STAR FORMATION–REGULATED GROWTH OF BLACK HOLES IN PROTOGALACTIC SPHEROIDS

ANDREAS BURKERT
Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany; burkert@mpia-hd.mpg.de

AND

JOSEPH SILK
Department of Astrophysics, University of Oxford, Nuclear and Astrophysics Laboratory, Keble Road, Oxford OX1 3RH, UK

Received 2000 November 28; accepted 2001 May 10; published 2001 June 19

ABSTRACT

The observed relation between central black hole mass and spheroid velocity dispersion is interpreted in terms of a self-regulation model that incorporates a viscous Keplerian accretion disk to feed the black hole, embedded in a massive, self-gravitating star-forming disk that eventually populates the spheroid. The model leads to a constant ratio between black hole mass and spheroid mass that is equal to the inverse of the critical Reynolds number for the onset of turbulence in the accretion disk surrounding the central black hole. Applying the fundamental plane correlation for spheroids, we find that the black hole mass has a power-law dependence on the spheroid velocity dispersion with a slope in the range of 4–5. We interpret the larger scatter in the Magorrian relation with respect to the black hole mass–spheroid velocity dispersion relationship as a result of secular evolution of the spheroid that primarily affects its luminosity and to a much lesser extent its velocity dispersion.

Subject headings: black hole physics — galaxies: kinematics and dynamics — galaxies: nuclei

1. INTRODUCTION

The remarkable dependence of black hole mass on spheroid velocity dispersion, with low dispersion (<10%) about a mean power-law slope between 4 and 5 (Ferrarese & Merritt 2000; Merritt & Ferrarese 2001; Gebhardt et al. 2000a, 2000b), merits serious attention by theorists. Several scenarios have recently been proposed to interpret this relation for either value of the slope, but none are entirely convincing (Silk & Rees 1998; Ostriker 2000; Haehnelt & Kauffmann 2000; Adams, Graf, & Richstone 2001). Silk & Rees (1998) had predicted such a dependence in a model that appealed to feedback for quasar outflows on the protogalactic gas reservoir. They found the relation

\[ M_{\text{bh}} = \alpha \sigma_{\text{sph}}^2 f_{\text{cold}}, \]

where \( f_{\text{cold}} \) is the cold gas fraction and \( \alpha \) depends weakly on the presence of an accretion disk (Haehnelt, Natarajan, & Rees 1998) and inversely on the ratio of the kinetic outflow energy to Eddington luminosity. This study requires a very fast accretion phase to feed the central black hole. Recent results by Merrifield, Forbes, & Terlevich (2000) provide evidence for a relationship between black hole mass and bulge age and indicate that the black hole formation timescale is longer, of order 1 Gyr.

A flatter relation was derived from cold dark matter–dominated cosmologies by Haehnelt & Kauffmann (2000). They assumed that a fixed fraction of the cold gas supply that forms spheroid stars in a merger will feed the central black hole and discussed the importance of feedback. Without feedback their derived scaling relation is \( M_{\text{bh}} \propto \sigma_{\text{sph}}^2 \). However, incorporating feedback leads to arbitrarily steeper slopes, depending on the choice of the free parameters, a situation that is not very satisfactory. A relatively flat slope was postulated by Ostriker (2000), for a model in which the central black hole formed from self-interacting dark matter, which is now largely discredited (Gnedin & Ostriker 2001; Yoshida et al. 2000). Adams, Graff, & Richstone (2001) proposed a new scenario in which the central black hole grew during the protogalactic collapse phase by capturing stars on nearly radial orbits. Their model is, however, questionable for several reasons: Merrifield’s result requires black hole growth on a timescale longer than that of protogalactic collapse; star formation is unlikely to have been highly efficient during this phase, and finally they adopt a characteristic velocity dispersion \( \sigma \) to determine the specific angular momentum of the stars that neglects the fundamental plane relation between \( \sigma \) and scale radius.

2. A MODEL OF SELF-REGULATED BLACK HOLE GROWTH

We propose an alternative model for black hole growth in forming spheroids that is based on the commonly adopted merging scenario for the formation of spheroids. As first proposed by Toomre & Toomre (1972), spheroids, like elliptical galaxies, formed via major mergers of gas-rich progenitors. These disk galaxies most likely had already formed central black holes in their bulges that satisfied the Magorrian relation (Magorrian et al. 1998). However, the black hole that results from the coalescence of the progenitor black holes during a spiral-spiral merger would fall short of the Magorrian relation by a large factor. A large fraction of its mass must therefore have been accreted during or after the merger event, presumably as gas.

