BINARY BLACK HOLE MERGERS FROM PLANET-LIKE MIGRATIONS

ANDREW GOULD
Ohio State University, Department of Astronomy, Columbus, OH, 43210-1173; gould@astronomy.ohio-state.edu

AND

HANS-WALTER RIX
Max-Planck-Institut für Astronomie, Heidelberg, 69117 Germany; rix@mpia-hd.mpe.de

Received 1999 December 6; accepted 2000 February 2; published 2000 February 25

ABSTRACT

If supermassive black holes (BHs) are generically present in galaxy centers, and if galaxies are built up through hierarchical merging, BH binaries are at least temporary features of most galactic bulges. Observations suggest, however, that binary BHs are rare, pointing toward a binary lifetime far shorter than the Hubble time. We show that, almost regardless of the detailed mechanism, all stellar dynamical processes are too slow in reducing the orbital separation once orbital velocities in the binary exceed the virial velocity of the system. We propose that a massive gas disk surrounding a BH binary can effect its merger rapidly, in a scenario analogous to the orbital decay of super-Jovian planets due to a proto-planetary disk. As in the case of planets, gas accretion onto the secondary (here a supermassive BH) is integrally connected with its inward migration. Such accretion would give rise to quasar activity. BH binary mergers could therefore be responsible for many or most quasars.

Subject headings: accretion, accretion disks — binaries: close — black hole physics — quasars: general

1. INTRODUCTION

Supermassive black holes (BHs) are nearly ubiquitous in nearby galaxy nuclei (e.g., Ho 1999). These BHs formed very early, probably during the epoch of quasars, \( z \approx 2 \), and are now largely dormant remnants of quasars. In the hierarchical picture of structure formation, present-day galaxies are the product of successive mergers (e.g., White 1996), and indeed there is evidence for many mergers in the high-z universe (Abraham et al. 1996). Hence it appears almost inevitable that modern galaxies should harbor, or at least should have once harbored, multiple BHs that were collected during their merger history (Kauffmann & Haehnelt 2000).

BHs of mass \( M \approx 10^7 M_\odot \) will quickly find their way to the center of a merger remnant by dynamical friction. Logically, there are only three possibilities. First, BH pairs could merge to form a single, larger BH. Second, the pairs of BHs could form binaries that would remain at galaxy centers to this day. Finally, a third BH could also fall in, leading to a three-body interaction violent enough to expel any number of the three BHs from the galaxy (Begelman, Blandford, & Rees 1980). While in principle this means that all three holes could be ejected, in practice such a violent ejection event is unlikely unless the binary’s internal velocity is much higher than the escape velocity from the galaxy (\( \approx 2000 \text{ km} \text{s}^{-1} \)); in this case, the binary would be in the late stages of merging anyway (see § 2). Since the broad lines of quasars are not often observed to be displaced from the narrow lines by such high velocities, the fraction of binaries with such high internal velocities cannot be large, and therefore triple ejection cannot be common. Hence, mergers generically produce BH binaries, and these binaries either merge on timescales short compared to a Hubble time, or they are present in galaxies today.

Observationally, there is evidence only for a few massive BH binaries (e.g., Lehto & Valtonen 1996), and in none of these cases is the evidence absolutely compelling. Systematic searches for evidence of binarity have turned up very little (Halpern & Eracleous 2000 and references therein). Theoretically, it has proved difficult to construct viable merger scenarios for these BH binaries. Here we first review this difficulty of driving the merger by any of the stellar dynamical means that are discussed in the literature. We then propose a gas-dynamical alternative.

2. NEAR-IMPOSSIBILITY OF STELLAR DYNAMICS-DRIVEN MERGERS

If a BH binary could (somehow) be driven to a sufficiently small orbit, then gravitational radiation would increasingly sap energy from the system and so engender a merger. For a circular orbit with an initial velocity \( v_{\text{orb}} \), the time \( T \) to a merger due to gravitational radiation is given by

\[
v_{\text{orb}} = c \left( \frac{5}{256} \frac{GM_{\text{tot}}^2}{\mu T^3} \right)^{1/8}
= 3400 \text{ km s}^{-1} \left( \frac{M_{\text{tot}}^2 / \mu}{8 \times 10^8 M_\odot} \right)^{1/8} \left( \frac{T}{10 \text{ Gyr}} \right)^{-1/8},
\]

where \( M_{\text{tot}} = M_1 + M_2 \) is the total mass and \( \mu = M_1 M_2 / M_{\text{tot}} \) is the reduced mass, and where we have normalized to the case \( M_1 = M_2 = 10^6 M_\odot \). Note that for fixed total mass, the equal-mass case gives a lower limit on this required velocity, and that the result depends only very weakly on the total mass.

