected by our own Galaxy. Velocities of ejection would be large, with many at 80

https://doi.org/10.1017/S1539299600007498 Published online by Cambridge University Press

or 90 per cent of the velocity of light. From our position about 30,000 light years off the center of the Galaxy, we would expect to see angular motion on the sky unless the objects are very far away. No angular motions were observed, requiring a minimum age of the explosion of about 10 million years. My own arguments against this hypothesis were mostly based on the total mass of the ejected quasar cloud. One might also expect in this case to see quasars around other galaxies--which should give rise to clustering of quasars and to blueshifts of quasars, both of which are not observed.

3.2 Gravitational Redshift

This interpretation was one that Greenstein and I considered carefully after the discovery of redshifts. On the basis of spectroscopic and dynamical arguments, we were able to show that the mass of quasars in this hypothesis had to be larger than that of the most massive galaxies known.

3.3 No Physical Interpretation

This was an extraordinary option, mostly based on Arp's finding of associations on the sky between quasars and galaxies at much smaller redshifts. These associations according to Arp show that quasars are at much smaller distances than would correspond to their cosmological redshift distance. If the cosmological redshift interpretation is correct, these associations on the sky must be just an accidental effect of projection of foreground objects against background objects. Much of the debate about this issue has centered on whether statistical arguments for or against the physical reality of these associations were valid.

3.4 Cosmological Redshift

If the redshifts of quasars are cosmological, then they are very luminous, highly compact objects. In early discussions of the nature of the redshifts, I tended to argue that the alternative interpretations led to objects that were at least as exotic as cosmological quasars. We can, however, show that quasars are at cosmological distances even ignoring their large redshifts!

The argument is based on the counts of quasars as a function of magnitude. Quasar surveys conducted by a number of teams of astronomers have shown that the number of quasars rises by a factor of 8 per magnitude, from magnitude 15 to 19, as shown in Figure 1. It can easily be shown that objects having a uniform distribution in local Euclidean space show a rise in numbers of a factor 4 per magnitude. Therefore, the space density of quasars must increase with distance, if they reside in local Euclidean space. Since quasars here are isotropically distributed, this means that we are located in a density minimum of quasars. This situation is in conflict with the Copernican principle, which states that it is very unlikely that the human observer is at a special location in the universe. We can only escape the Copernical condemnation if we

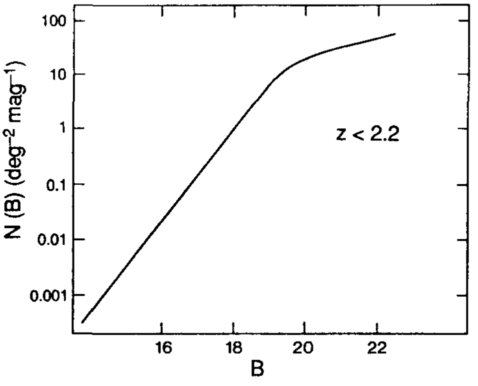

Fig. 1. Counts of quasars with redshifts less than 2.2, as a function of blue magnitude B. The curve represents the work of many astronomers.

put quasars at distances so large that their light takes most of the age of the expanding universe to reach us. In that case, fainter more distant quasars reflect the situation in the universe at earlier epochs. Since we accept that the expanding universe shows evolution, the increasing density toward earlier epochs is no objection. Ryle used this argument in the 1950s to show that extragalactic radio sources had to be very distant, even before most of their redshifts had been measured. I believe that this line of reasoning provides the most direct support for the cosmological redshift of quasars.

4. NATURE OF QUASARS

On the cosmological interpretation of the redshifts, quasars are at distances from 1 to 13 million light years, assuming a Big Bang model of the universe with an age of 15 thousand million years. Their luminosities range from 10^{11} to 10^{14} solar luminosities, while the size of the central luminous part is probably less than a light week as we mentioned earlier. Compared to the most luminous galaxies, quasars are up to a thousand times more luminous, yet their size is 100,000 times smaller! It is not surprising then that for many years the relation between quasars and galaxies remained unclear, in particular since no quasars were found in rich clusters of galaxies.

A useful clue has been provided by the nebulosity, such as seen around 3C 48. Optical imaging of quasars at modest redshift has shown that they are all surrounded by nebulosity and it is now believed that all quasars are resident in a host galaxy. The dominant light of the quasar makes such studies difficult. It is expected that the Hubble Space Telescope with its high spatial resolution will make a major contribution to the study of host galaxies of quasars.

The spectral energy distribution of quasars is remarkably broad.

Each octave of frequency or wavelength contributes a roughly equal amount of energy all the way from gamma-ray or X-ray energies to infrared or radio wavelengths. Individual quasars differ strongly in their behavior at radio wavelengths. Only a small fraction emit strong radio radiation, such as shown by 3C 48 and 3C 273. Essentially all quasars are strong X-ray emitters.