Major merger simulations (Mihos & Hernquist 1996) demonstrate that the gas rapidly settles, on a short dynamical timescale, into a central, self-gravitating disk. By that time, the progenitor black holes will also have merged in the center by dynamical friction (A. Burkert & R. Sunyaev 2001, in preparation). The inner disk has Keplerian rotation because the gravitational potential is dominated by the mass of the black hole. This region loses angular momentum by viscous drag, and the resulting gas inflow will feed the central black hole. At the same time, the outer self-gravitating disk part is gravitationally unstable to fragmentation and subsequent star formation.

Within the framework of our model, we can estimate the critical radius \( r_c \) that separates the two regions dominated by accretion and fragmentation and that is determined by the radius...
out to which the black hole dominates the gravitational potential,

\[ r_{\text{crit}} = \frac{GM_{\text{bh}}}{\sigma_{\text{sph}}^2}. \]

Here \( \sigma_{\text{sph}} \) is the characteristic velocity dispersion of the spheroid and \( M_{\text{bh}} \) is the black hole mass. At the critical radius, the Keplerian accretion disk rotates with a velocity \( v_{\text{esc}} \approx \sigma_{\text{sph}} \). Typical timescales for the gas in the whole inner disk to accrete onto the black hole are given by the viscous drag timescale \( t_{\text{vis}} = r_{\text{crit}}^2 / v \). For a viscosity prescription, we adopt the formulation \( v \approx R_{\text{sph}}^2 v_{\text{esc}} / r_{\text{esc}} \), where the critical Reynolds number for the onset of turbulence is \( R_{\text{cr}} \approx 100-1000 \) (Duschl, Strittmatter, & Biermann 2000). Adopting \( v_{\text{esc}} = \sigma_{\text{sph}} \), the viscous time can be written as

\[ t_{\text{vis}} = \frac{G M_{\text{bh}}}{\sigma_{\text{sph}}^3}. \]

that is, the viscous time is the product of the dynamical time of the spheroid times the Reynolds number.

The black hole will grow as long as there is a gas supply for the inner disk. This gas reservoir is replenished by viscous inflow from the outer disk into the accretion region and augmented by the increase in the inner disk radius due to the growth in the black hole mass. In the absence of any limiting factor, all the gas supply will eventually be accreted onto the black hole. In this case one would expect a large variance in the black hole mass with respect to the local velocity dispersion unless the initial gas fraction of the merger components was relatively fine-tuned. This seems unphysical on theoretical grounds, nor is any relation observed between gas fraction and velocity dispersion.

We propose that star formation in the outer disk provides the self-regulation that limits the mass of the central black hole. Black hole growth saturates because of the competition with star formation, which determines the gas fraction in the disk that is available for accretion. Let \( M_{\text{f}} \) be the total gas mass available for accretion and \( M_{\text{f}} \) be the mass in stars within the corresponding radius. We can identify the gas fraction \( \epsilon \) as \( \epsilon = M_{\text{f}} / M_{\text{f}} \), where \( M_{\text{f}} = M_{\text{f}} + M_{\text{s}} \) is the disk mass at a given radius \( r \), to demonstrate explicitly how black hole growth is limited by the gas supply. Black hole growth occurs only during the gas-rich phase of the protospheroid, that is, on the characteristic star formation timescale \( t_{\text{sf}} \).

Whatever gas is in the inner disk within \( r_{\text{esc}} \) will be accreted onto the black hole without significant star formation, since self-gravity is unimportant. The average growth rate of the black hole can then be estimated by

\[ M_{\text{bh}} = \epsilon M_{\text{f}} t_{\text{sf}}^{-1}. \]

where \( M_{\text{f}} \), \( \epsilon \), and \( t_{\text{sf}} \) are defined at \( r_{\text{esc}} \). The total gas fraction within \( r_{\text{esc}} \) will contribute to the black hole mass. We therefore set \( M_{\text{f}} = \epsilon M_{\text{f}} \) equal to the black hole mass \( M_{\text{bh}} \). This assumes the initial black hole mass to be negligible with regard to the accreted mass. Viscosity (eq. [3]) then limits the black hole growth rate to

\[ M_{\text{bh}} = \sigma_{\text{sph}}^3 R_{\text{sph}}^3 G^{-1} t_{\text{sf}}. \]

