

# Self-gravity and quasi-stellar object discs

Jeremy Goodman<sup>★</sup>

*Princeton University Observatory, Princeton, NJ 08544, USA*

Accepted 2002 October 28. Received 2002 October 22; in original form 2002 March 21

## ABSTRACT

The outer parts of standard steady-state accretion discs around quasi-stellar objects (QSOs) are prone to self-gravity, and they might be expected to fragment into stars rather than feed the central black hole. Possible solutions to this well-known problem are examined with an emphasis on general dynamic constraints. Irradiation by the QSO is insufficient for stability even if the outer disc is strongly warped. Marginal local gravitational instability enhances viscous transport but extends the stable regions only modestly. Compton cooling in the observed QSO radiation field rules out hot thick discs unless the local accretion rate is vastly super-Eddington. The formation of stars or stellar-mass black holes, and the release of energy in these objects by fusion or accretion, may help to stabilize the remaining gas in an otherwise standard disc. But at fixed mass accretion rate, the energy inputs required for stability increase with radius; beyond a parsec, they approach the total QSO luminosity and are probably unsustainable by stars. Magnetic torques from a wind or corona, and gravitational torques from bars or global spirals, may increase the accretion speed and reduce the density of the disc. But dynamical arguments suggest that the accretion speed is at most sonic, so that instability still sets in beyond about a parsec. Alternatively, the QSO could be fed by stellar collisions in a very dense stellar cluster, but the velocity dispersion would have to be much higher than observed in nearby galactic nuclei containing quiescent black holes. In view of these difficulties, we suggest that QSO discs do not extend beyond a thousand Schwarzschild radii or so. Then they must be frequently replenished with gas of small specific angular momentum.

**Key words:** accretion, accretion discs – gravitation – quasars: general.

## 1 INTRODUCTION

The outer parts of steady, geometrically thin, optically thick, viscously-driven accretion discs around quasi-stellar objects (QSOs) are predicted to be self-gravitating (e.g. Shlosman & Begelman 1987). This is due to the high mass accretion rate required to feed a QSO and the relatively shallow gravitational potential at large radius. Under standard assumptions, the Toomre stability parameter  $Q$  falls below unity at  $r \gtrsim 10^{-2}$  pc  $\sim 10^3 R_S$  (Section 2). It seems unlikely that a strongly self-gravitating disc can persist much longer than an orbital time.

There are, however, several reasons to suspect that QSO discs are larger than  $10^{-2}$  pc. The mass of the disc within this radius is much less than that of the black hole. Hence, in order to increase the mass of the hole substantially, or equivalently, to fuel QSO activity for a Salpeter (1964) time

$$t_{\text{Sal}} \approx 5 \times 10^7 \varepsilon_{0.1}^{-1} \text{ yr}, \quad (1)$$

a small disc would have to be replenished. The spectral energy distribution of typical QSOs shows a strong broad peak in the rest-

frame infrared, which is interpreted as the reprocessing of light from the central source by an extended warped disc (Sanders et al. 1989). Nearby Seyferts and other active galaxies, thought to be low-luminosity analogues of QSOs, sometimes have resolvable nuclear discs on scales  $\sim 10^2$  pc. Most spectacularly, very long baseline interferometry (VLBI) observations of maser emission in NGC 4258 and 1068 indicate discs on parsec scales and have been used to measure the mass of their central black holes (Nakai, Inoue & Miyoshi 1993; Greenhill et al. 1995; Greenhill & Gwinn 1997). It is usually difficult to know from observation whether the outer disc is as dense as would be required to support the central luminosity on the assumption of a radially constant accretion rate.

At least on larger galactic scales, it is believed that strong self-gravity leads to star formation (Martin & Kennicutt 2001). If the same is true of QSO discs, there is a danger that most of the gas would form stars, leaving little to fuel the QSO.

Several theoretical attempts have been made to modify the standard  $\alpha$ -disc model so as to extend gaseous QSO discs self-consistently into the self-gravitating regime. The options include: driving accretion with global bars or density waves so as to increase the radial velocity and lower the surface density needed to sustain

<sup>★</sup>E-mail: jeremy@astro.princeton.edu

**Table 1.** Accretion solutions and their limitations.

Type	Where discussed	Limitations
Standard steady viscous disc	Section 2, Appendix	$Q < 1$ at $r \gtrsim 0.02$ pc
As above but warped	Section 2.3	$Q < 1$ at $r \gtrsim 0.02$ pc
Unsteady viscous disc	Section 2.2	$\bar{r}_{Q=1}$ increases $\lesssim \times (t/t_{\text{Sal}})^{8/27} \sim 4$
Optically thin ADAF	Section 4.1	Inconsistent at QSO luminosity
Hot optically-thick disc or ADAF	Section 2.1	$\dot{M} \gg \dot{M}_{\text{E}}$ at $r \gtrsim 10^3 R_{\text{S}}$
Wind-driven disc	Section 4.3	Accretion subsonic, hence $r_{Q=1} \lesssim 1$ pc
Disc w. magnetized corona	Section 4.4	Accretion subsonic, hence $r_{Q=1} \lesssim 1$ pc
Disc w. bar or global spiral	Section 4.5	Accretion subsonic, hence $r_{Q=1} \lesssim 1$ pc
Clumpy disc	Section 4.6	Clumps collapse unless heated from within
Constant- $Q$ , star-forming disc	Section 3	$r_{\text{max}} \lesssim 1$ pc if heated by fusion
Collisional star cluster	Section 4.2	Requires $\sigma \gtrsim 700$ km s $^{-1}$
Disc truncated at $\lesssim 10^3 R_{\text{S}}$	Section 5	$M_{\text{disc}} \ll M_{\text{bh}}$ ; needs infall w. $j_{\text{gas}} \lesssim 70$ km s $^{-1}$ pc

a given mass accretion rate (Shlosman & Begelman 1989); heating or mechanically stirring the disc with embedded stars or black holes so as to raise the temperature and lower the density of the gas (Collin & Zahn 1999a,b); or allowing the disc to fragment into colliding gas clumps whose epicyclic motions stabilize it against self-gravity, at least on scales larger than the clumps themselves (Kumar 1999).

Our purpose is to re-examine self-gravity in QSO accretion discs with an emphasis on dynamical and energetic constraints. These constraints are most severe for massive and luminous systems, so our interest is in black-hole masses  $M \gtrsim 10^8 M_{\odot}$  and luminosities close to the Eddington limit. Because many possibilities must be considered, we use simple methods and often sacrifice mathematical precision to simplicity and generality. Vertical disc structure is represented by algebraic rather than differential relations, permitting easy derivation of scaling laws. In general, however, we are disappointed in our original hope of finding an hospitable theoretical environment for extended discs.

The plan of the paper is as follows. In Section 2 we review self-gravity in steady  $\alpha$  discs, including irradiation from the central source (Section 2.3), and enhancements to  $\alpha$  by local gravitationally-driven turbulence in a moderately self-gravitating disc (Section 2.4). In Section 3 we consider discs that are stabilized by additional heating beyond that due to the dissipation of orbital energy, but in which angular momentum continues to be transported by an  $\alpha$  viscosity. In contrast to the usual situation, the total energy released per unit mass accreted is strongly dependent on the outer radius of the disc (Section 3). In Section 4 we briefly explore various alternatives to viscous thin-disc accretion. Schemes that involve a thin disc but invoke faster-than-viscous angular-momentum transport include global spiral waves (Section 4.5) and magnetized disc winds (Section 4.3). In every such case, we argue that the accretion velocity is bounded by the sound speed. Other alternatives explored in Section 4 are less like thin discs: quasi-spherical flows (Section 4.1), collisional star clusters (Section 4.2) and clumpy discs (Section 4.6).

All of these alternatives to the standard  $\alpha$  disc face severe theoretical difficulties, or else seem unlikely to permit a centrifugally supported accretion flow beyond  $\sim 0.1$  pc. The situation is summarized in Table 1. We conclude in Section 5 that the gas probably does not circularize beyond this radius and must be supplied to the galactic nucleus with low specific angular momentum.

## 2 STEADY $\alpha$ DISCS

To begin, we review some salient features of a standard theoretical accretion disc: one that is steady, optically thick, and geometrically thin (Pringle 1981). The mass accretion rate ( $\dot{M}$ ) is constant with radius and time. Angular momentum is carried inward by the accreting gas at the same rate that it is transferred outward by viscosity, and the radiation of energy from the surfaces of the disc is balanced by viscous dissipation.

Some of the assumptions above are relaxed in later sections. The following assumptions, however, apply throughout this paper. The rotation curve is dominated by a Newtonian point mass  $M$ , as relativistic corrections are important only at small radii where self-gravity is negligible. Elsewhere, though self-gravity may be important for local stability, the disc is thin and its mass  $< M$  so that the rotation curve is not much affected. We do not consider radii so large that the potential of the stars becomes important ( $r \gg 1$  pc). Except where otherwise stated (Section 4.3), magnetic pressure contributes little to the thickness of the disc even if magnetic stresses dominate angular-momentum transport, as indicated by numerical simulations of magnetorotational instability (MRI; Balbus & Hawley 1998). Contrary suggestions, however, do exist (Pariev, Blackman & Boldyrev 2002).

Under these assumptions,

$$\dot{M} = 3\pi\nu\Sigma = 3\pi\alpha\beta^b c_s^2 \Omega^{-1} \Sigma, \quad (2)$$

where  $\nu$  is the effective viscosity and has been expressed in terms of the isothermal sound speed at the disc midplane,  $c_s = \sqrt{p/\rho}$ , and the orbital angular velocity,  $\Omega = (GM/r^3)^{1/2}$ , and the dimensionless (Shakura & Sunyaev 1973) viscosity parameter  $\alpha$ .  $\Sigma$  is the surface mass density, and  $\beta \equiv p_{\text{gas}}/p = p_{\text{gas}}/(p_{\text{gas}} + p_{\text{rad}})$  is the ratio of gas pressure to total pressure at the midplane. The mechanism likely to be responsible for the viscosity of most discs is magnetorotationally-driven turbulence, for which simulations indicate  $\alpha = 10^{-3}$ – $10^{-1}$  (Balbus & Hawley 1998). As discussed below, where the disc is self-gravitating,  $\alpha$  may be as large as  $\sim 0.3$ . Therefore, we often write  $\alpha_{0.01} \equiv 10^2\alpha$  or  $\alpha_{0.3} \equiv \alpha/0.3$ . The parameter  $b$  is a switch that determines whether the viscosity is proportional to gas pressure ( $b = 1$ ) or total pressure ( $b = 0$ ). In the latter case, radiation-pressure-dominated regions of  $\alpha$  discs are viscously unstable (Lightman & Eardley 1974), but it is possible that  $b = 0$  in an average sense. Although the average surface density  $\bar{\Sigma}$  would be lower and the

corresponding  $\bar{Q}$  higher for  $b = 0$  than for  $b = 1$ , the viscous instability is expected to produce overdense rings in which these trends may well be reversed.

