THE ASTROPHYSICAL JOURNAL, 460:L81–L84, 1996 April 1

# QUASARS FROM GALAXY COLLISIONS WITH NAKED BLACK HOLES

MASATAKA FUKUGITA
Institute for Advanced Study, Princeton, NJ 08540, and Yukawa Institute, Kyoto University, Kyoto 606, Japan

AND

EDWIN L. TURNER
Princeton University Observatory, Princeton, NJ 08544
*Received 1995 December 18; accepted 1996 January 23*

## ABSTRACT

Motivated by a set of new *Hubble Space Telescope* observations by Bahcall and collaborators, we propose that the universe contains a substantial independent population of supermassive black holes and that QSOs are a phenomenon that occurs when they collide with galaxies or gas clouds in the intergalactic medium. We argue that this new model is not in conflict with any available constraints and would also naturally explain one puzzle for the conventional model, the very rapid QSO population decline toward low redshift. The hypothesis therefore bears further investigation.

*Subject headings:* black hole physics — cosmology: theory — quasars: general

## 1. INTRODUCTION

In the now well established conventional view (see Rees 1984 and references therein), quasi-stellar objects (QSOs) and related active galactic nuclei (AGNs) phenomena are explained as the result of accretion of plasma onto giant black holes which are postulated to form via gravitational collapse of the high-density regions in the centers of massive host galaxies. This model is supported by a wide variety of indirect evidence and seems quite likely to apply at least to some observed AGN phenomena. However, one surprising set of new *Hubble Space Telescope* (*HST*) observations directly challenges the conventional model. We propose here an alternative possibility: the universe contains a substantial independent population of supermassive black holes, and QSOs are due to their collisions with galaxies or gas clouds in the intergalactic medium (IGM). This hypothesis would naturally explain why the QSO population declines very rapidly toward low redshift, as well as the new *HST* data.

The recent direct observation, which calls the traditional model into question, is the result of attempts to image the host galaxies of low-redshift QSOs using *HST*. So far, 20 such systems have been imaged deeply with WFPC2 (Bahcall, Kirhakos, & Schneider 1995a, 1995b, 1995c). A few of the images show the expected normal, giant galaxy with the QSO shining from its nucleus; however, the remainder show a somewhat bewildering array of different local environments for the QSO activity: In some cases the QSO is positioned in the midst of what appears to be a system of galaxies in collision but not associated with any obvious galaxy nucleus. In many cases, there are a few dwarf galaxies within several kiloparsec of the QSO but none detectable (to limits well below $L^*$) directly associated with it. There are even cases in which the QSO seems to have no particular association with any visible galaxy, aside from being a part of some possible loose galaxy group. In any case, the data indicate that a major fraction of at least low-redshift QSOs do not conform to the most straightforward predictions of the conventional scenario.

In fact, the search for host galaxies has a long history (Kristian 1973; Gehren et al. 1984; Malkan 1984; Smith et al. 1986; Veron-Cetty & Woltjer 1990; Hutchings & Neff 1992) of ground-based studies. The results, however, have not been definitive. A few observations have recently made using the *HST* with the hope of resolving this issue (Hutchings et al. 1994; Disney et al. 1995; Bahcall et al. 1995a, 1995b, 1995c). The *HST* results, however, also do not seem conclusive even with very high resolution afforded by *HST* due to a number of difficulties (Hutchings 1995): while Hutchings et al. and Disney et al. have reported the presence of normal galaxies, Bahcall et al. find that host galaxies are at most subluminous or entirely absent for a number of cases. McLeod & Rieke (1996) have reanalyzed some of the data of Bahcall et al., using a different method and claimed the presence of luminous but low surface brightness host galaxies. Their method, fitting the azimuthally averaged profile with the sum of a point spread function and a spherically symmetric disk, however, will yield a positive disk component if there is any excess signals near the point image. Hence, it does not actually contradict the claim by Bahcall et al. Here we do not attempt to resolve this observational issue, but rather we examine the implications of Bahcall et al. results, since their analysis seems at least as careful and convincing as the others.