However, as we now show, it is almost impossible to achieve this velocity by any conceivable stellar dynamical process. The basic problem is that when the orbital velocity \( v_{\text{orb}} \) is about equal to the stellar velocity dispersion \( \sigma \approx 200 \text{ km} \text{s}^{-1} \), the total mass in stars within a volume circumscribed by the BH orbital radius \( (a \approx 5 \text{ pc} M_{\text{tot}} / 10^8 M_\odot) \approx M_{\text{tot}} \). If all of these stars were expelled from the BH binary at speed \( v_{\text{orb}} (M_{\text{tot}} / M)^{1/2} \) (Rajagopal & Romani 1995 and references therein), the binding energy of the binary would increase by only a factor \( \approx e \). How-
ever, to get from a virial velocity of \( \sim 200 \text{ km s}^{-1} \) to \( v_{gr} \)
(eq. [1]) would require \( N \sim 6 \) e-foldings in binding energy. Hence, the binary will clear out a hole in the stellar distribution, and dynamical friction will be shut down (Quinlan 1996; Quinlan & Hernquist 1997; Makino 1997; Merritt 2000).

The most efficient conceivable process to rejuvenate the orbital decay would be to equip the binary with an intelligent “captain.” Like a fisherman working in overfished waters, whenever the captain saw that the binary was running out of stars to expel, she would steer the binary to the densest unexploited region of the galaxy. To effect the merger, this would mean systematically moving through and expelling all the stars within a region containing about \( N \sim 6 \) in stars. For a galaxy with an \( r^{-2} \) density profile, this implies expelling all the stars within a radius \( r = \frac{N \sigma^2}{G M_{tot}} \) pc, where we have made the evaluation for \( M_{tot} = 2 \times 10^8 \) M\(_\odot\) and \( a = 200 \) km s\(^{-1}\).

The real difficulty of the captain’s work is best understood by considering the last \( e \)-folding before gravitational radiation can take over. For \( v_{\text{esc}} \gg \sigma \), the cross section for hard interactions (including gravitational focusing) is \( \pi a^2 \left( \frac{v_{\text{esc}}}{\sigma} \right)^2 \). If each incident particle is expelled with speed \( v_{\text{esc}} (M_1/M)^{1/2} \) (Rajagopal & Romani 1995), then the binding energy \( E_b \) decays at \( \dot{E}_b \approx 2 \pi a^2 \sigma v_{\text{esc}} / \rho M_0 = 2 \pi G \rho a \), where \( \rho \) is the local density. The last \( e \)-folding alone would require a time \( t \approx \frac{2 \pi G a \rho}{v_{\text{esc}}^2} \) \( \sim 250 \) Myr, where we have assumed \( v_{\text{esc}} \approx 60 \) pc and our other canonical parameters. Taking account of both the dynamical friction time and the gravitational adiabatic time (eq. [1]), the absolute minimum time for coalescence for an equal-mass black hole would be

\[
T_{\text{min}} = \int_0^{2 \pi G a \rho / v_{\text{esc}}^2} \left( \frac{16G^3 M_0^3}{5a^3 c^5} + \frac{a^2 \sigma^2}{r^2} \right)^{-1} dr = \left( \frac{5}{16} \right)^{1/5} \left( \frac{a^2 \sigma^2}{r^2} \right)^{1/5} \approx 640 \text{ Myr} \frac{M_0}{2 \times 10^8 M_\odot} \left( \frac{200 \text{ km s}^{-1}}{\sigma} \right)^{-4}. \tag{2}
\]

Thus, even with the captain’s careful guidance, a significant fraction of galaxies would harbor BH binaries, and if she suffered even modest inefficiencies (see below), the coalescence would take a large fraction of a Hubble time. Moreover, comparing this decay rate with the standard formula for the decay of translational energy \( E_t \) (Binney & Tremaine 1987) yields

\[
\frac{d \ln E_t}{dt} \leq 0.1 \left( \frac{\sigma^3}{v_{\text{esc}}^2} \right) \frac{d \ln E_e}{dt}. \tag{3}
\]