Because viscous dissipation draws upon the orbital energy of the gas and is assumed to be balanced by radiation, the local effective temperature is

$$\begin{aligned} T_{\text{eff}} &= \left( \frac{3}{8\pi\sigma} \frac{GM\dot{M}}{r^3} \right)^{1/4} \\ &\approx 2.9 \times 10^3 (M_8 \dot{M}_\odot)^{1/4} (r/10^{-2} \text{ pc})^{-3/4} \text{ K}, \\ &\approx 6.2 \times 10^5 \left( \frac{l_E}{\varepsilon_{0.1} M_8} \right)^{1/4} \left( \frac{r}{R_S} \right)^{-3/4} \text{ K}. \end{aligned} \quad (3)$$

We have introduced the abbreviations  $M_8 \equiv M/(10^8 M_\odot)$ ,  $\dot{M}_\odot \equiv \dot{M}/(1 M_\odot \text{ yr}^{-1})$ , dimensionless luminosity  $l_E \equiv L/L_E$ , radiative efficiency  $\varepsilon \equiv L/\dot{M}c^2 \equiv 0.1\varepsilon_{0.1}$ , and Schwarzschild radius  $R_S = 2GM/c^2 \approx 10^{-5} M_8 \text{ pc}$ . The mass accretion rate can then be written as

$$\dot{M} = \frac{4\pi GM}{\kappa_{\text{e.s.}} c} \frac{l_E}{\varepsilon} \approx 2.2 \varepsilon_{0.1}^{-1} l_E M_8 \dot{M}_\odot, \quad (4)$$

where  $\kappa_{\text{e.s.}} \approx 0.4 \text{ cm}^2 \text{ g}^{-1}$  is the electron-scattering opacity.

If vertical transport of heat is by radiative diffusion, then the midplane and surface temperatures are related approximately by

$$T^4 \approx \frac{\kappa \Sigma}{2} T_{\text{eff}}^4. \quad (5)$$

The numerical factor is somewhat arbitrary, as it depends upon the vertical dependence of the heating function; it ought to be  $3/8$  if the heating rate per unit mass and the opacity are constant. In fact, simulations of magnetorotational turbulence indicate that the vertical scale height of the turbulent dissipation is larger than that of the gas density (Miller & Stone 2000), so that the factor should be even smaller. Hence equation (5) probably overestimates the midplane temperature. Other things being equal, this would cause self-gravity to be underestimated, but we assume that the error is not significant.

## 2.1 Onset of self-gravity

The Toomre stability parameter for Keplerian rotation is

$$Q = \frac{c_s \Omega}{\pi G \Sigma} \approx \frac{\Omega^2}{2\pi G \rho}, \quad (6)$$

and local gravitational instability occurs where  $Q < 1$ . We have taken  $\Sigma = 2h\rho$  with the disc half-thickness  $h = c_s/\Omega$ . Combining equation (6) with equation (2), we obtain the simple relation

$$GMQ = 3\alpha\beta^b c_s^3. \quad (7)$$

For a flat rotation curve,  $v_{\text{circ}} = r\Omega = \text{constant}$ , the numerical factor 3 is replaced by  $2\sqrt{2}$ .

To see why self-gravity is inevitable, consider the radiation-pressure-dominated case,  $\beta \ll 1$ , so that the isothermal sound speed  $c_s^2 = 4\sigma T^4/3c\rho$ . From equations (3) and (5),  $T^4 = 3\kappa \Sigma \Omega^2 \dot{M}/16\pi\sigma$ . Eliminating  $T$  between these relations and using  $\Sigma = 2c_s\rho/\Omega$  leads to

$$c_s = \frac{\kappa \Omega \dot{M}}{2\pi c} = \frac{l_E \kappa}{\varepsilon \kappa_{\text{e.s.}}} \Omega R_S \quad (\beta \ll 1), \quad (8)$$

whence  $h = (l_E/\varepsilon)R_S = \text{constant}$  for  $\kappa = \kappa_{\text{e.s.}}$ . Using this for  $c_s$  and equation (4) for  $\dot{M}$  in equation (7) leads to

$$Q = \frac{3\alpha\beta^b}{8\pi\sqrt{2}} \left( \frac{l_E}{\varepsilon} \right)^2 \hat{\kappa}^3 \left( \frac{\kappa_{\text{e.s.}} c^4}{G^2 M} \right) \left( \frac{R_S}{r} \right)^{9/2} \quad (\beta \ll 1), \quad (9)$$

in which  $\hat{\kappa} \equiv \kappa/\kappa_{\text{e.s.}}$ . Hence discs around more massive black holes are more prone to self-gravity. For  $b = 0$ , there is a characteristic mass above which an Eddington-limited disc would be self-gravitating even at its inner edge. This mass is enormous ( $\sim 10^{19} M_\odot$  for  $\alpha_{0.01} = \varepsilon_{0.1} = 1$ ), but even for realistic black-hole masses,  $Q < 1$  at radii greater than

$$r_{Q=1} \approx 2.1 \times 10^3 \left( \frac{\alpha_{0.3} \beta^b l_E^2 \hat{\kappa}^3}{\varepsilon_{0.1}^2 M_8} \right)^{2/9} R_S \quad (\beta \ll 1). \quad (10)$$

Notice that  $\alpha$  has been scaled to 0.3 rather than  $10^{-2}$  because of the expected enhancement by gravitational turbulence (Section 2.4).

In the appendix, we briefly derive the radial run of  $\Sigma$  and  $T$  that follow from the standard assumptions laid out in this section. From these we derive expressions for  $\beta(r)$  (equation A3) and  $Q(r)$  (equation A4).  $\beta$  increases monotonically with radius, and  $Q$  decreases. For typical parameters,  $p_{\text{gas}} = p_{\text{rad}}$  at about the same radius where  $Q = 1$ ,  $r \sim 10^3 R_S \sim 10^{-2} \text{ pc}$ . This appears to be a coincidence, as the two radii  $r_{Q=1}$  and  $r_{\beta=1}$  depend differently on the accretion parameters, especially  $\alpha$  and  $\dot{M}$ . We generally presume  $\dot{M} \sim L_E/c^2$ . But if the disc were not in a steady state or if there were a massive wind, then the accretion rate at small radii, where most of the QSO luminosity presumably originates, could be very different from  $\dot{M}$  at large radii where self-gravity is problematic. Therefore, it is interesting to calculate the local value of  $\dot{M}$  that would yield  $Q = 1$  at each radius. We begin by expressing  $\beta$  in terms of  $Q$  using equations (4), (6) and (7):

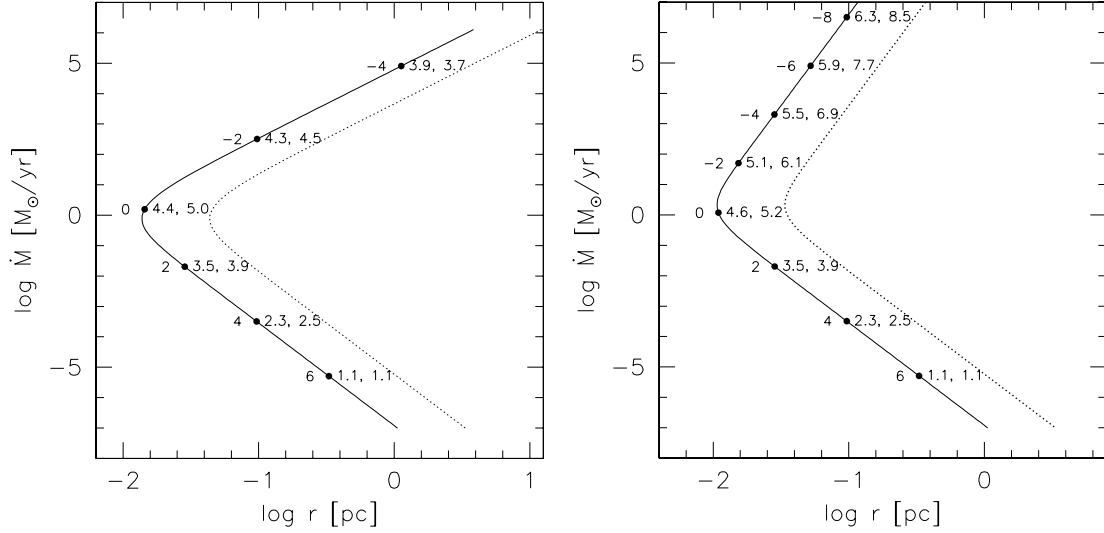
$$\begin{aligned} \frac{\beta^{1-(b/2)}}{(1-\beta)^{1/4}} &= 2^{-5/2} \left( \frac{3}{\pi} \right)^{3/4} \left( \frac{\varepsilon \alpha}{l_E} \right)^{1/2} Q^{-3/4} \left( \frac{k_B^4 c^9 \kappa_{\text{e.s.}}^2}{m^4 G^7 \sigma} \right)^{1/4} \\ &\times \left( \frac{r}{R_S} \right)^{-3/4} M^{-1} \\ &\approx 0.39 \left( \frac{\varepsilon_{0.1} \alpha_{0.3}}{l_E} \right)^{1/2} Q^{-3/4} \left( \frac{r}{2100 R_S} \right)^{-3/4} M_8^{-1}. \end{aligned} \quad (11)$$

Unlike equation (A3), this equation does not derive from equations (3) and (5); it does not assume a balance between radiative losses and viscous dissipation. The opacity  $\kappa_{\text{e.s.}}$  enters equation (11) only because it is used in the definition of the Eddington ratio  $l_E$ . Equation (11) remains true regardless of the heat source, whereas equation (A3) is specific to viscous dissipation.