The spectral energy distribution shows a number of peaks that suggest a number of components. Studies of variability of the different components are very useful in the diagnostics of a complex situation; these require coordinated ground-based and space observations.

The enormous energy output of quasars, coupled with their very small size, requires a compact, massive object. An attractive hypothesis is that the central body is a black hole of about 1000 million solar masses. Material falling into the black hole from the surrounding galaxy would be responsible for the enormous energy output of quasars. The total size is about equal to the orbit of Saturn in our solar system. The emission lines are formed in clouds at greater distances, of a few light years.

From the beginning, our understanding of quasars has been hampered by the fact that the central object could not be resolved optically. How well would we understand the Andromeda nebule with its spiral arms, cepheids, clusters, rotation, etc., if it presented itself to us as a point source?

At radio wavelengths, the situation is more favorable, for two reasons. First, radio quasars often do show structure on a scale of seconds of arc. Jets are usually prominent, sometimes stretching over enormous distances into the intergalactic medium. Second, radio astronomy has succeeded in creating images that show detail a thousand times finer than optical images. This is achieved by using radio telescopes located continents apart as an interferometer. This allows studies on a scale of light years, close to the central engine of quasars.

The resulting radio maps of quasars show variations from month to month. Generally, a component appears that moves away from the central source at high speed. In fact, the speed naively calculated is usually five to twenty times the speed of light, hence the term superluminal motion. Three radio blobs in 3C 273 moved apart 25 light years in the three year interval 1977-1980! Since speeds larger than the velocity of light are frowned upon, this finding at first seemed to argue strongly against the large distances that follow from the cosmological redshifts. It may be shown, however, that relativisitc ejection of material in a direction close to the line of sight to the observer leads to superluminal motion. This does not mean that the frequent occurrence of superluminal motion among quasars is fully understood. The orientation of these sources relative to the observer is the subject of much discussion.

5. QUASARS AS PROBES

Besides the emission lines observed in quasar spectra, they also exhibit absorption lines, mostly at redshifts well below that of the emission lines. A single quasar can show absorption lines at many different redshifts: these are formed by gas clouds at different locations on our line of sight to the quasar. There are two main types of absorption-line systems. The first one shows both hydrogen lines and absorption lines of heavier elements such as carbon, silicon, iron, etc. These systems are caused by galaxies that are on the line of sight to the observer. The second type of absorption system consists only of the Lyman-alpha line of hydrogen. They are observed at wavelengths below that of the Lyman-alpha emission line. These systems are probably gas clouds that have not formed stars yet and therefore have not experienced metal enrichment through stellar evolution, supernova explosions, etc. The only evidence for the existence of these pure hydrogen gas clouds is in the absorption spectra of quasars.

Quasars have also led to the first observation of the gravitational lens effect that was expected on the basis of Einstein's General Theory of Relativity. The lens effect typically manifests itself in multiple images of the quasar, with each image having an identical spectrum and redshift. The different images reach us through different paths around the lens. Since the light travel time along the two paths may differ by many years, any time variation in the spectrum of quasars would show up as a difference in the spectra of the images. This makes the distinction between a pair of quasars at the same redshift and a true lens effect sometimes difficult in individual cases. Most of the lenses are combinations of a galaxy in a cluster and the cluster itself. Standard theory requires that the number of images must be odd. Ironically, most if not all cases show an even number of images--perhaps the last image is very faint or very close to the quasar. Here, again, the Hubble Space Telescope with its high spatial resolution will add much to our knowledge.

6. SPACE DISTRIBUTION OF QUASARS

You will recall that we discussed earlier the steep increase of the counts of quasars with magnitude and how that increase could be used to argue against the local hypothesis. I would now like to discuss the space distribution of quasars based on the cosmological interpretation of their redshifts. In order to do this, we have to find and count quasars, i.e., to take their census, before we can derive their space density.

Let me first explain how we search for quasars. In the 1960s, it was noticed that all quasars identified from radio sources had a blue color, or so-called ultraviolet excess. This turned out to be a property of all quasars of redshift less than 2.2, regardless of their radio emission. As an example, consider the Palomar-Green survey, started by Green by taking two-color exposures of 266 fields in the northern sky with the Palomar 18-inch Schmidt telescope. The millions of objects on these exposures were digitally recorded and those of the appropriate blue color selected. This produced a sample of several thousand objects, of which individual spectra were taken with the 60 and 200-inch telescopes at Palomar. This long-term study produced the PG catalog, which is a complete sample over 10,714 square degrees of

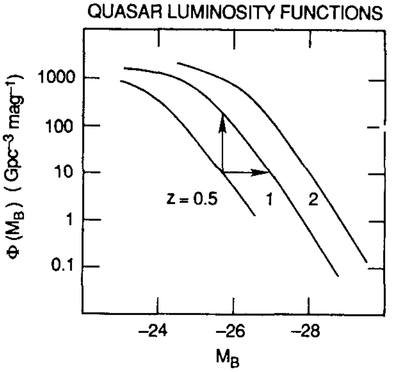

Fig. 2. Schematic representation of quasar luminosity functions based on a review by Boyle et al. Density evolution is represented by the vertical arrow, luminosity evolution by the horizontal arrow.

around 1700 blue objects to an average limiting blue magnitude of 16.2. Among them were 115 quasars which constitute the Palomar Bright Quasar Survey.