That is, \( d \ln E_t / d \ln E_e \approx 10^{-3} \), so that the binary would be driven by dynamical friction back to the center of the Galaxy before it had completed 10\(^{-3}\) of an \( e \)-folding of energy loss. Hence, the captain would have to initiate 10\(^3\) “course changes” in the last \( e \)-folding alone. Since the “captain” must in fact be some random process, the only source of such “course changes” is Brownian motion due to continuous interaction with other compact objects. However, for stars of mass \( m \) in an \( r^{-2} \) profile, the range of such Brownian motion is \( \Delta r \sim m / M_{tot} \), i.e., too small by several orders of magnitude. In contrast to ordinary Brownian motion, the present system has an “external” energy source, the binary’s binding energy. However, it follows from equation (3) that even if all of this donated energy were acquired by the binary’s transverse motion, the Brownian motion would be only slightly augmented. In any event, most of the donated energy goes to the stars, not the binary. Infall of globular clusters might well give the binary an occasional jolt, but these would be far too infrequent to drive the merger.

A more commonly discussed alternative to the intelligent-captain scenario is “filling the loss cone.” However, by Liouville’s (1837) theorem, the phase-space density of the material flowing in to fill the loss cone cannot be any higher than the phase-space density of the material scavenged by the intelligent captain. Hence, exactly the same limit (eq. [2]) applies. The maximal efficiency for filling the loss cone would occur if every stellar orbit within radius \( r = 60 \) pc filled the entire volume within \( r \) and therefore ultimately would come sufficiently close to the center to interact with the BH binary. In this case, the timescale would be \( \sim (4\pi r^3/3)a^2(v_{\text{esc}}^2) \sim 1.4 \) Gyr for a self-consistent \( v_{\text{esc}} \approx 4400 \text{ km s}^{-1} \) and our other canonical parameters. This is the absolute minimum time. Again, it is difficult to believe that this limit is actually saturated because there is no mechanism to guarantee maximal efficiency to the loss-cone filling process. In brief, viable mechanisms to drive a merger by ordinary dynamical friction, while not strictly ruled out, are rather contrived.

The only loophole to this argument is that we have assumed circular binary orbits. If an instability existed that systematically drove the BH binaries toward eccentricity \( e \rightarrow 1 \) orbits, then either the binaries would suffer enhanced gravitational radiation (for a fixed semimajor axis) or they could even merge in a head-on collision. Fukushige, Ebisuzaki, & Makino (1992) first suggested such an instability based on the following qualitative argument: dynamical friction is more effective at low speeds than at high speeds, and hence, in the regime where the ambient particles interact with the binary mainly by encounters with its individual members \( v_{\text{esc}} \leq \sigma \), the binary would suffer more drag at apocenter than at pericenter, tending to make the orbit more eccentric. Fukushige et al. (1992) presented numerical simulations that gave initial support to this conjecture. There are, however, two reasons for believing that this effect cannot drive mergers. First, several groups have conducted more sophisticated simulations, and these do not show any strong tendency for \( e \rightarrow 1 \) (Makino et al. 1993; Rajagopal & Romani 1995; Quinlan & Hernquist 1997). Second, once the binary entered the regime \( v_{\text{esc}} \gg \sigma \), the ambient particles would interact with the binary as whole, and so there is no reason to expect any drive toward high eccentricities.

3. GASDYNAMICAL SOLUTION

Begelman et al. (1980) were the first to suggest that gas infall may “lead to some orbital evolution.” But at the time it was not clear that all other mechanisms to overcome the BH hangup would most likely fail.

To resolve the above dilemma, we suggest that gasdynamics play the decisive role in orbital decay, forcing the secondary BH to “migrate” in toward the primary in a manner analogous to the migration of planets. Such migration has been proposed to account for the discovery of Jovian mass and super-Jovian mass planets at \( \sim 1 \) AU from solar-type stars, while it is generally believed that such massive planets can only be created several astronomical units from the stars (Trilling et al. 1998). Artyomowicz & Lubow (1994, 1996) simulated interactions between moderately unequal-mass binaries and accretion disks, which is more directly relevant to the present case than extreme-ratio (planetary) systems. They did not follow the orbital evo-
lution as has been done in more recent work on planets, but only evaluated the instantaneous effect of the torques. They found that a migration to higher eccentricities was a larger effect than migration to smaller orbits. Regardless of which effect dominates, one would expect the final merger to be from circular rather than radial orbits: if the binary is driven toward radial orbits, its emission of gravitational radiation near pericenter will eventually pull in the apocenter of the orbit, decoupling the binary from the disk and allowing the gravitational radiation to circularize the orbit before final coalescence.