Eliminating  $\beta$  between these two equations, we obtain loci of constant  $Q$  in the  $(r, \dot{M})$  plane, as shown in Fig. 1. These have been drawn for constant opacity  $\kappa = \kappa_{\text{e.s.}}$  regardless of temperature, but, because of the very strong dependence of  $\dot{M}$  on  $r$ , we expect that a more realistic opacity law would yield similar results. The unstable region where  $Q < 1$  is the interior of the wedge. Thus, below some radius depending on  $\alpha$  and  $M$ , there are no gravitationally unstable solutions for any  $\dot{M}$ . Beyond this critical radius, there are two marginally stable solutions, which diverge very quickly from one another in  $\dot{M}$ . The upper branch is very strongly radiation-pressure dominated ( $\beta \ll 1$ ) and very optically thick ( $\kappa \Sigma \gg 1$ ), so that it bears a resemblance to the supercritical solutions studied by Begelman & Meier (1982). Their solutions, however, besides assuming constant  $\dot{M}$ , are quasi-spherical and advection dominated out to the ‘trapping radius’

$$r_{\text{tr}} = \frac{\dot{M} c^2 R_S}{L} \frac{1}{2},$$

whereas those displayed in Fig. 1 are calculated from the thin-disc equations assuming local balance between radiative losses and viscous heating. The relative thickness  $h/r$  increases with  $r$  on the



**Figure 1.** Loci of constant  $Q$  for black-hole masses  $M = 10^8 M_\odot$  (solid) and  $M = 10^{9.5} M_\odot$  (dashed),  $\alpha = 0.3$ , and  $\kappa = 0.4 \text{ cm}^2 \text{ g}^{-1}$ . Left panel:  $b = 0$ , i.e. viscosity proportional to total pressure. Right panel:  $b = 1$ ,  $\nu \propto p_{\text{gas}}$ . Filled circles are labelled (left to right) with  $\log(p_{\text{gas}}/p_{\text{rad}})$ ,  $\log T$  (K),  $\log \Sigma$  ( $\text{g cm}^{-2}$ ).

upper branch. For  $b = 0$  and  $M_8 = 1$ ,  $h/r = 0.1$  at 1.2 pc and  $h/r \rightarrow 1$  at 7.3 pc; for  $b = 1$ ,  $h/r = 1$  is already reached at 0.11 pc. Thus, the vertical limit of both plots has been set approximately where the thin-disc assumption breaks down.

## 2.2 Time dependence

Steady-state accretion is mathematically convenient but the evidence for it is scant, because no QSO has been monitored for a significant fraction of its growth time (equation 1). The observed luminosity constrains  $\dot{M}$  at small  $r$ , but  $\dot{M}$  at large  $r$  could be different. Suppose the black hole shines at  $L \sim L_E$  with a duty cycle  $f \ll 1$ , and that it scarcely accretes at other times. A steady inflow rate  $\sim f L_E / \epsilon c^2$  at large radii is sufficient to support this behaviour, if  $\epsilon$  is the radiative efficiency during the bright phase. Assuming that the black hole gains its mass by accretion,  $f \gtrsim t_{\text{Sal}}/t \gtrsim 10^{-2}$ , where  $t$  is the age of the universe when the QSO shone, estimated for  $z_{\text{QSO}} \gtrsim 2$  and standard cosmological parameters. Thus, the accretion rate in the outer disc can typically be reduced by at most two orders of magnitude compared to the peak rate (equation 4). From Fig. 1, however, we see that a reduction of at least four orders of magnitude is needed to stabilize a disc at 0.1 pc if  $M_8 = 1$ .

From this we conclude that a variable accretion rate will not solve the problem of self-gravity in the outer disc.

## 2.3 Irradiation

Although flared or warped outer regions can be warmed by irradiation from the inner parts (where most of the total disc luminosity originates), we now show that this effect is not enough to stabilize the disc at  $r \sim r_{Q=1}$ .

To do so, irradiation must raise the midplane temperature  $T$ , not just the surface temperature  $T_{\text{eff}}$ , substantially. As a result, the disc will be nearly isothermal from surface to midplane. The vertical pressure gradient, and therefore the disc thickness, will be dominated by the gas even if  $\beta \ll 1$ : that is,  $h^2 \approx p_{\text{gas}}/\rho\Omega^2$ . If the effective viscosity derives from local magnetorotational turbulence, then it probably scales more directly with thickness than with sound speed because MRIs do not require compression. Then in all relevant respects the disc behaves as if  $\beta = 1$  and  $c_s^2 \rightarrow p_{\text{gas}}/\rho = k_B T/m$ .

From equation (7), it then follows that the minimum temperature that the irradiation must provide to ensure gravitational stability is

$$T_{Q=\beta=1} = \frac{m}{k_B} \left( \frac{GM}{3\alpha} \right)^{2/3} \approx 5.6 \times 10^4 \left( \frac{L_E M_8}{\alpha_{0.3} \epsilon_{0.1}} \right)^{2/3} \text{ K.} \quad (12)$$

To be conservative, we have scaled to  $\alpha = 0.3$ ; for  $\alpha = 0.01$ , this temperature would be an order of magnitude larger. On the other hand, the temperature of the disc in equilibrium with radiation from the central source is

$$T_{\text{eq}} = \left( \frac{c^5 l_E \cos \theta}{4\sigma \kappa_{e,s} GM} \right)^{1/4} \left( \frac{r}{R_S} \right)^{-1/2} \\ \approx 3.8 \times 10^5 \left( \frac{l_E \cos \theta}{M_8} \right)^{1/4} \left( \frac{r}{R_S} \right)^{-1/2} \text{ K,} \quad (13)$$

where  $\theta$  is the angle between the local normal to the disc and the radial rays. For a flat disc,  $\cos \theta \sim \max(h, R_S)/r$  at  $r \gg R_S$ , but even for a severely warped disc with  $\cos \theta \sim 1$ ,  $T_{\text{eq}} \ll T_{Q=\beta=1}$  at  $r/R_S \gg 44(\alpha_{0.3} \epsilon_{0.1})^{4/3} l_E^{-5/6} M_8^{-11/6}$ . Because this last inequality certainly holds at all  $r \geq r_{Q=1}$ , we conclude that irradiation is not important for the self-gravity of the disc.

## 2.4 Local gravitational turbulence

It has occasionally been proposed that a partially self-gravitating disc may transport angular momentum by spiral waves (e.g. Cameron 1978; Paczyński 1978; Larson 1984; Lin & Pringle 1987). Others have suggested that gravitational instabilities are intrinsically global and therefore not reducible to a local viscosity (Balbus & Papaloizou 1999). It may be that either opinion can be correct depending upon the ratio of disc thickness to radius, because the most unstable wavelength in a  $Q = 1$  disc is  $\sim h$ . We consider local gravitational turbulence here, as it can be accommodated within the  $\alpha$  model; a (much larger) upper bound on angular momentum transport by global spirals is given in Section 4.5.

By careful two-dimensional (2D) simulations, Gammie (2001) finds that gravitationally-driven turbulence can be local and can support  $\alpha$  approaching unity; in fact, in the absence of any other viscosity mechanism, his discs self-regulate themselves so that

$$\alpha_{\text{grav}} \approx \frac{1}{(\gamma_{2\text{D}} - 1)\Omega t_{\text{th}}}, \quad (14)$$

where  $t_{\text{th}} \equiv \Sigma k_{\text{B}} T / \sigma T_{\text{eff}}^4$  is the local thermal time-scale, and  $\gamma_{2\text{D}}$  is the 2D adiabatic index:

$$\gamma_{2\text{D}} = \frac{\partial}{\partial \log \Sigma} \log \left[ \int_{-\infty}^{\infty} p(r, \theta, z) dz \right].$$

[Our equation (14) differs from that of Gammie by a factor  $\gamma_{2\text{D}}$  because we define  $\alpha$  in terms of isothermal rather than adiabatic sound speed.] Gammie finds that self-regulation fails and the disc fragments if  $\Omega t_{\text{th}} \lesssim 3$  for  $\gamma_{2\text{D}} = 2$ . The latter result is not entirely surprising; in the absence of pressure support, the natural time-scale for collapse is  $t_{\text{dyn}} = \Omega^{-1}$ , and we would not expect collapse to be prevented by thermal energy unless  $t_{\text{th}} > t_{\text{dyn}}$ . The converse statement, that fragmentation can be indefinitely postponed if  $\Omega t_{\text{th}} > 3$ , is not at all obvious but is supported by Gammie's simulations.

We are not aware of any direct numerical simulations of discs that are unstable to both gravitational and magnetorotational modes at the same radii, as is likely to be the case for QSO discs (Menou & Quataert 2001). It would be interesting to explore this, as there might be a synergy between the two instabilities.

QSO discs are expected to be very thin in the regions of interest to us, so that Gammie's results may be applicable. Equation (8) implies  $h = l_{\text{E}} R_{\text{S}} / 6\epsilon$  in a radiation-pressure dominated disc, and hence  $h/r \lesssim 0.003$  at the innermost radius where  $Q \sim 1$  (equation 10). We expect that, from this radius outward,  $\alpha$  will rise smoothly from the value supported by magnetorotational turbulence (perhaps  $\alpha_{\text{m.h.d.}} \sim 10^{-2}$ ) up to the maximum allowed by gravitational turbulence,  $\alpha_{\text{grav,max}} \sim 0.3$ , in such a way that  $Q \approx \text{constant}$ . It follows from equation (9) that the ratio of outer to inner radii of this region is rather modest:  $\sim (\alpha_{\text{grav,max}} / \alpha_{\text{m.h.d.}})^{2/9} \sim 2$ . At still larger radii, additional sources of energy are required in order to prevent catastrophic fragmentation.

### 3 CONSTANT- $Q$ DISCS

It is not a new idea that accretion discs may regulate themselves at  $Q \approx 1$ . The heating required to offset radiative cooling and thereby maintain  $Q$  is usually supposed to come at the expense of the orbital energy of the gas (e.g. Paczyński 1978; Gammie 2001). In that case, gravitational instabilities and turbulence can be subsumed into the  $\alpha$  parameter, at least to the extent that they act locally (Section 2.4). Bertin (1997) and Bertin & Lodato (1999) have suggested replacing the energy equation (3) with the constraint of constant  $Q$  (equation 6), while retaining  $\alpha$  to describe angular-momentum transport (equation 7). In this section we adopt that approach and work out some consequences for the structure of the disc, with due attention to the effects of radiation pressure, which Bertin et al. seem to have neglected. It begs the question where the energy required to heat the disc may come from. Bertin & Lodato (1999) and especially Lodato & Bertin (2001) appear to suggest that the energy might ultimately derive from the gas orbits but yet be independent of local angular-momentum transport. We do not understand how this is possible. Instead, we attribute any surplus of the local radiative cooling above viscous dissipation to thermonuclear sources in stars, or perhaps to accretion on to small black holes that these stars evolve into. In Section 3.2 we calculate this surplus in terms of  $\dot{M}$  and the outer radius of the disc, and compare it to what might be available by converting a significant fraction of the disc gas into stars.