The difficulties for the standard scenario raised by observed QSO population evolution (see Hartwick & Schade 1990) are neither so direct nor so clearly recognized, but they may also represent important clues. There are in fact three such puzzles. The first is that at $z > 4$, when the universe was less than 10% of its present age (for $\Omega_0 = 1$), the most distant QSOs we have so far located were as luminous and roughly as numerous as those present at any later epoch and far brighter and more common than they are at the present epoch (Turner 1991a, 1991b). Moreover, this population of objects may well be present at even higher redshifts; they are difficult to locate not because they are faint (some $z > 4$ QSOs are brighter than 18th magnitude!), but because the comoving volume per redshift interval is decreasing or increasing less rapidly than at low redshift (depending on the cosmological model) and the QSO light is rapidly shifting into the near IR. The second problem for the conventional model is to explain why the giant black hole remnants of the QSOs, which were so luminous at

$z \sim 2$, are so dark and inactive at the present, despite the presence of a dense ISM and stellar population in the nuclei of many giant galaxies (Goodman & Lee 1989). The third puzzle is the remarkably fast drop in the QSO population at redshifts below about 2; during this period the comoving emissivity of luminous QSOs drops by orders of magnitude and with a halving time substantially shorter than the concurrent cosmic expansion timescale (Turner 1991b).

It is useful to recall that the natural *a priori prediction* of the conventional model would be quite different. Most structure formation scenarios (particularly in high $\Omega_0$ universes) predict that increasingly massive objects form at successively later epochs. Moreover, processes of gravitational collapse and accretion would be expected to accelerate and produce ever more massive and rapidly growing black holes, especially in the very high-density environments at the cusps of the nuclei of bright galaxies. This would lead to the expectation of a QSO population becoming more luminous and numerous with time as structure formation and the nonlinear evolution of galactic nuclei proceed. Even if one postulates some limit to terminate the luminous QSO phase, such as exhaustion of the accretion fuel supply or inability to tidally disrupt passing stars (Rees 1990), it is difficult to see why it should apply so synchronously as observed. Of course, given our inability to reliably predict the details of the complex, nonlinear evolution of galactic nuclei and black hole accretion processes, it has proven possible to a posteriori explain the QSO population evolution in various, sometimes ad hoc, ways (Efstathiou & Rees 1988; Small & Blandford 1992; Haehnelt & Rees 1993; Narayan & Yi 1995).

Motivated by the Bahcall et al. (1995a, 1995b, 1995c) observations and these shortcomings of the conventional model we here investigate the alternative hypothesis that the universe contains a substantial population of massive ($\sim 10^8$ $M_\odot$) black holes existing independently of any host galaxy and perhaps even formed by rather different physical mechanisms. QSOs are then identified with the accretion luminosity and other activity generated when one of these "naked" black holes collides with a galaxy or a massive IGM cloud. In addition to trying to account for the unexpected results of the *HST* study, it is expected that such a scenario will be able to naturally account for the large and extremely rapid decrease in QSO activity in the recent history of the universe. Thus, although the postulated "naked" black hole population may seem rather ad hoc, it offers the possibility of understanding at least two otherwise quite puzzling observations.

This simple scenario raises a host of detailed theoretical considerations. These include the efficiency with which the black hole can accrete gaseous material from the galaxy or cloud with which it is colliding (as a function of gas density, angular momentum, composition, . . . ), the required number density of the "naked" black holes (relative to available baryons and other limits), their formation mechanisms and epochs, possible effects on the cosmic radiation background (CRB) spectrum, gravitational lensing consequences, relation to those QSOs and AGNs that do seem to reside in the nuclei of $L \sim L^*$ galaxies, possible explanations for QSO phenomenologies and classifications, and so forth. In this discussion, we consider a few of the more critical points briefly, but we do not claim to have resolved any of these theoretical issues conclusively. Rather, our main purpose is to suggest a new qualitative scenario for the nature of at least some QSOs and to show that it is not immediately excluded by any simple considerations.