Ivanov, Papaloizou, & Polnarev (1999) studied BH binaries surrounded by an accretion disk, focusing on the case where the total disk mass is less than the mass of the secondary. They found that the binary’s interaction with the disk tends to sap the binary’s angular momentum, and that the timescale for orbital evolution is always smaller than the accretion timescale.

For migration to drive the binary all the way to coalescence, the galaxy merger that creates the BH binary must eventually dump at least $M_2$ worth of gas into the inner ~5 pc of the merger remnant where the binary coalescence has gotten “hung up.” Whether this happens on timescales short compared to a dynamical time at 5 pc ($\sim 10^7$ yr), leading to tremendous gas densities and ensuing rapid star formation (Taniguchi & Wada 1996), or whether the gas accumulates over a longer timescale and so does not trigger a starburst, the basic scenario will be the same.

There is every reason to expect that mergers effect such a gas accumulation. First, quasars must gorge themselves on gas to reach their present size. Hence, regardless of whether our picture of binary mergers is correct, this much gas must find its way to central BHs. Second, there is substantial evidence that many quasars are in either recent merger remnants or at least significantly disturbed galaxies (Kirhakos et al. 1999 and references therein). Hence, it seems likely that mergers are the most efficient means to drive gas to the center. Third, many spiral bulges and ellipticals have cuspy profiles populated by metal-rich stars whose total mass is comparable to that of their massive BHs (van der Marel 1999). Thus, it must be possible to funnel huge amounts of gas to the centers of galaxies.

In planet migration, the migration timescale is similar to the accretion timescale for growing the planet because the two processes are governed by the same phenomena, gravitational torques and dissipation (Trilling et al. 1998; A. Nelson 1999, private communication). Ivanov et al. (1999) find for the case $S = M_2/M_1 < 1$ that $t_{\text{ev}} \sim S^{\eta} t_{\text{acc}}$, where $M_1$ is the disk mass, $t_{\text{ev}}$ is the migration timescale, $t_{\text{acc}}$ is the accretion timescale, and $\eta \sim 0.27$ is a parameter. Since they restricted consideration to $S < 1$, they concluded that $t_{\text{ev}}$ is always less than $t_{\text{acc}}$. However, if their results can be extrapolated to the regime we are considering, $S \sim 1$, then (as would be inferred from analogy with planets) the two timescales are comparable.

A major difference between BH and planet migration is that the BH secondary does not form out of the accretion disk and therefore does not automatically open a “gap” in it. Rather, both the gas and the secondary BH fall toward the primary BH because of the same event (galaxy merger), but by completely different paths. Thus, either may arrive first. Clearly, if the gas settles on scales $\ll$5 pc before the secondary arrives, then it cannot affect the BH migration. If the gas arrives first, but settles on scales $\gg$5 pc, then the secondary will be driven through the disk by stellar dynamical friction. Thus, there should be a grand accretion disk around the primary with a “gap” opened up by the secondary. Material should be transported across this gap to a second, smaller accretion disk surrounding the secondary BH. If the secondary arrives first, then the accretion disk will lie primarily outside the binary orbit, i.e., approximately in line with the scenario examined by Ivanov et al. (1999). Since migration and accretion are both driven by the gas outside the secondary, the secondary will still be surrounded by a smaller accretion disk and will still migrate inward. The total energy liberated by this smaller accretion disk should be $\sim \epsilon M_2 c^2 \sim 2 \times 10^{47} (M_2/10^5 M_\odot)$ ergs, where we have taken the efficiency to be $\epsilon = 0.1$, producing a quasar-like appearance during this phase.

Since the secondary BH and the gas both arrive as a result of the same galaxy merger, the angular momentum vectors of the BH binary and the accretion disk are likely to be roughly aligned with the angular momentum vector of the merger, and hence with each other. This alignment will not be perfect initially. Since the disk has substantially more angular momentum than the binary BH, their interaction will inexorably force the binary into alignment with the disk. Ivanov et al. (1999) study this process in the related case of light disks.

4. DISCUSSION

While our suggestion, driven by the lack of alternatives, makes few unambiguous predictions, it does open several lines of investigation that could help test and flesh out our picture. First, merging binaries would appear very much like quasars, since our picture of the migrating secondary is essentially identical to the standard picture of a quasar. The one difference is that the jet from a migrating BH could precess if the orbit of the secondary were substantially misaligned relative to the accretion disk. At present, however, we have no method of estimating how often significant misalignment should occur.