Although they may seem artificial, these assumptions are probably appropriate for the discs of spiral galaxies, and for the local interstellar medium (ISM) in particular. The local Galactic magnetic field is consistent with simulations of magnetorotational turbulence: namely, a magnetic energy density somewhat less than the thermal pressure, a predominantly toroidal orientation, and fluctuations comparable to the mean (Brandenburg et al. 1995). It is plausible therefore that there is a non-zero average magnetic stress  $-\langle B_r B_\phi \rangle / 4\pi = \alpha p_{\text{gas}}$  that systematically transfers angular momentum outward (Sellwood & Balbus 1999). Taking  $p_{\text{gas}}/k_{\text{B}} \approx 2000 \text{ cm}^{-3} \text{ K}^{-1}$ ,  $\rho \approx 0.3 m_{\text{H}} \text{ cm}^{-3}$  (Spitzer 1978), and circular velocity  $V_0 \approx 200 \text{ km s}^{-1}$  (Binney & Merrifield 1998), the implied radial drift velocity  $v_r \approx -\alpha p_{\text{gas}} / \rho V_0 \approx -0.3 \alpha \text{ km s}^{-1}$  is small enough to have escaped detection. Perhaps coincidentally, equation (7) predicts  $\dot{M} \approx 3 \times 10^{-2} \alpha_{0.1} Q^{-1} \dot{M}_{\odot} \text{ yr}^{-1}$ , about half the Eddington rate for the Galaxy's  $2.5 \times 10^6 M_{\odot}$  central black hole. The implied viscous heating rate  $\alpha p_{\text{gas}} \Omega_0 \approx 2 \times 10^{-29} \alpha_{0.1} \text{ erg cm}^{-3}$  is negligible compared to the inferred radiative cooling rate of the gas,  $\sim 2 \times 10^{-26} \text{ erg cm}^{-3}$  (Spitzer 1978). Presumably, the temperature of the ISM is maintained by stars. An important difference between the local ISM and QSO discs is that the former is very optically thin, especially to absorption, which means that the energy input from stars is inefficiently radiated. We estimate below that constant- $Q$  QSO discs remain optically thick out to at least 0.1 pc. A more careful treatment will be given in a later paper using realistic opacity tables (Sirko & Goodman 2003).

#### 3.1 Density and temperature

The midplane density in a constant- $Q$  disc follows from equation (6)

$$\rho = \frac{M}{2\pi Q r^3} = 1.2 M_8^{-2} Q^{-1} \left( \frac{R_{\text{S}}}{r} \right)^3 \text{ g cm}^{-3}, \quad (15)$$

so that the density at  $r_{Q=1}$  (equation 10) is  $\sim 10^{-8} M_8^{-4/3} \text{ g cm}^{-3}$ . The ratio  $\beta / (1 - \beta) = 3c\rho / 4\sigma T^3$  is determined by equation (11), which shows that  $\beta \ll 1$  at  $r \gg 10 R_{\text{S}}$  if  $Q \sim O(1)$ . The temperature itself is

$$\begin{aligned} T &= 2^{-7/6} \left( \frac{3}{\pi} \right)^{1/12} \left( \frac{l_{\text{E}}}{\alpha \epsilon Q^{1/2}} \right)^{1/6} c^{19/12} G^{-5/12} \sigma^{-1/4} \kappa_{\text{e.s.}}^{-1/6} M^{-1/3} \\ &\quad \times \left( \frac{R_{\text{S}}}{r} \right)^{3/4} \\ &\approx 6.9 \times 10^6 \left( \frac{l_{\text{E}}}{\alpha_{0.3} \epsilon_{0.1} Q^{1/2}} \right)^{1/6} M_8^{-1/3} \left( \frac{R_{\text{S}}}{r} \right)^{3/4} \text{ K} \quad \text{if } b = 0; \end{aligned} \quad (16)$$

$$\begin{aligned} T &= 2^{-1/3} \left( \frac{\pi}{3} \right)^{1/6} \left( \frac{l_{\text{E}} Q^{1/2}}{\alpha \epsilon} \right)^{1/3} c^{5/6} G^{1/6} \sigma^{-1/6} \kappa_{\text{e.s.}}^{-1/3} (k_{\text{B}}/m)^{-1/3} \\ &\quad \times \left( \frac{R_{\text{S}}}{r} \right)^{1/2} \\ &\approx 1.4 \times 10^6 \left( \frac{l_{\text{E}} Q^{1/2}}{\alpha_{0.3} \epsilon_{0.1}} \right)^{1/3} \left( \frac{R_{\text{S}}}{r} \right)^{1/2} \text{ K} \quad \text{if } b = 1. \end{aligned} \quad (17)$$

The surface density is

$$\begin{aligned}\Sigma &= \left( \frac{l_E c^8}{12\pi^2 \sqrt{2} \varepsilon \alpha Q^2 G^4 \kappa_{e.s.} M^2} \right)^{1/3} \left( \frac{R_S}{r} \right)^{3/2} \\ &\approx 7.4 \times 10^9 \left( \frac{l_E}{\alpha_{0.3} \varepsilon_{0.1} Q^2} \right)^{1/3} M_8^{-2/3} \\ &\quad \times \left( \frac{R_S}{r} \right)^{3/2} \text{ g cm}^{-2} \quad \text{if } b = 0;\end{aligned}\quad (18)$$

$$\begin{aligned}\Sigma &= \frac{2}{3} \left( \frac{2l_E m}{\varepsilon \alpha k_B \kappa_{e.s.}} \right)^{2/3} \left( \frac{3 \sigma c^7}{\pi Q G} \right)^{1/6} \left( \frac{R_S}{r} \right) \\ &\approx 5.4 \times 10^8 Q^{-1/6} \left( \frac{l_E}{\alpha_{0.3} \varepsilon_{0.1}} \right)^{2/3} \left( \frac{R_S}{r} \right) \text{ g cm}^{-2} \quad \text{if } b = 1.\end{aligned}\quad (19)$$

As a check on the self-consistency of the thin-disc approximations, it is useful to calculate the fractional thickness. Because  $h \equiv \Sigma/2\rho$  and  $\rho$  can be eliminated via equation (6),

$$\frac{h}{r} = \frac{\pi r^2 \Sigma}{QM} = \begin{cases} 0.033 Q^{-5/3} (l_E/\alpha_{0.3} \varepsilon_{0.1})^{1/3} M_8^{-1/6} r_{\text{pc}}^{1/2} & b = 0, \\ 0.79 Q^{-7/6} (l_E/\alpha_{0.3} \varepsilon_{0.1})^{2/3} r_{\text{pc}} & b = 1. \end{cases}\quad (20)$$

Here we have expressed the radius in absolute units because  $h/r$  then depends weakly or not at all on black-hole mass.

Beyond  $10^4$ – $10^5 R_S \approx 0.1$ – $1 M_8$  pc, the above formulae predict  $T \lesssim 5000$  K, so that the opacity  $\kappa \ll \kappa_{e.s.}$  and the disc becomes optically thin. (This assumes  $M_8 = l_E = \varepsilon_{0.1} = \alpha_{0.3} = 1$ . Dust will raise opacity again at  $T \lesssim 1700$  K.) The disc must then be supported by gas pressure, notwithstanding equation (11) which presumes that the radiation is trapped. But in a gas-pressure dominated disc, the minimum temperature for gravitational stability is equation (12), which is about an order of magnitude larger than predicted by the formulae above. Note that this temperature is nearly independent of the disc rotation curve, whether dominated by the mass of the black hole or of the bulge stars, as  $\Omega$  is subsumed into  $Q$  and  $M$ . So, in a marginally gravitationally stable disc, there must be an extended region where the temperature adjusts itself within a limited range ( $5000 \text{ K} \lesssim T \lesssim 10^4 \text{ K}$ ) so that the disc is marginally optically thin. At the low densities relevant here, the maximum opacity is  $\kappa_{\text{max}} \approx 10\kappa_{e.s.}$  and is achieved at  $T \approx 10^4$  K (Kurucz 1992; Keady & Kilcrease 2000). Hence the outer edge of the region in question should end at  $\Sigma \approx \kappa_{\text{max}}^{-1} \approx 0.3 \text{ cm}^2 \text{ g}^{-1}$ , which occurs (assuming  $Q = 1$  and  $T = 10^4$  K) at

$$r_{\text{thin}} \approx \frac{c_s \kappa_{\text{max}} v_{\text{circ}}}{\pi G} \approx 170 M_8^{0.24} \text{ pc}, \quad (21)$$

provided, of course, that the disc extends to such large radii. We have taken the circular velocity as  $\sigma_{\text{bulge}} \sqrt{2} \propto M^{0.24}$  (see equations 32 and 33) rather than  $\sqrt{GM/r}$ , because  $r_{\text{thin}}$  lies outside the black hole's sphere of influence (equation 25). Optically-thin constant- $Q$  discs have been considered by Bertin & Lodato (2001) in a somewhat abstract context, but apparently without regard to radiation pressure, which remains important for the vertical hydrostatic balance even at small  $\tau$ ; see Sirko & Goodman (2003).

The parts of the disc beyond  $r_{\text{thin}}$  would radiate predominantly in optical emission lines with velocity widths  $\sim \sigma_{\text{bulge}}$ , so that they might be identified with the QSO narrow-line region. However, we now see that the energy required to maintain the disc at constant  $\dot{M}$  and  $Q \gtrsim 1$  all the way out to  $r_{\text{thin}}$  would be prohibitive.

### 3.2 Energetics

We define a local disc efficiency by

$$\varepsilon'(r) \equiv \frac{4\pi r^2 \sigma T_{\text{eff}}^4}{3Mc^2}. \quad (22)$$

For a viscously heated disc in Newtonian gravity,  $\varepsilon' c^2$  reduces to the local binding energy per unit mass,  $GM/2r = c^2 R_S/4r$  (see equation 3), which is largest at small radii; given a torque-free inner boundary at  $r = r_{\text{min}}$ , the global efficiency is  $\varepsilon = \varepsilon'(r_{\text{min}})$ . But the constant- $Q$  discs require additional energy inputs, so that  $\varepsilon'$  is generally larger than  $R_S/4r$  and tends in fact to *increase* with radius.