## 2. THE MODEL

Let us discuss our model. The typical black hole mass of interest is $\sim 10^8$ $M_\odot$ at around $z \sim 2$. The characteristic luminosity is set by the Eddington limit, at which radiation pressure on free electrons balances gravitational forces:

$$L_{\rm E} = 4\pi G M_h m_p/\sigma_{\rm T} = 1.3 \times 10^{46} M_8 \ {\rm ergs\ s^{-1}}, \qquad (1)$$

where $G$ is Newton's constant, $M_h$ is black hole mass, $m_p$ is proton mass, $\sigma_{\rm T}$ is the Thomson scattering cross section, and $M_8$ is the black hole mass in units of $10^8$ $M_\odot$. This is a typical luminosity for bright QSOs. Accretion rates needed to maintain this luminosity are of order $2\epsilon_{0.1}^{-1}$ $M_\odot$ yr$^{-1}$, where $\epsilon_{0.1}$ is the fraction of the accreted material's rest energy, which is emitted radiatively in units of 0.1. In particular, it is expected that a black hole moving with a characteristic velocity $v$ through a diffuse medium of density $n$ will produce an accretion luminosity

$$L = 1.0 \times 10^{45} M_8^2 \epsilon_{0.1} \left(\frac{v}{100\ {\rm km\ s^{-1}}}\right)^{-3} \left(\frac{n}{3\ {\rm cm^{-3}}}\right) {\rm ergs\ s^{-1}}, \qquad (2)$$

(Hunt 1971; Lacey & Ostriker 1985), which is of order the luminosity of the low-redshift QSOs observed by Bahcall et al. Some combination of higher densities in the ambient medium, lower encounter velocities and/or more massive black holes would be required to achieve $10^{46}$ ergs s$^{-1}$ luminosities characteristic of high-redshift QSOs.

The very sensitive dependence of $L$ on $v$ is a particularly interesting feature of this scenario. The characteristic infall velocity of several hundred km s$^{-1}$ into a typical giant galaxy would depress the expected luminosity well below characteristic bright QSO values. Collisions with low-mass (dwarf) galaxies might therefore produce higher luminosity events, a correlation opposite the one usually attributed to the conventional QSO scenario in which a loose correlation of QSO luminosity with host galaxy mass is often supposed. It might also contribute to the explanation of the surprisingly low luminosities of many of the host galaxies detected in the *HST* studies. Of course, small relative velocity between gas in the galaxy and the infalling black hole might also be achieved in other ways. Dynamical friction of the black hole with material in the galaxy will act to reduce the relative velocity, especially for high-mass black holes and relatively dense galaxies. Alternatively, random chance coincidence of the black hole velocity vector with the internal motions of gas in the galaxy, as for example in a "prograde" collision with the disk of a spiral galaxy, can produce a small $v$ even if the black hole is moving quite rapidly with respect to the galaxy center of mass.

What cosmic density of black holes is required to produce the observed QSO population? Let us write the comoving number density of black holes as $n_h$, and that of galaxies as $n_g$. The encounter rate is then written

$$R = n_h n_g (1+z)^6 \sigma v\,, \qquad (3)$$

where $\sigma$ is the cross section such that an encounter gives a QSO and $v$ is the relative velocity. For an order-of-magnitude estimate, let us assume that 30% of baryons are captured in galaxies and 10% becomes black holes of mass $10^8$ $M_\odot$. We assume that masses of galaxies are distributed according to the Schechter luminosity function with $M_{\rm baryon}/L \sim 10$ and that the relative velocities of a galaxy and a black hole are typically