The growth continues until star formation exhausts the gas supply both by turning gas into stars and by heating the gas in the disk due to newly formed high-mass stars and dispersing it. As this occurs on the star formation timescale \( t_{\text{sf}} \), the final black hole mass will be

\[ M_{\text{bh}} = \sigma_{\text{sph}}^3 R_{\text{sph}}^3 G^{-1} t_{\text{sf}} = 1.9 \times 10^8 \left( \frac{\sigma_{\text{sph}}}{200 \text{ km s}^{-1}} \right)^3 \times \left( \frac{R_{\text{sph}}}{100} \right) \left( \frac{t_{\text{sf}}}{10^7 \text{ yr}} \right) M_{\odot}. \]

As a first rough estimate, note that star formation occurs in molecular gas on timescales of order \( t_{\text{sf}} \approx 10^7 \text{ yr} \) (Blitz & Shu 1980; Pringle, Allen, & Lubow 2001), leading to a relationship that fits the observed normalization of the \( (M_{\text{bh}}, \sigma_{\text{sph}}) \) relation for reasonable critical Reynolds number \( R_{\text{cr}} \approx 100 \) for the onset of turbulence. This timescale is also of order the local dynamical timescale of spheroids. For example, typical giant elliptical galaxies with effective radii \( r_e \approx 2 \text{ kpc} \) and velocity dispersions \( \sigma_{\text{sph}} \approx 200 \text{ km s}^{-1} \) have dynamical timescales \( t_{\text{dyn}} \approx \sigma_{\text{sph}} \approx 10^7 \text{ yr} \). Interestingly, low-mass bulges have velocity dispersions and effective radii that are a factor of 2 smaller, leading to the same value for \( t_{\text{dyn}} \). We can therefore write the star formation timescale as \( t_{\text{sf}} = \eta t_{\text{dyn}} = \eta \sigma_{\text{sph}}^2 \), where \( \eta \approx 1 \). Hence,

\[ M_{\text{bh}} = \eta \sigma_{\text{sph}}^3 R_{\text{sph}}^3 G^{-1} r_e = M_{\text{bh}} \eta \sigma_{\text{sph}}^3. \]

This provides a natural explanation of the Magorrian relation.

The ratio of black hole mass to spheroid stellar mass is given by a “universal” constant, \( \sim R_{\text{sph}}^3 \), which in principle should be derivable from fundamental theory and should be insensitive to galaxy parameters. It is remarkable that the ratio of two global masses might be regulated by a parameter that is determined by the microphysics of viscous transport of angular momentum. The mass ratio could therefore provide a direct measure for the typical Reynolds number \( R_{\text{cr}} \) in accretion disks around massive black holes.

The predicted dependence of \( M_{\text{bh}} \) on \( \sigma_{\text{sph}} \) and \( r_e \) follows the virial theorem expectation \( M_{\text{bh}} = r_e \sigma_{\text{sph}}^2 G^{-1} \) for the spheroid. This leads to a fundamental plane-like projection on spheroid parameters for the black hole mass, with a constant offset that is consistent with the Magorrian relation. Note that \( r_e \) depends on \( \sigma_{\text{sph}} \) and on surface brightness: \( r_e \propto \sigma_{\text{sph}}^2 \). This follows, for example, by combining the Faber-Jackson relation, \( M_{\text{bh}} \propto \sigma_{\text{sph}}^4 \), with the virial theorem. We infer that \( M_{\text{bh}} \propto \sigma_{\text{sph}}^4 \), which is in the observed range.

The crucial issue, however, is that of the remarkable reduction in dispersion in the relation between \( M_{\text{bh}} \) and \( \sigma_{\text{sph}} \) relative to that of the spheroid luminosity \( L_{\text{sph}} \) versus \( \sigma_{\text{sph}} \). We note that up to one-half the scatter in this fundamental plane projection may be attributed to population age differences (Forbes, Ponman, & Brown 1998) that affect primarily the mass-to-light ratios of spheroids. Another source of scatter that has also been discussed by Gebhardt et al. (2000a) is due to projection effects that lead to variations in the effective radius and velocity dispersion by a factor of order the typical ellipticity of the spheroid, up to a factor of 2 (A. Burkert & T. Naab 2001, in preparation). The inferred intrinsic dispersion in the dependence of bulge mass, and also of black hole mass, on velocity dispersion must therefore have been quite small compared with the observed spread. Note that projection effects would result in scatter in the Magorrian relation as well as in scatter in the \( M_{\text{bh}} - \sigma_{\text{sph}} \) relationship.