Using the results of Section 3.1 for the midplane temperature  $T$ , and assuming that vertical radiative transport obeys equation (5), we find that

$$\begin{aligned}\varepsilon'(r) &= \frac{\pi^{1/3}}{2^{7/6} 3^{1/3}} \hat{\kappa}^{-1} \left( \frac{Q \varepsilon^2 G R_S}{\alpha l^2 c^2 \kappa_{e.s.}} \right)^{1/3} \left( \frac{r}{R_S} \right)^{1/2} \\ &\approx 0.85 \times 10^{-3} Q^{1/3} \alpha_{0.3}^{-1/3} (\varepsilon_{0.1}/l_E)^{2/3} \\ &\quad \times \hat{\kappa}^{-1} M_8^{-1/6} r_{\text{pc}}^{1/2} \quad \text{if } b = 0; \\ &= \frac{\pi^{5/6}}{2^{1/6} 3^{5/6}} \frac{\kappa_{e.s.}^{1/3}}{\kappa} \frac{Q^{5/6} \varepsilon^{1/3}}{\alpha^{2/3} l_E^{1/3}} \left( \frac{G}{c} \right)^{5/6} \left( \frac{m}{k_B} \right)^{2/3} \sigma^{1/6} r \\ &\approx 1.2 \times 10^{-2} Q^{5/6} \alpha_{0.3}^{-2/3} (\varepsilon_{0.1}/l_E)^{1/3} \hat{\kappa}^{-1} r_{\text{pc}} \quad \text{if } b = 1.\end{aligned}\quad (23)$$

It would appear from these relations that  $\varepsilon' \rightarrow \infty$  as  $\kappa \rightarrow 0$ , but they are not valid when the disc is optically thin. If the absorption optical depth  $\tau = \kappa \Sigma < 1$ , then  $T_{\text{eff}}^4 \approx \tau T^4$  rather than  $T^4/\tau$ . So in fact  $\varepsilon' \rightarrow 0$  as  $\kappa \rightarrow 0$ . Hence the largest radius at which equations (23) or (24) are valid is the smallest among the radii  $r_{\text{thin}}$ ,  $R_N$  and  $r_{\text{out}}$ , where the first is the radius at which a  $Q = 1$  disc would become optically thin (equation 21), the second is the radius within which the black hole dominates the potential,

$$R_N \equiv \frac{GM}{2\sigma_{\text{bulge}}^2} \approx (6 \pm 1) M_8^{0.5 \pm 0.1} \text{ pc}, \quad (25)$$

and the last is the actual outer edge of disc, which depends upon the angular momentum of the gas supplied to it. The final expression in equation (25) has used the recently reported  $M$ – $\sigma_{\text{bulge}}$  relation (Gebhardt et al. 2000; Merritt & Ferrarese 2001, see Section 4.2 below). Actually, equation (25) probably underestimates the distance to which the black hole dominates the circular velocity, because the density profile in the cusp of bright bulges is considerably shallower than  $r^{-2}$ .

At  $r = 10 \text{ pc} \approx 10^6 M_8^{-1} R_S$ , equations (23) and (24) imply  $\varepsilon' \approx 0.003 M_8^{-1/6}$  and  $\varepsilon' \approx 0.1$ , respectively. The former efficiency is barely compatible with thermonuclear burning even if all of the disc gas is processed through high-mass stars. The latter is unsustainable by stars and in fact comparable to the efficiency of the central engine.

Even if the energy required to maintain  $Q \sim 1$  can be found, these discs may have severe thermal instabilities. In a steady viscously heated disc, the thermal time (defined as the internal energy per unit area divided by the power radiated per unit area) is  $t_{\text{th}} \sim (\alpha \Omega)^{-1}$ , hence longer than the local dynamical time if  $\alpha < 1$  (Pringle 1981). While this does not guarantee thermal stability, it at least ensures a clear distinction between thermal and dynamical perturbations. For these constant- $Q$  discs, however, the thermal time is shorter by the ratio of viscous heating to total heating. In view of equation (22) and the remarks following it, we can write  $\alpha \Omega t_{\text{th}} \sim \varepsilon'(r)(4r/R_S)$ , and from equations (23)–(24), it follows that  $t_{\text{th}} \ll \Omega^{-1}$  at  $r \sim 1 \text{ pc}$ . In such circumstances, it is far from clear whether  $Q = 1$  is the correct stability criterion. Unfortunately, we cannot say much more about

this without understanding how the auxiliary heat sources respond to changes in local density and temperature, and with what time lag.

#### 4 NON-STANDARD DISCS AND OTHER ALTERNATIVES

As shown in Section 2, a geometrically thin, optically thick accretion disc in a typical high-luminosity QSO would be self-gravitating at radii  $r \gtrsim 10^2\text{--}10^3 R_S$ , or  $10^{-3}\text{--}10^{-2}$  pc, where  $R_S = 2GM/c^2$  is the Schwarzschild radius of the central black hole. Large global radiative efficiency,  $\varepsilon \gtrsim 0.1$ , probably requires a thin disc near  $R_S$ , but it does not much depend upon the nature of the flow at large radius. So, in this section, we consider whether the self-gravitating part of the disc can be replaced by some other form of accretion.

##### 4.1 Hot, quasi-spherical flows

These have been proposed as models for accretion at very low accretion rate ( $\dot{M}$ ) and low radiative efficiency  $\varepsilon \equiv L/\dot{M}c^2$  (Rees et al. 1982; Ichimaru 1977; Narayan & Yi 1994). At sufficiently low density, the gas cooling time and ion–electron thermal equilibration time are longer than the accretion time, so that the ion temperature ( $T_i$ ) is approximately virial and the thickness of the disc is comparable to its radius. We suppose that angular-momentum transport is efficient (viscosity parameter  $\alpha \sim 1$ ), or that the angular momentum of the gas is very small to begin with, so that the accretion velocity ( $v_r$ ) is comparable to the free-fall velocity. The large  $v_r$  and  $T_i$  combine with the low  $\dot{M}$  to make the density low, as required. It is unclear just how low  $\dot{M}$  must be for this mode of accretion to sustain itself, because angular-momentum transport is not well understood and collisionless processes may enhance the thermal coupling of ions and electrons.

A quasi-spherical flow is not viable when the luminosity is close to the Eddington limit, however, because of inverse-Compton cooling. For spherical free fall on to a source of luminosity  $L = l_E L_{\text{Edd}}$ , the inverse-Compton cooling rate of free electrons is

$$t_C^{-1} \approx \frac{2\sigma_T L}{3\pi r^2 m_e c^2} = \frac{8 m_p}{9 m_e} l_E \left(\frac{R_S}{r}\right)^{1/2} t_{\text{ff}}^{-1}. \quad (26)$$

We have introduced the free-fall time of a radial parabolic orbit from radius  $r$ ,

$$t_{\text{ff}} \equiv \frac{2}{3} \left(\frac{r^3}{2GM}\right)^{1/2} \approx 2.2 \times 10^{10} M_8^{-1/2} r_{\text{pc}}^{3/2} \text{ s}, \quad (27)$$

in which  $r_{\text{pc}}$  is the distance from the source in parsec and  $M_8 \equiv M/(10^8 M_\odot)$ . Actually, radiation pressure increases the free-fall time by  $(1 - l_E)^{-1/2}$ , but this factor is neglected for simplicity. From equation (26),  $t_C < t_{\text{ff}}$  at  $r < 25 l_E^2 M_8$  pc. Hence the electrons assume the colour temperature of the central source (or even less, see below):  $T_C \lesssim 10^5 M_8^{-1/4}$ . This is much less than the virial temperature,  $T_{\text{vir}} \approx 10^7 M_8 r_{\text{pc}}^{-1}$ , at all radii of interest to the present paper.<sup>1</sup> The electron density is

$$n_e = \left(\frac{l_E}{\varepsilon}\right) \frac{1}{3c\sigma_T t_{\text{ff}}}, \quad (28)$$

<sup>1</sup>The time-scale for Poynting–Robertson drag is  $\approx (m_p/m_e)t_C$ , hence  $\gg t_{\text{ff}}$  at  $r \gg R_S$ , i.e. the radiation field does not remove angular momentum from the gas fast enough to prevent it from circularizing.

so that the electron–ion equilibration time due to Coulomb collisions alone (Spitzer 1978) is much shorter than the flow time:

$$\frac{t_{\text{eq}}}{t_{\text{ff}}} \approx 2 \times 10^{-5} \varepsilon l_E \left(\frac{T_e}{10^5 \text{ K}}\right)^{3/2}.$$

Of course,  $T_e \sim 10^5$  K is the peak of the cooling curve (Spitzer 1978). If  $n_e$  obeys equation (28), then radiative cooling is actually faster than inverse-Compton cooling:

$$\frac{t_{\text{rad}}}{t_C} \approx 6. \varepsilon \left(\frac{R_S}{r}\right)^{1/2} (T_e = 10^5 \text{ K}).$$

In short, both the ions and the electrons of a quasi-spherical flow on to a near-Eddington QSO cool in much less than a free-fall time. A thin disc will form unless the specific angular momentum of the flow is negligible.

##### 4.2 Collisional stellar cluster

Dense stellar clusters have occasionally been nominated as precursors to QSO black holes, either by relativistic collapse (Zel’dovich & Podurets 1966; Shapiro & Teukolsky 1985; Ebisuzaki et al. 2001) or by collisions among non-degenerate stars (Spitzer & Saslaw 1966; Rees 1978). Our interest, however, is in stellar collisions as the main source of fuel for an already very massive black hole. This has been studied by McMillan, Lightman & Cohn (1981) and Illarionov & Romanova (1988), among others, who show that in order to supply a  $10^8 M_\odot$  black hole at its Eddington rate, the velocity dispersion of such a cluster must be  $\gtrsim 10^3$  km s<sup>-1</sup>. To demonstrate the robustness of this conclusion, we make some oversimplified but conservative estimates here.

In order to provide high radiative efficiency ( $\varepsilon$ ), stars must be disrupted and their gaseous debris circularized before accretion. If  $M_{\text{b.h.}} \lesssim 10^8 M_\odot$ , main-sequence stars scattered on to loss-cone orbits are likely to be tidally disrupted rather than swallowed whole (Hills 1975). About half of the tidal debris is unbound and promptly escapes from the black hole (Lacy, Townes & Hollenbach 1982; Evans & Kochanek 1989), and much of the remainder is likely to be swallowed at low radiative efficiency before the gas circularizes (Cannizzo, Lee & Goodman 1990; Ayal, Livio & Piran 2000). We limit our discussion to  $M_{\text{b.h.}} \gtrsim 10^8 M_\odot$ , as required for the most luminous QSOs, and we assume that stars are disrupted by stellar collisions. We ignore loss-cone effects because stars swallowed whole do not contribute to the QSO luminosity.