100 km $s^{-1}$. We also assume that QSO activity takes place when black holes crosses the galaxy within its Holmberg radius (at a baryonic surface density $\approx 10\ M_\odot\ \mathrm{pc}^{-2}$); we estimate the size of a galaxy to be $\sim 14(M_B/10^{11}\ M_\odot)^{0.4}$ kpc. We take the slope of the Schechter function to be $\alpha = -1.5$, in agreement with the numerous dwarfs reported in several recent studies (Impey, Bothun, & Malin 1988; Turner et al. 1993; Dalcanton 1995). The integral over the Schechter function is dominated by objects with $10^8$–$10^{10}\ M_\odot$ of gaseous material.

We note that the mass of $10^8\ M_\odot$ is close to the minimum mass needed for bright QSOs; a less massive galaxy would not supply sufficient fuel to sustain the QSO luminosity through a characteristic crossing time. A fuel reservoir of $10^9\ M_\odot$ could shine for about $10^9$ yr at the Eddington luminosity if $\epsilon_{0.1} = 1$. From equation (3) we find that the encounter rate $R \sim 4.0 \times 10^{-22}(1+z)^6\ h^{-2}\ \mathrm{s}^{-1}\ \mathrm{Mpc}^{-3}$ or equivalently, $1.3 \times 10^{-4}(1+z)^6\ h^{-3}\ \mathrm{Mpc}^{-3}$ per inverse Hubble time, where $h$ is Hubble's constant in units of 100 km $\mathrm{s}^{-1}$ $\mathrm{Mpc}^{-1}$. Thus, the cumulative number of collisions (which we identify as QSO outbursts) around $z \sim 2$ is about 0.018 $h^{-3}$ $\mathrm{Mpc}^{-3}$ in physical density or 0.0007 $h^{-3}$ $\mathrm{Mpc}^{-3}$ in comoving units. These numbers are a conservative estimate, since in reality, we expect some correlation in the distribution of QSOs and galaxies which would significantly increase the rate of collisions. In addition, QSO activity might result from the collision of black holes with some of the denser clouds in the IGM, rather than with a galaxy, thus further increasing the collision rate or reducing the required black hole population. (It might reasonably be objected that the metallicity inferred from typical QSO emission line spectra is far higher than that attributed to IGM absorption line systems, but we note that these measurements refer to volumes of very different scales, densities and physical conditions. In other words, the gas immediately surrounding the QSO in its emission line region could well be greatly enriched in metal content via relatively small amounts of star formation, either directly or indirectly associated with the QSO activity itself.)

We note that the total mass density of black holes that ever shone as QSOs can be estimated reliably from their cumulative observed flux (Soltan 1982). A modern estimate is $n_h \sim 0.001$–$0.002 M_8^{-1}$ $\mathrm{Mpc}^{-3}$ (Chokshi & Turner 1992), in comoving coordinates. The fact that this number density based on the observed QSO population agrees with that inferred above from the collision rate calculation, within the substantial uncertainties of the input parameters (including $h$) and our very simplified treatment, is encouraging. This density is about 0.1 $h^{-3}$ that of luminous galaxies and is at least $\sim 10^2\ h^{-3}$ times higher than the peak number density of QSOs, which already suggests that QSOs are made and fade one after another.

A particularly notable feature is the $(1+z)^6$ dependence of the encounter rate. Since we expect that the lifetime of $10^8$–$10^9$ $M_\odot$ QSOs is of the order of $0.6$–$6 \times 10^8$ yr, considerably shorter than the Hubble time, we predict that the number density of QSOs decreases as $(1+z)^6$ toward $z = 0$ in gross qualitative agreement with observations. Of course, since the other factors in equation (3) ($n_g$, $\sigma$, $v$) might evolve with redshift, the situation could be considerably more complex, but these effects will be dominated by the $(1+z)^6$ factor unless the evolution is extreme.