The following effect is likely to play an additional important role. Minor mergers add stars but should have little effect on
the central black hole mass. These mergers could occur after the initial protospheroid formation phase, which characterizes the formation of a central disk and the accretion of gas onto the central black hole but before the gas in the outer regions of the spheroid was completely consumed. The mergers induce a change in the star formation rate that is disproportionately larger than would be inferred directly from the amount of stars and gas added. This is because the tidal shocks stimulate the larger preexisting outer gas supply to form stars earlier at an accelerated pace. Indeed, the near-ultraviolet–optical color-magnitude relation for early-type cluster galaxies is incompatible with a monolithic scenario for star formation at high redshift (Ferreras & Silk 2001). An increased scatter is found in the color-magnitude relation at the faint end, resulting in a significant fraction of faint blue–early-type systems, implying that less massive galaxies undergo more recent episodes of star formation. Such episodes will produce scatter in the fundamental plane by affecting the galaxy luminosities and outer radii but will not add appreciably to the scatter in the \((M_{bh}, \sigma_{ph})\) correlation because the velocity dispersion in the inner regions is not changed.

There is a further effect to be explained, namely, the correlations between fundamental plane residuals (Kormendy 2001), specifically an anticorrelation between \(\Delta \sigma_{ph}\) and \(\Delta r\). We attribute this as being in part due to projection effects of nonaxisymmetric spheroids. In addition, late mass loss or mass infall may well occur. Substantial mass loss or infall would modify both the galaxy size and spheroid velocity dispersion. In the limit of an adiabatic response of the host galaxy, we predict that

\[
\frac{\Delta r}{r} = - \frac{\Delta \sigma}{\sigma} = -0.4 \Delta M_B, \tag{8}
\]

where \(M_B\) is the absolute blue magnitude. This equation is consistent with the observed anticorrelation between \(\Delta \sigma_{ph}\) and \(\Delta r\). The tightness of the \((M_{bh}, \sigma_{ph})\) correlation therefore means that late infall into or outflow from spheroids can be limited to about 10% of the current stellar mass. This complements a similar conclusion for disks, based on their observed thickness (Toth & Ostriker 1992).

Finally, let us introduce cosmology via merger-induced feeding of the accretion disk. If the accretion disk lifetime exceeds the characteristic time between mergers, the gas inflow is disrupted. For example, a binary black hole merger would sweep out the local environment. The merger, by forming a transient bar, will subsequently resupply cold gas to the central object causing quasar activity. Whatever the details, it seems reasonable to equate the quasar lifetime \(t_q = \gamma t_{edd}\) to the merger timescale \(t_{merger}\) thereby introducing a dependence on the cosmological model. Here \(\gamma\) (expected to be of order 0.1) is the ratio of \(t_q\) to the Eddington timescale \(t_{edd} \equiv 0.4\) Gyr. Consider hierarchical merging of primordial fluctuations described by a power spectrum \(\delta \rho / \rho \sim M^{-1/2}\). Let \(f_{cold}\) be the cold gas fraction of the total baryon content in the merging galaxy. Following White & Rees (1978), one can then write

\[
t_{merger} = \left( \frac{\sigma_{ph}}{\sigma_0} \right)^{3(3+n)/(1-n)} f_{cold}^{-4/(1-n)} t_0, \tag{9}
\]

where the normalization assumes that systems with velocity dispersion \(\sigma_i = 1000\alpha\) km s\(^{-1}\), with \(\alpha \approx 1\), are forming at the present epoch \(t_0\). The quasar lifetime is determined by the supply of cold gas to the circumquasar accretion disk. Following Haehnelt & Kauffmann (2000), we assume that a constant fraction of the cold gas supply constitutes the cold gas reservoir that feeds the quasar. It follows that we may rewrite the expression for the black hole mass by introducing \(t_q\), which we expect to be determined by the star formation timescale, since star formation competes for the gas supply, to generate

\[
M_{bh} = t_q \left( \frac{\sigma_{ph}}{\sigma_0} \right)^{3(3+n)/(1-n)} f_{cold}^{-4/(1-n)} t_0 \frac{\sigma_{ph}^3}{GR_{ce}}. \tag{10}
\]

The expected range for protogalaxies, \(-1.5 \leq n \leq -2.5\), again yields the range of 4–5 for the slope of the \(M_{bh}-\sigma_{ph}\) relation if \(t_q \approx t_{edd}\).