The total collision rate among  $N$  stars forming a cluster of structural length  $a$  is

$$\dot{N} = C \bar{\sigma}^7 N^2 (GM)^{-3} R_*^2, \quad (29)$$

where  $C$  is a dimensionless coefficient,  $\bar{\sigma}$  is the rms velocity dispersion in one dimension averaged over the cluster,  $N$  is the number of stars, and  $R_* \sim R_\odot$  is the radius of an individual star.  $M$  is the total mass that determines  $\bar{\sigma}$  via the virial theorem, so that, if  $m_*$  is the mass of individual stars and  $M_* \equiv Nm_*$ , then  $M \approx M_{\text{bh}} + \frac{1}{2} M_*$ , the factor of 1/2 being needed to avoid double-counting the gravitational interactions among the stars. We take  $4\pi R_*^2$  for the collision cross-section; this allows for grazing collisions that probably would not disrupt the stars (Spitzer & Saslaw 1966), but it will lead to a conservative estimate of  $\bar{\sigma}$ . Gravitational focusing is unimportant at the high velocity dispersions relevant here.

A second relation for  $\dot{N}$  follows by requiring the collisional debris to sustain the QSO at a fraction  $l_E \equiv L/L_E$  of its Eddington luminosity:

$$2M_* \dot{N} = \frac{l_E}{\varepsilon} \frac{4\pi G M m_p}{c \sigma_T} \quad (30)$$

Eliminating  $\dot{N}$  between equations (29) and (30) leads to

$$\bar{\sigma}^7 = \frac{2\pi l_E}{C \varepsilon} \frac{G^4 m_p}{\sigma_T c} \frac{m_*}{R_*^2} \left( \frac{M_{\text{b.h.}} M^3}{M_*^2} \right).$$

The term in parentheses achieves its minimum value  $27M_{\text{b.h.}}^2/16$  at  $M_* = 4M_{\text{b.h.}}$ .

The coefficient  $C$  depends upon the density profile  $\rho(r)$  of the stellar cluster. It can be arbitrarily large if  $\rho(r)$  is sufficiently steep, but then the collision rate is dominated by stars at small radius. These tightly-bound stars represent only a small fraction of the cluster mass and would be consumed in much less than the growth time of the black hole, unless their total mass ( $M_{\text{t.b.}}$ ) is  $\gtrsim M_{\text{b.h.}}$ , in which case the tightly-bound population is substantially self-gravitating and we redefine  $M_* \equiv M_{\text{t.b.}}$ . We might suppose that  $M_{\text{t.b.}} \ll M_{\text{b.h.}}$  and that the tightly-bound stars were continuously replenished by two-body relaxation from a reservoir of more weakly bound stars; but the larger cluster would then have to expand to conserve energy, reducing the collision rate and the fueling of the QSO. For these reasons, we consider clusters for which  $C$  is dominated by stars near the half-mass radius. For example, if the stars have isotropically distributed orbits with a common semimajor axis  $a$  in a Keplerian potential, then  $\rho(r) \propto \sqrt{(2a/r)-1}$ ,  $\bar{\sigma}^2 = GM/3a$ , and  $C \approx 196$ . This leads to

$$\bar{\sigma} \approx 760 \left( \frac{l_E}{10\varepsilon} \right)^{1/7} \left( \frac{g_*}{g_\odot} \right)^{1/7} M_8^{2/7} \text{ km s}^{-1}, \quad (31)$$

where  $g_*$  is the stellar surface gravity. Because of the one-seventh root, the result is not very sensitive to our assumptions. A Plummer sphere of the same total mass,  $M_* = 5 \times 10^8 M_\odot$ , yields  $\bar{\sigma} \approx 730 \text{ km s}^{-1}$ .

Equation (31) can be compared with recently-discovered empirical relations between inactive black holes – presumably QSO relics – and their host bulges. Gebhardt et al. (2000) find

$$M_{\text{b.h.}} = 1.2(\pm 0.2) \times 10^8 \left( \frac{\sigma_e}{200 \text{ km s}^{-1}} \right)^{3.75 \pm 0.3} M_\odot, \quad (32)$$

where  $\sigma_e$  is the line-of-sight velocity dispersion at one effective radius. Merritt & Ferrarese (2001) use the central velocity dispersion, which is usually little different from  $\sigma_e$ :

$$M_{\text{b.h.}} = 1.30(\pm 0.36) \times 10^8 \left( \frac{\sigma_c}{200 \text{ km s}^{-1}} \right)^{4.72 \pm 0.36} M_\odot. \quad (33)$$

The scaling exponent  $d \log M/d \log \sigma = 3.5$  implied by equation (31) is similar to the empirical equations (32) and (33), but the normalization is very different. Extrapolated to  $700 \text{ km s}^{-1}$ , the empirical relations predict  $M_{\text{b.h.}} \gtrsim 10^{10} M_\odot$  instead of  $10^8 M_\odot$ .

Therefore, if QSOs were fuelled by dense stellar clusters, these clusters must have been an order of magnitude more tightly bound than the surrounding bulge, and they would have been a dynamically distinct stellar component. There seems to be very little trace of this tightly-bound stellar population in present-day bulges. How would such a component form? A likely possibility is gaseous dissipation followed by star formation. As will be seen, accretion in a thin viscous disc may lead to just such a result. We have discussed fuelling the QSO by a disc and fuelling it by stellar collisions as though these were mutually exclusive possibilities, but perhaps the two occur in concert.

### 4.3 Wind-driven discs

In principle at least, a magnetized wind can remove angular momentum from a thin disc rather efficiently. Compared to a viscous disc of the same sound speed ( $c_s$ ) and accretion rate ( $\dot{M}$ ), a wind-driven disc might have an accretion velocity that is larger by a factor  $\sim r\Omega/c_s = r/h$ . The surface density would be correspondingly reduced, as would the tendency toward self-gravity.

If viscous transport can be neglected, the vectorial angular momentum flux is

$$\mathbf{F}_J = r \left( \rho v_\phi \mathbf{v} - \frac{B_\phi \mathbf{B}}{4\pi} \right).$$

We integrate this over an annular surface enclosing the disc from its inner edge at  $r_{\text{min}}$  out to an arbitrary radius  $r > r_{\text{min}}$ ,

$$(r', z') : r_{\text{min}} \leq r' \leq r, \quad -h(r) \leq z' \leq h(r).$$

In steady state, the angular momentum enclosed by this surface is constant, whence the surface integral of  $\mathbf{F}_J$  vanishes:

$$\begin{aligned} & [\dot{M}(r') \Omega(r') r'^2]_{r'=r_{\text{min}}}^{r'=r} \\ &= \int_{r_{\text{min}}}^r \left[ \frac{\overline{B_z B_\phi}(r', -h) - \overline{B_z B_\phi}(r', h)}{2} + \frac{d\dot{M}}{dr} \Omega(r') \right] r'^2 dr'. \end{aligned} \quad (34)$$

The left-hand side is the angular momentum carried through the side walls of the annulus by the accreting gas ( $\dot{M} > 0$  for inflow). The right-hand side is the angular momentum escaping through the top and bottom faces by magnetic stresses and mass outflow. The overbars average over azimuth and time, as the magnetic field is likely to be complicated on small scales. Mass conservation has been invoked, so that the mass flux through the top and bottom faces balances the net flux through the sides. Let us also assume that  $B_\phi$  is odd in  $z$  while  $B_z$  is even; the field lines go through the disc but bend in the direction opposite to its rotation at both faces. Furthermore, we take  $r \gg r_{\text{min}}$  so that  $\Omega(r)r^2 \gg \Omega(r_{\text{min}})r_{\text{min}}^2$  and we neglect the angular momentum carried through  $r_{\text{min}}$ . Finally, we assume that the wind is magnetically dominated so that the term involving  $\dot{M}$  on the right-hand side can be neglected compared to the magnetic stresses. Then the radial accretion speed

$$v_{\text{acc}} \sim \overline{B_z B_\phi} / 4\pi \Sigma \Omega.$$

The main point to notice about equation (34) is that the area of the top and bottom faces is much larger than that of the sides. This means that, for the same field strength, a magnetized wind can be much more effective at removing angular momentum from the accreting gas than a field confined to the gas layer. The accretion speed driven by a magnetic ‘viscosity’,

$$v_{\text{acc}} = \frac{1}{4\pi \Sigma \Omega r} \int_{-h}^h \overline{B_r B_\phi} dz,$$

is smaller than the rate implied by equation (34) by a factor  $\sim (h/r)(B_r/B_z)$ .

Viscous stresses are normally presumed to be smaller than thermal pressure, i.e.  $\alpha < 1$ . For a magnetic viscosity,  $\alpha \approx -\overline{B_r B_\phi} / 4\pi p$ , where  $\alpha < 1$  if the field strength is less than the equipartition value. A superequipartition field would shut off the crucial MRI (Balbus & Hawley 1998), and would probably escape the disc by interchange or Parker instabilities. We now give two arguments to show that, even in a wind-driven accretion disc, the field must be at or below equipartition.



Vertical hydrostatic equilibrium entails

$$-\left[\frac{B_r^2 + B_\phi^2 - B_z^2}{8\pi} + p\right]_{z=0}^{z=h} \approx \int_0^h \rho \Omega^2 z \, dz.$$

For dipolar symmetry,  $B_z$  is approximately constant with  $z$  on the scale  $h \ll r$ , whereas the horizontal components vanish at the mid-plane. Because the right-hand side above is positive, and because  $p(z=h) \ll p(z=0)$ , it follows that

$$\frac{B_r^2 + B_\phi^2}{8\pi}(z = \pm h) \lesssim p(z=0). \quad (35)$$

Henceforth  $B_r$  and  $B_\phi$  are evaluated at  $z = h$ , and  $p$  at  $z = 0$ . As pointed out by Blandford & Payne (1982), in order that centrifugal force should drive the wind outward, the poloidal field lines must make an angle of at most  $60^\circ$  with the surface of the disc, so that  $B_r/B_z \geq \sqrt{3}$ . Therefore

$$\frac{\mathbf{B}^2}{8\pi} \leq \frac{4B_r^2 + B_\phi^2}{8\pi} \leq 4p.$$

The stress on the faces of the disc is therefore bounded by

$$\left|\frac{B_z B_\phi}{4\pi}\right| \leq \sqrt{3} \left|\frac{B_r B_\phi}{4\pi}\right| \leq \sqrt{3} \frac{B_r^2 + B_\phi^2}{8\pi} < \sqrt{3} p.$$

Using equation (34), we have

$$|v_{\text{acc}}| \lesssim \frac{\sqrt{3}}{\Omega r^3 \Sigma} \int_{r_{\text{min}}}^r \frac{\Sigma c_s^2}{h}(r') r'^2 \, dr' \sim \sqrt{3} \frac{c_s^2}{\Omega h}. \quad (36)$$

So the inflow could be marginally supersonic. For a viscous disc, on the other hand,  $|v_{\text{acc}}| \approx \alpha c_s^2 / \Omega r$ , which is strongly subsonic as long as  $\alpha \ll r/h$ .