## 3. DISCUSSION

The *HST* observations show that some QSOs are located in the nuclei of host galaxies, sometimes spirals and sometimes ellipticals. This does not necessarily contradict the basic model presented here. While most QSOs would lose their activity rather quickly, either due to the end of the collision or the exhaustion of fuel in low-mass galaxies, close encounters with more massive galaxies will sometimes lead to capture of the black hole via dynamical friction and produce much longer lived activity. The dynamical friction will eventually bring the black hole into the center of a galaxy. The timescale for a $10^8$ $M_\odot$ black hole to spiral into the nucleus of a typical giant galaxy from an initial radius of $\sim 10$ kpc is of order $10^{10}$ yr (Begelman, Blandford, & Rees 1980). It is interesting to see that the nucleus of 3C 273 is not at the center, and there are a few other examples seen in the QSO sample of Bahcall et al. In a predictive sense, when the active nucleus is found at or near the center of a galaxy, we would expect a massive, high-density host galaxy capable of producing strong dynamical friction.

As for the formation of black holes, our suggestions are not more than speculative. According to standard hierarchical clustering models, small mass objects collapse earlier; typically one expects large numbers of bound objects of $\sim 10^5$–$10^7\ M_\odot$ before $z \sim 10$. Alternatively, more unconventional structure formation models such as PBI (Peebles 1987) or cosmic textures (Gooding, Spergel, & Turok 1991) can produce very nonlinear structure formation on small mass scales at early epochs. Of course, the black hole formation mechanism might be quite unrelated to those that form galaxies and other familiar structures. It is also worth noting that early formation of a galaxy-independent population of massive black holes has been invoked and investigated in a variety of other astrophysical connections (Lacey & Ostriker 1985; Carr & Rees 1984; Gnedin & Ostriker 1992; Loeb 1993; Umemura, Loeb, & Turner 1993; Loeb & Rasio 1994; Eisenstein & Loeb 1995).

In any case, we may suppose that a small fraction of baryons go into black holes when Compton cooling is very efficient (i.e., $z > 10$), as a generalization of the Efstathiou & Rees (1988) scenario. Some of these black holes may eventually grow to $10^8\ M_\odot$ by $z = 3$–4. Let us assume that the initial black hole mass is $M_{\mathrm{hi}}$ formed at around $z \sim 10$–20. If a fraction $f$ of the binding energy of the forming black holes is deposited into the CRB; then the amount of distortion to be observed as Zel'dovich-Sunyaev effect is $y_c = \delta\rho/4\rho_{\mathrm{CRB}}$, which must be smaller than the observed limit $2.5 \times 10^{-5}$ (Mather et al. 1994). This means that the initial mass of black hole be smaller than $\sim 10^5\ M_\odot$ for $f = 0.1$. On the other hand, the characteristic accretion time is $t_E = 4 \times 10^8$, and the mass of a black hole can grow as fast as $M_h(t) \sim \exp(t/\epsilon t_E)$ where $\epsilon$ is the radiative efficiency, usually assumed of the order of 0.1. Hence, the available time is enough for more than 20 $e$-folds, sufficient to bring the mass to $10^8\ M_\odot$ well before $z \sim 4$, though the availability of an adequate accretion fuel supply is a nontrivial requirement (Turner 1991a). This exponential increase of the hole mass, and hence accretion luminosity, would explain the rapid rise of a bright QSO population before $z = 3$–4 (Schneider, Schmidt, & Gunn 1989).

Gravitational lensing provides a potential direct method for detecting the postulated black hole population (Press & Gunn 1973). Unfortunately, however, the angular splittings of roughly $0''.01$ and per source probabilities of multiple imaging of less than $10^{-3}$ at $z = 2$ are so small as to preclude any useful tests of our hypothesis based on available datasets.

Our primary conclusions can then be stated as follows: The conventional and in many respects successful model for QSOs is severely challenged by recent *HST* data and has difficulty accounting for their well-established population evolution. The alternative scenario suggested above could better account for these observations and does not manifestly violate any other empirical constraints. Thus, we believe it merits further exploration.