Another important survey is that by Boyle et al., which covers 4.2 square degrees and is complete to a magnitude of 20.9. The method of selection was essentially the same as for the Bright Quasar Survey. Since the fields were relatively small, spectra could be taken with a fiber-fed spectrograph allowing dozens of spectra to be observed simultaneously.

With the date from these suveys and many others in hand, the space density of quasars of given luminosity at given redshift can be derived. Before I show the results, let me emphasize that they reflect many years of effort and telescope time by many groups of astronomers.

Since quasar numbers depend on absolute luminosity, we express quasar space densities in terms of the luminosity function, which represents the space density per interval of luminosity. We see from Figure 2 that at given redshift, say 0.5, the luminosity function declines steeply with increasing luminosity: quasars of the highest luminosity are very rare. We also see that the luminosity function of quasars at different redshifts is different. At given luminosity, say absolute magnitude -26, the space density increases by a factor of about 130 from a redshift of 0.5 to 2 (cf. Fig. 3).

An alternative way of looking at these luminosity functions is in terms of a horizontal shift, as illustrated in Figure 2. The shift between the luminosity functions at redshift 1 and 2 is about a magnitude. In other words, the change can be understood if all quasars brightened by about a magnitude between redshifts 1 and 2. In fact, all the luminosity functions can be understood if quasar brightened by about 5 magnitudes, or a factor of 100, from redshift zero to redshift 2.

These two alternative ways of looking at quasar evolution are usually labeled density evolution and luminosity evolution. Much discussion has been going on lately about the relative merits of these points of view. Luminosity evolution could be understood if quasars

https://doi.org/10.1017/S1539299600007498 Published online by Cambridge University Press

are long-lived, that is they all formed at or before redshift 2 and all dimmed uniformly by 5 magnitudes to the present. It is easy to show that if quasars radiate 10 per cent of the rest-mass energy of the accreted mass, then under the hypothesis of luminosity evolution the typical quasar would have a mass of around 10^{11} solar masses. This is much larger than the mass deduced from the dynamics of the emission-line region of quasars, which is around 10^9 solar masses.

It is essential for the hypothesis of luminosity evolution that the luminosity function of Seyfert galaxy nuclei be added to the quasar luminosity function since at large redshifts many of these would brighten to become quasars. This creates another problem. Seyfert galaxies contribute at X-ray energies about 30 per cent of the total X-ray background at 2 keV. If they brighten as required in the luminosity evolution scenario, their X-ray contribution would exceed that of the observed background.

For these reasons, I prefer to think in terms of density evolution, *with quasars typically living only ten or a hundred million years. In* this scenario, quasars were more numerous at larger redshifts because their birth rates were higher at earlier cosmic times.

7. THE REDSHIFT LIMIT

Even though the space density of quasars of given luminosity rises sharply with redshift, there has been a suspicion for a long time that quasars of very high redshift may not exist. To illustrate, consider the quasar KPS 2000-030 with a redshift of 3.78. Its optical magnitude is around 18. Such a quasar at a redshift of 5 would be about magnitude 19, and therefore easily observable, yet not a single one has been found yet. In order to investigate this, we have to look carefully into the ways quasars of large redshift are discovered.

Quasars at redshifts beyond 2.2 do not have blue colors any more, therefore a different search technique has to be used. The Cambridge group headed by Hewitt uses many different colors and separates out those that show color combinations unlike those of the majority of the stars. Slit spectra are taken of all the candidates. Quasars with colors similar to those of stars will be missed in this technique.

An alternative method is the objective prism or grism technique. Each object in the field is recorded as a short, low-resolution spectrum. Candidate quasars are selected by searching for an emission line in the short spectra. This technique has been used for many years using photographic plates and visual inspection of the spectra. Even though around a thousand quasars have been found this way, the resulting space densities are of little value as a census of quasars, because the selection effects can not be quantified.