The angular momentum extracted from the face of the disc by the field must be carried off by the wind. It is problematic whether a strong wind can be launched from the disc (Ogilvie & Livio 2001), and whether rapid wind-driven accretion is stable (Cao & Spruit 2002). Apart from these difficulties, the consideration of the magnetic flux,

$$\Phi(r) \equiv \int_{r_{\text{min}}}^r B_z(r', 0) 2\pi r' \, dr',$$

leads to important constraints on the field strength and accretion rate. Presumably  $\Phi(r)$  should not change secularly. If  $B_z$  is predominantly of one sign, then the increase of  $|\Phi|$  by advection must be balanced by diffusion of the lines through the inflowing gas. The drift velocity of the lines  $v_{\text{drift}} = -v_{\text{acc}} \sim \eta_{\text{eff}}/r$  if  $B_z$  varies on scales  $\sim r$ , where  $\eta_{\text{eff}}$  is the effective diffusivity of the gas. In significantly ionized discs, the microscopic diffusivity is negligible, so  $\eta_{\text{eff}}$  is due to turbulence, and we expect  $\eta_{\text{eff}} \sim \alpha' c_s h$  with  $\alpha' \lesssim 1$ . Hence, for a steady wind-driven disc threaded by net magnetic flux,  $|v_{\text{acc}}| \lesssim (\alpha' h/r) c_s$ , which is probably very much less than the upper limit (36) and comparable to the accretion velocity of a viscous disc. Alternatively, the net flux could be essentially zero if  $B_z$  changes sign on scales  $\ll r$ . In the latter case, the higher flow speed (36) may be achievable. But so irregular a field would probably have to be sustained by dynamo action within the disc rather than inherited from whatever region supplies the accreting gas. This probably requires MRI and gives another argument for a subequipartition field.

To summarize this subsection, accretion driven by magnetized winds is even less well understood than viscous accretion but might allow substantially higher accretion velocities and lower surface densities, perhaps by factors up to  $\sim r/\alpha h$ .

#### 4.4 Thin discs with strongly magnetized coronae

This is a variant of Section 4.3 in which most of the field lines are not open but re-attach to the disc at large distances  $\Delta r \gg h$  (Galeev et al. 1979; Heyvaerts & Priest 1989). The vertical magnetic scale height is then  $H \sim \Delta r \gg h$ . The angular momentum flux carried through the corona,

$$J_{\text{cor}}(r) = -2 \int_{z=h}^{z=\infty} dr \int_0^{2\pi} d\phi \frac{B_r B_\phi}{4\pi} \sim H \frac{B_r B_\phi}{4\pi}(r, h),$$

can be larger than the flux within the gas layer by a factor  $\sim H/h$ , so that the effective value of  $\alpha$  might be as large as  $r/h$  (for  $H \sim r$ ) without fields exceeding equipartition.

The evidence for magnetized coronae that dominate angular-momentum transport is suggestive but inconclusive. Local simulations of MRI generally find that the scale height of the field exceeds that of the gas, but only by factors of order unity; they also find  $\alpha \sim 10^{-2}$ – $10^{-1}$  rather than  $\sim r/h$  (Brandenburg et al. 1995; Stone et al. 1996; Miller & Stone 2000). Possibly,  $H/h$  is limited by the fact that the smallest dimension of the computational domain is  $\lesssim h$ . Global simulations of MRI have been performed for relatively thick discs only, so that it is difficult to distinguish scalings with  $h$  from scalings with  $r$  (Matsumoto & Shibata 1997; Hawley 2000). Global simulations of *thin* discs in three dimensions may not be available for some time because of the very large numbers of grid cells needed to resolve both the disc and the corona. Merloni & Fabian (2001) argue that X-ray observations of accreting black holes – active galactic nuclei (AGN) and X-ray binaries – demand a strongly magnetized corona, at least in the innermost part of the disc. On the other hand, observations of eclipsing cataclysmic variables indicate that X-rays are emitted from the disc–star boundary layer rather than an extended corona (Mukai et al. 1997; Ramsay et al. 2001).

#### 4.5 Global spiral waves

As is well known, a trailing  $m$ -armed spiral density wave

$$\Sigma(r, \phi) = \Sigma_0(r) + \Sigma_m(r) \cos(m\phi + \mu \ln r) \quad (37)$$

exerts an outward (positive) gravitational angular-momentum flux

$$\Gamma \approx \pi^2 G \Sigma_m^2 r^3 \frac{m\mu}{|\mu|^3}$$

(Lynden-Bell & Kalnajs 1972). The above approximation is good for tightly-wrapped waves,  $\mu \gg m \neq 0$ , which carry relatively little flux for a given density contrast  $\Sigma_m/\Sigma_0$ . An exact formula for logarithmic spirals with  $\Sigma_m \propto r^{-3/2}$  is

$$\Gamma = -\pi^2 G \Sigma_m^2 r^3 m \frac{\partial}{\partial \mu} K(\mu, m), \quad (38)$$

where  $K(\mu, m)$  is the Kalnajs function (Kalnajs 1971):

$$K(\mu, m) = \frac{1}{2} \left| \frac{\Gamma \left[ \frac{1}{2} \left( m + i\mu + \frac{1}{2} \right) \right]}{\Gamma \left[ \frac{1}{2} \left( m + i\mu + \frac{3}{2} \right) \right]} \right|^2 \quad (\text{real } \mu). \quad (39)$$

With  $\Sigma_m/\Sigma_0 = 1$ , the largest ratio for which the surface density is everywhere positive, we find that the torque is maximized at  $m = 1$  and a pitch angle  $\tan^{-1}(m/\mu) \approx 48^\circ.2$ , so that

$$\Gamma_{\text{max}} \approx 0.961 \pi G \Sigma_0^2 r^3. \quad (40)$$

Suppose that the gravitational torque is balanced by the advection of angular momentum with the accreting gas. In other words,

$\Gamma \approx \dot{M}\Omega r^2$ , so that there is no secular change in the angular momentum within radius  $r$ . The gravitationally-driven accretion speed is then

$$|v_r| \leq \frac{\Gamma_{\max}}{2\pi r^3 \Omega \Sigma_0} \approx 0.15 Q^{-1} c_s \quad (41)$$

in which  $Q$  has been calculated from the azimuthal average of the surface density,  $\Sigma_0$ . In principle therefore, accretion may occur at a significant fraction of the sound speed. But the existence of a self-consistent wave-driven flow has been assumed rather than proved. Non-linear single-armed spirals in Keplerian discs have been found by Lee & Goodman (1999), but only for weak self-gravity ( $Q \gg 1$ ) and without dissipation or accretion.

In steady accretion on to a central mass that dominates the rotation curve, the advected angular-momentum flux  $\Gamma = \dot{M}\Omega r^2 \propto r^{1/2}$ , so  $\Sigma \propto r^{-5/4}$  rather than  $r^{-3/2}$ . Presumably the slight change in power-law index does not change the results (40) and (41) much.

#### 4.6 Clumpy discs

Can QSO discs persist at  $Q \ll 1$  without fragmenting entirely into stars? Lin & Pringle (1987) have suggested that complete fragmentation does not necessarily result from such a low  $Q$ . Numerical simulations suggest that it does (Monaghan & Lattanzio 1991; Gammie 2001).

The question is all the more urgent because of rather direct evidence for parsec-scale accretion discs in nearby AGN, if not QSOs, from VLBI observations of maser emission. If the nuclear disc of NGC 1068 is in a steady state, then the nuclear luminosity implies  $Q \sim 10^{-3}$  at  $r \sim 1$  pc (Kumar 1999). NGC 4258 is much less luminous, and estimates of  $\dot{M}$  range from  $7 \times 10^{-5} \alpha_{0.1} \dot{M}_\odot$  based on modeling the maser emission itself (Neufeld & Maloney 1995), to  $10^{-2} \dot{M}_\odot$  (Gammie, Narayan & Blandford 1999; Kumar 1999) for an assumed central advection dominated accretion flow (ADAF); at the former rate,  $Q \sim 1$  at the outer edge of the masing region,  $\sim 0.2$  pc (Maoz 1995), while in the latter,  $Q \sim 10^{-2}$ .

Kumar (1999) has suggested a clumpy rather than smooth disc, in which accretion occurs by gravitational scattering and physical collisions among clumps rather than an  $\alpha$  viscosity. These clumps are supposed to be gas clouds rather than fully formed stars, in order to provide appropriate conditions for maser amplification. Although the model deals with the stability and accretion rate of the clumpy disc as a whole, it does not ask what prevents the individual clumps from collapsing. In the application to NGC 1068, the masses, radii and surface temperatures of the clumps are quoted as  $M_c \sim 10^3 M_\odot$ ,  $R_c \sim 0.1$  pc and  $T_{\text{eff},c} \approx 500$  K; the virial temperature implied by this mass and radius is  $\sim 2000$  K. These clumps would be moderately optically thick at the wavelength corresponding to  $T_{\text{eff},c}$ . Their characteristic thermal time – the time required to radiate their binding energy – is

$$t_{\text{th},c} \approx \frac{GM_c^2/R_c}{4\pi R_c^2 \sigma T_{\text{eff},c}^4} \sim 10^5 \text{ s},$$

hence many orders of magnitude less than the orbital time ( $\Omega^{-1} \sim 10^{11}$  s at  $r \sim 1$  pc) or interclump collision time. Of course, the surface temperatures of these objects are assumed to be maintained by irradiation from the central source, but without a fast-responding feedback mechanism, the thermal equilibrium is unstable. If a clump starts to contract, its surface temperature will rise, and collapse will proceed on the clump's internal free-fall time-scale (as this is larger than  $t_{\text{th},c}$ ).