We would like to thank John Bahcall and Sofia Kirhakos for valuable discussions. Avi Loeb also provided important comments on a draft of the manuscript for which we are very grateful. This research was supported in part by US-NSF grants PHY94-07194 at the Institute for Theoretical Physics in Santa Barbara and AST94-19400 at Princeton University. M. F. also thanks the Fuji-Xerox Corporation for generous support.

## REFERENCES

Bahcall, J. N., Kirhakos, S., & Schneider, D. P. 1995a, ApJ, 435, L11
———. 1995b, ApJ, 450, 486
———. 1995c, ApJ, 447, L1
Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, Nature, 287, 307
Carr, B. J., & Rees, M. J. 1984, MNRAS, 206, 315
Chokshi, A., & Turner, E. L. 1992, MNRAS, 259, 421
Dalcanton, J. 1995, Ph.D. thesis, Princeton Univ.
Disney, M. J., et al. 1995, Nature, 376, 150
Efstathiou, G., & Rees, M. J. 1988, MNRAS, 230, 5P
Eisenstein, D. J., & Loeb, A. 1995, ApJ, 443, 11
Gehren, T., Fried, J., Wehinger, P. A., & Wyckoff, S. 1984, ApJ, 278, 11
Gnedin, N. Y., & Ostriker, J. P. 1992, ApJ, 400, 1
Gooding, A. K., Spergel, D. N., & Turok, N. 1991, ApJ, 372, L5
Goodman, J., & Lee, H. M. 1989, ApJ, 337, 84
Haehnelt, M. G., & Rees, M. J. 1993, MNRAS, 263, 168
Hartwick, F. D. A., & Schade, D. 1990, ARA&A, 28, 437
Hunt, R. 1971, MNRAS, 154, 141
Hutchings, J. B. 1995, Nature, 376, 118
Hutchings, J. B., et al. 1994, ApJ, 429, L1
Hutchings, J. B., & Neff, S. G. 1992, AJ, 104, 1
Impey, C., Bothun, G., & Malin, D. 1988, ApJ, 330, 634
Kristian, J. 1973, ApJ, 179, L61
Lacey, C. G., & Ostriker, J. P. 1985, ApJ, 299, 633
Loeb, A. 1993, ApJ, 403, 542
Loeb, A., & Rasio, F. A. 1994, ApJ, 432, 52
Malkan, M. A. 1984, ApJ, 287, 555
Mather, J. C., et al. 1994, ApJ, 420, 439
McLeod, K. K., & Rieke, G. H. 1995, ApJ, 441, 96
Narayan, R., & Yi, I. 1995, ApJ, 452, 710
Peebles, P. J. E. 1987, ApJ, 315, L73
Press, W. H., & Gunn, J. E. 1973, ApJ, 185, 397
Rees, M. J. 1984, ARA&A, 22, 471
———. 1990, Science, 247, 817
Schneider, D. P., Schmidt, M., & Gunn, J. E. 1989, AJ, 98, 1951
Small, T. A., & Blandford, R. D. 1992, MNRAS, 259, 725
Smith, E. P., Heckman, T. M., Bothun, G. D., Romanishin, W., & Balick, B. 1986, ApJ, 306, 64
Soltan, A. 1982, MNRAS, 200, 115
Turner, E. L. 1991a, AJ, 101, 5
———. 1991b, in ASP Conf. Ser. 21, The Space Distribution of Quasars, ed. D. Crampton (San Francisco: ASP), 361
Turner, J. A., Phillips, S., Davies, J. I., & Disney, M. J. 1993, MNRAS, 261, 39
Umemura, M., Loeb, A., & Turner, E. L. 1993, ApJ, 419, 459
Véron-Cetty, M.-P., & Woltjer, L. 1990, A&A, 236, 69