Note that the expected narrow dispersion in \(\sigma_{ph}\) remains since most of the dependence on \(\sigma_{ph}\) (the part that varies as \(\sigma_{ph}^3\)) is cosmology-independent. Note, moreover, that there is no reliance on the uncertainties inherent in feedback, other than in the dependence on cold gas fraction, which may be adopted from the cold gas fraction for the observed damped Ly\(\alpha\) clouds. This varies approximately as \((1+z)^{-3/2}\) for \(0 < z < 3\). We finally infer an approximately linear dependence of \(M_{bh}\) on age of the stellar population, presumed to have formed when the cold gas fraction was supplied by the last major merger in accordance with observations (Merrifield et al. 2000).

3. DISCUSSION

A simple model of star formation–regulated black hole growth seems to explain the observed dependence of black hole mass on spheroid mass and spheroid velocity dispersion, and especially the observed dispersions. The proposed model leads to several predictions that might merit further investigation. We find that the Magorrian relation depends on two parameters, the critical viscous Reynolds number and the timescale of star formation in units of the dynamical timescale. As \(\eta\) should be of order unity, we infer that black hole growth must be regulated by a turbulent accretion disk with a Reynolds number of a few hundred. In addition, the accretion and star formation timescales are closely coupled and in our interpretation are of order the quasar lifetime.

Our model predicts a tight relationship between black hole mass and the product of \(\sigma^2 \times r_\text{c}\), which can be tested observationally. The scatter of the Magorrian relation directly reflects the scatter in the fundamental plane parameters \(r_\text{c}\) and \(M_{bh}\), which we postulate to be enhanced by secular events after black hole and spheroid formation. Thus, the original scatter in the fundamental plane should be as small as that observed in the black hole mass–spheroid velocity dispersion relationship. This provides a possible observable prediction for high-redshift field galaxies and also for clusters where most of the merging may have been suppressed.

We wish to thank UCSC, UC Berkeley, IAP, Chez Papa, and Cafe Peret for hospitality and for providing a stimulating atmosphere. We thank the referee Oleg Gnedin for valuable comments.
REFERENCES

Adams, F. C., Graff, D. S., & Richstone, D. O. 2001, ApJ, 551, L31
Blitz, L., & Shu, F. H. 1980, ApJ, 238, 148
Duschl, W. J., Strittmatter, P. A., & Biermann, P. L. 2000, A&A, 357, 1123
Ferrarese, L., & Merritt, D. 2000, ApJ, 539, L9
Ferreras, I., & Silk, J. 2001, ApJ, in press (astro-ph/0008493)
Forbes, D. A., Ponman, T. J., & Brown, R. J. N. 1998, ApJ, 508, L43
Gebhardt, K., et al. 2000a, ApJ, 539, L13
———. 2000b, ApJ, 543, L5
Gnedin, O., & Ostriker, J. 2001, ApJ, submitted (astro-ph/0101436)
Haehnelt, M. G., & Kauffmann, G. 2000, MNRAS, 318, L35
Haehnelt, M. G., Natarajan, P., & Rees, M. 1998, MNRAS, 300, 817
Kormendy, J. 2001, in ASP Conf. Ser. 230, Galaxy Disks and Disk Galaxies, ed. J. G. Funes & E. M. Corsini (San Francisco: ASP), in press (astro-ph/0007401)

Merrifield, M. R., Forbes, D. A., & Terlevich, A. I. 2000, MNRAS, 313, 29
Merritt, D., & Ferrarese, L. 2001, MNRAS, 320, 30
Mihos, C., & Hernquist, L. 1996, ApJ, 464, 641
Ostriker, J. P. 2000, Phys. Rev. Lett., 84, 5258
Pringle, J. E., Allen, R. J., & Lubow, S. H. 2001, preprint
Silk, J., & Rees, M. 1998, A&A, 331, L1
Toomre, A., & Toomre, J. 1972, ApJ, 178, 623
Toth, G., & Ostriker, J. P. 1992, ApJ, 389, 5
White, S., & Rees, M. 1978, MNRAS, 183, 341
Yoshida, N., Springel, V., White, S., & Tormen, G. 2000, ApJ, 535, L103