Several years ago, Schneider, Gunn and I decided to start a program using this method in an objective way. We employ CCDs as detectors, and search for emission lines in the digitally recorded spectra by computer. Soon it developed that we needed to cover larger areas of the sky to discover a substantial number of quasars. This was a problem, because a CCD at the 200-inch telescope covers less than

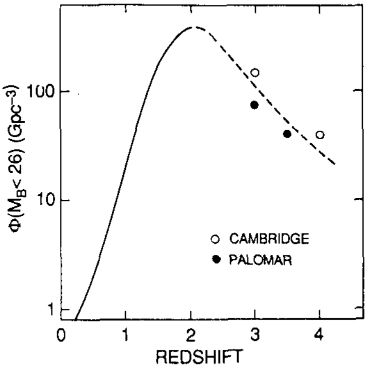

Fig. 3. Co-moving space density of quasars more luminous than M_B = -26, as a function of redshift. The points represent preliminary results of two searches for high redshift quasars.

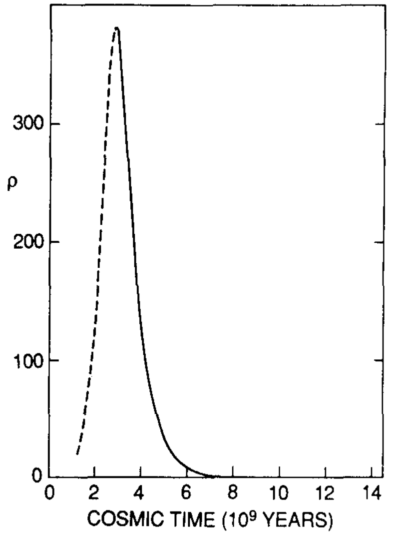

Fig. 4. Co-moving space density of quasars more luminous than M_B = -26, versus cosmic time.

one-hundredth of a square degree. We resolved the problem by stopping the tracking of the telescope: the array of four CCDs are read out exactly at the sidereal rate with which objects move across them. The

https://doi.org/10.1017/S1539299600007498 Published online by Cambridge University Press

exposure time is very short, of course, generally less than a minute. In a 7-hour exposure a long strip of sky can be observed with a total area of 14 square degrees.

There are many details of this approach that warrant mentioning; suffice it to say that one night's surveying records tens of thousands of short spectra, and that the computer search for emission lines produces typically hundreds of candidates. In the course of this work we have found two quasars with redshifts over 4.

We now have tentative results for one survey of 14 square degrees, in which we found 44 quasars. Among these quasars, 15 have a redshift larger than 3. Since the candidate search was done by computer algorithm, the selection effects pertaining to this survey can be evaluated quantitatively. It turns out that our survey essentially covered quasars of absolute mangitude brighter than -26 at redshifts above 2.7. This is the same luminosity for which we found an increase in co-moving space density by a factor of 130 from redshift 0.5 to 2. The points to the right of Figure 3 are from our group and from first results announced by the Cambridge group. I want to emphasize that our results are still tentative. Also, I am presenting the Cambridge results in a different fashion than they did.

It is clear from these results that the co-moving space density of quasars brighter than -26 is reaching a maximum at a redshift of 2 or 2.5 and that it then declines rapidly. Some of our colleagues have suggested that we are missing high redshift quasars in our census as a consequence of dust absorption in the distant universe. This appears unlikely, since extragalactic radio sources, not affected by dust absorption, also peak at a redshift of 2.

At this stage it is important to realize that at increasing redshifts we are looking back to progressively earlier epochs. It makes sense, then, to replace the redshift by the appropriate cosmic time: 0 for the Big Bang and 15 thousand million years for the present. Also, I will use a linear scale for the density rather than the logarithmic scale used here.

The curves drawn in Figure 4 correspond precisely to those in Figure 3. Here for the first time, we see the rise and fall of quasars. Quasars started appearing perhaps a thousand million years after the Big Bang, reached a sharp maximum at 3 thousand million years, and since that time have declined to almost nothing. Clearly, the fireworks has long been over. The fact that we see the height of the fireworks at all is due to the finite velocity of light, which allows us to look back to these early epochs.

The sharp maximum must signify an extraordinary event in the early universe. Now, as you know, astronomers have been looking for many years for the epoch of galaxy formation. Theoreticians place this event anywhere in the range of redshift 2-10. When galaxies start forming, the collapse of the inner parts must provide plenty of material to accrete onto any black hole present in the center of the galaxy, producing a bright quasar. Perhaps the black holes were formed in the same period. In any case, I would argue that the peak of quasar activity and the formation of galaxies are probably related, directly or indirectly.

https://doi.org/10.1017/S1539299600007498 Published online by Cambridge University Press

I am sure you realize that his scenario is very speculative on my part. For that reason, I am awaiting with some trepidation Professor Martin Rees' judgment on these matters in his Invited Discourse on the formation of galaxies next Monday night.

https://doi.org/10.1017/S1539299600007498 Published online by Cambridge University Press