Second, the redshift of the broad lines from a migrating BH’s accretion disk should be offset from the redshift of the host galaxy (as traced perhaps by the narrow lines). Since the migration probably accelerates with time, most migrating quasars should have $v_{\text{orb}} \sim \sigma$. Nevertheless, some should have substantially higher offsets, and measuring the distribution of these offsets would allow one to trace the migration process. However, if no offsets were observed, this would not in itself rule out our hypothesis. It could be, for example, that migrating binaries in merger remnants are preferentially buried in a larger, roughly spherical cloud of dusty gas. In this case, they would have more similarity to ultraluminous infrared galaxies (ULIRGs) than to quasars, and the line centers of their emission would be at the galaxy velocity, not that of the secondary.

Third, it is at least possible that one would see two broad-line systems, one from the primary and one from the secondary. Since broad lines are by definition broad ($\geq 3000$ km s$^{-1}$), the existence of two systems would not easily be recognized for $v_{\text{orb}} \sim \sigma$. However, distinct peaks might be discernible when the BHs were closer to merger. On the other hand, it may be that the major supply of gas lies outside the orbit of the secondary, and hence the primary does not generate a significant broad-line region.

Fourth, it will be important to carry out simulations to determine whether the gas dissipation timescale is short enough for the accreting material to follow the binary inward. This is certainly the case for the simulations that have been done for extreme mass-ratio (planetary) systems, but the less extreme case also needs to be checked.

Finally, we suggest that migrating BH binaries may simply be the quasars, or at least most of them. They have the same
integrated energy output as quasars, they have the same accretion disk fuel source as quasars, and, like quasars, they turn on in the wake of mergers. It may be easier to move gas inward from ~5 pc scales for a binary BH than for a single BH because the binary would excite spiral density waves in the grand accretion disk and so augment viscous drag. Accretion in the inner disk around the secondary might also be easier than for an isolated BH because of the tidal effects of the primary. If this hypothesis is correct, then quasars should generically show offsets between the centers of their broad and narrow lines with a root mean square of (2/3)^{1/2} σ.

We thank Andy Nelson and Martin Rees for valuable discussions. Comments by an anonymous referee helped clarify several critical points. A. G. thanks the Max-Planck-Institut für Astronomie for its hospitality during a visit when most of the work of this Letter was completed. His work was supported in part by grant AST 97-27520 from the NSF.

REFERENCES

Abraham, R. G., van den Bergh, S., Glazebrook, K., Ellis, R. S., Santiago, B. X., Surma, P., & Griffiths, R. E. 1996, ApJS, 107, 1
Artyomowicz, P., & Lubow, S. 1994, ApJ, 421, 651
———. 1996, ApJ, 467, L77
Begelman, M. C., Blandford, R. D., & Rees M. J. 1980, Nature, 287, 307
Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton Univ. Press)
Fukushige, T., Ebisuzaki, T., & Makino, J. 1992, PASJ, 44, 281
Halpern, J., & Eracleous, J. P. 2000, AJ, in press (astro-ph/9910118)
Ho, L. C. 1999, ApJ, 516, 672
Ivanov, P. B., Papaloizou, J. C. B., & Polnarev, A. G. 1999, MNRAS, 307, 79
Kauffmann, G., & Haehnelt, M. 2000, MNRAS, 311, 576
Kirhakos, S., Bahcall, J. N., Schneider, D. P., & Kristian, J. 1999, ApJ, 520, 67
Lehto, H., & Valtonen, M. J. 1996, ApJ, 460, 207
Liouville, J. 1837, J. Math. Pures Appl., 2, 16
Makino, J. 1997, ApJ, 478, 58
Makino, J., Fukushige, T., Okumura, S. K., & Ebisuzaki, T. 1993, PASJ, 45, 303
Merritt, D. 2000, in Galaxy Dynamics: From the Early Universe to the Present, ed. F. Combes, G. Mamon, & V. Charmandaris (San Francisco: ASP), 221
Quinlan, G. D. 1996, NewA, 1, 35
Quinlan, G. D., & Hernquist, L. 1997, NewA, 2, 533
Rajagopal, M., & Romani, R. W. 1995, ApJ, 446, 543
Taniguchi, Y., & Wada, K. 1996, ApJ, 469, 581
Trilling, D. E., Benz, W., Guillot, T., Lunine, J. I., Hubbard, W. B., & Burrows, A. 1998, ApJ, 500, 428
van der Marel, R. P. 1999, AJ, 117, 744
White, S. D. M. 1996, in Les Houches, Session LX, Cosmology and Large Scale Structure, ed. R. Schaeffer et al. (Amsterdam: North-Holland), 121