To answer the question raised above, we therefore believe that no QSO accretion disc, whether smooth or clumpy, can persist at  $Q \ll 1$ .

## 5 DISCUSSION

Although the accretion velocity of a thin viscous disc is very strongly subsonic, with a Mach number  $\mathcal{M} \equiv v_r/c_s \approx \alpha h/r$ , the discussion of Section 4 suggests that  $\mathcal{M}$  might approach  $\sim 0.1$  if accretion is driven by large-scale magnetic or gravitational fields rather than a local effective viscosity. Self-consistent solutions that achieve this bound would be very interesting to pursue.

Even at near-sonic accretion speeds, QSO discs become self-gravitating at radii less than a parsec if there is no important source of heating other than dissipation of orbital energy. Combining equations (3) and (5) with  $\dot{M} = 2\pi r \Sigma \mathcal{M} c_s$  instead of the viscous relation (2), and assuming that gas pressure dominates ( $\beta \approx 1$ ), we find that

$$\begin{aligned} Q &= 2^{-25/18} \pi^{-1} \mathcal{M}^{7/9} \hat{\kappa}^{2/9} \kappa_{e,s}^{7/9} G^{-1} \left( \frac{c k_B}{\sigma^{1/4} m} \right)^{8/9} R_S^{-11/9} \left( \frac{r}{R_S} \right)^{-25/18} \\ &\approx 0.8 \left( \frac{r}{10^4 R_S} \right)^{-25/18} \left( \frac{\mathcal{M}_{0.1}^7 \varepsilon_{0.1}^5 \hat{\kappa}^2}{I_E^5 M_8^{11}} \right)^{1/9} \\ &\approx 0.8 \left( \frac{r}{0.1 \text{ pc}} \right)^{-25/18} \left( \frac{\mathcal{M}_{0.1}^7 \varepsilon_{0.1}^5 \hat{\kappa}^2}{I_E^5} \right)^{1/9} M_8^{1/6}. \end{aligned} \quad (42)$$

On the other hand, we have seen (Section 3.2) that, when angular-momentum transport is viscous, then it is unlikely that stars can supply enough additional heat to stabilize the disc beyond 1 pc, especially if viscosity is proportional to gas pressure rather than total pressure as viscous stability probably demands. In the latter case, even low-mass black holes embedded in the disc are probably inadequate. These statements assume that enough free gas remains in the disc to supply the central black hole at its Eddington rate. If all the gas converts to stars, stability may result, but the quasar is quenched. Perhaps a combination of stellar (or embedded-black-hole) heating and a superviscous accretion speed may allow an extended gravitationally stable disc; we hope to explore this possibility in a future paper.

Given the serious theoretical difficulties of all proposed mechanisms for keeping  $Q \gtrsim 1$  at large radii, we are forced to take seriously the only remaining possibility: that QSO discs do not exist much beyond  $R_{\text{sg}} \sim 10^{-2}$  pc (equation 10) – at least not in a state of centrifugal support, vertical hydrostatic equilibrium, and steady accretion. Yet the mass within this radius is smaller than that of the central black hole by a factor  $\sim (h/r) \sim 10^{-2}$  (since the midplane density  $\approx M/2\pi r_3$  at  $Q = 1$ ). Hence, in order to grow the black hole, the disc must be replenished, either steadily or intermittently, by infall of low-angular-momentum material. In order that the gas not circularize outside  $10^{-2} r_{\text{pc}}$ , its specific angular momentum must be  $\lesssim 70 M_8^{1/2} \text{ km s}^{-1} \text{ pc}$ , or some three orders of magnitude smaller than that of most stars in ellipticals and bulges (Binney & Merrifield 1998). Such a small ratio may seem unlikely, but on the other hand,  $M_{\text{bh}}/M_{\text{bulge}} \approx 10^{-3}$  (McLure & Dunlop 2002). So perhaps the QSO is fuelled by the low-angular-momentum tail of gas that forms the bulge. This gas would arrive at the outer edge of the disc in a vertically broad infall, perhaps already carrying dust formed from metals injected by outflows from the bulge or QSO disc itself, and hence taking the place of a warped outer disc as the source of reprocessed infrared light. The picture this calls to mind is similar, except in

scale, to the standard scenario for the formation of a protostar (Shu, Adams & Lizano 1987).

## ACKNOWLEDGMENTS

I thank Charles Gammie, Pawan Kumar, Kristen Menou, Ramesh Narayan, Bohdan Paczyński and Jonathan Tan for helpful discussions. This work was supported by the NASA Origins Programme under grant NAG 5-8385.

## REFERENCES

- Ayal S., Livio M., Piran T., 2000, *ApJ*, 545, 772  
 Balbus S. A., Hawley J. F., 1998, *Rev. Mod. Phys.*, 70, 1  
 Balbus S. A., Papaloizou J. C. B., 1999, *ApJ*, 521, 650  
 Begelman M. C., Meier D. L., 1982, *ApJ*, 253, 873  
 Bertin G., 1997, *ApJ*, 478, L71  
 Bertin G., Lodato G., 1999, *A&A*, 350, 694  
 Bertin G., Lodato G., 2001, *A&A*, 370, 342  
 Binney J., Merrifield M., 1998, *Galactic Astronomy*. Princeton Univ. Press, Princeton  
 Blandford R. D., Payne D. G., 1982, *MNRAS*, 199, 883  
 Brandenburg A., Nordlund A., Stein R. F., Torkelsson U., 1995, *ApJ*, 446, 741  
 Cameron A. G. W., 1978, in *Protostars and Planets: Studies of Star Formation and of the Origin of the Solar System Physics of the Primitive Solar Nebula and of Giant Gaseous Protoplanets*, pp. 453–487  
 Cannizzo J. K., Lee H. M., Goodman J., 1990, *ApJ*, 351, 38  
 Cao X., Spruit H. C., 2002, *A&A*, 385, 289  
 Collin S., Zahn J., 1999a, *A&A*, 344, 433  
 Collin S., Zahn J., 1999b, *Ap&SS*, 265, 501  
 Ebisuzaki T. et al., 2001, *ApJ*, 562, L19  
 Evans C. R., Kochanek C. S., 1989, *ApJ*, 346, L13  
 Galeev A. A., Rosner R., Vaiana G. S., 1979, *ApJ*, 229, 318  
 Gammie C. F., 2001, *ApJ*, 553, 174  
 Gammie C. F., Narayan R., Blandford R., 1999, *ApJ*, 516, 177  
 Gebhardt K. et al., 2000, *ApJ*, 539, L13  
 Greenhill L. J., Gwinn C. R., 1997, *Ap&SS*, 248, 261  
 Greenhill L. J., Jiang D. R., Moran J. M., Reid M. J., Lo K. Y., Claussen M. J., 1995, *ApJ*, 440, 619  
 Hawley J. F., 2000, *ApJ*, 528, 462  
 Heyvaerts J. F., Priest E. R., 1989, *A&A*, 216, 230  
 Hills J. G., 1975, *Nat*, 254, 295  
 Ichimaru S., 1977, *ApJ*, 214, 840  
 Illarionov A. F., Romanova M. M., 1988, *AZh*, 65, 682  
 Kalnajs A. J., 1971, *ApJ*, 166, 275  
 Keady J. J., Kilcrease D. P., 2000, in Cox A. N., ed., *Allen's Astrophysical Quantities*, 4th edn. AIP Press, Springer, New York, ch. 5, pp. 95–120  
 Kumar P., 1999, *ApJ*, 519, 599  
 Kurucz R. L., 1992, *Rev. Mex. Astron. Astrofis.*, 23, 181  
 Lacy J. H., Townes C. H., Hollenbach D. J., 1982, *ApJ*, 262, 120  
 Larson R. B., 1984, *MNRAS*, 206, 197  
 Lee E., Goodman J., 1999, *MNRAS*, 308, 984  
 Lightman A. P., Eardley D. M., 1974, *ApJ*, 187, L1  
 Lin D. N. C., Pringle J. E., 1987, *MNRAS*, 225, 607  
 Lodato G., Bertin G., 2001, *A&A*, 375, 455  
 Lynden-Bell D., Kalnajs A. J., 1972, *MNRAS*, 157, 1  
 Maoz E., 1995, *ApJ*, 455, L131  
 Martin C. L., Kennicutt R. C., 2001, *ApJ*, 555, 301  
 Matsumoto R., Shibata K., 1997, in *ASP Conf. Ser. Vol. 121, Accretion Phenomena and Related Outflows*. Astron. Soc. Pac., San Francisco, p. 443  
 McLure R. J., Dunlop J. S., 2002, *MNRAS*, 331, 795  
 McMillan S. L. W., Lightman A. P., Cohn H., 1981, *ApJ*, 251, 436  
 Menou K., Quataert E., 2001, *ApJ*, 552, 204  
 Merloni A., Fabian A. C., 2001, *MNRAS*, 321, 549  
 Merritt D., Ferrarese L., 2001, *ApJ*, 547, 140  
 Miller K. A., Stone J. M., 2000, *ApJ*, 534, 398  
 Monaghan J. J., Lattanzio J. C., 1991, *ApJ*, 375, 177  
 Mukai K., Wood J. H., Naylor T., Schlegel E. M., Swank J. H., 1997, *ApJ*, 475, 812  
 Nakai N., Inoue M., Miyoshi M., 1993, *Nat*, 361, 45  
 Narayan R., Yi I., 1994, *ApJ*, 428, L13  
 Neufeld D. A., Maloney P. R., 1995, *ApJ*, 447, L17  
 Ogilvie G. I., Livio M., 2001, *ApJ*, 553, 158  
 Paczyński B., 1978, *Acta Astron.*, 28, 91  
 Pariev V. I., Blackman E. G., Boldyrev S. A., 2002, *MNRAS*, submitted(astro-ph/0208400)  
 Pringle J. E., 1981, *ARA&A*, 19, 137  
 Ramsay G. et al., 2001, *A&A*, 365, L288  
 Rees M. J., 1978, *The Observatory*, 98, 210  
 Rees M. J., Phinney E. S., Begelman M. C., Blandford R. D., 1982, *Nat*, 295, 17  
 Salpeter E. E., 1964, *ApJ*, 140, 796  
 Sanders D. B., Phinney E. S., Neugebauer G., Soifer B. T., Matthews K., 1989, *ApJ*, 347, 29  
 Sellwood J. A., Balbus S. A., 1999, *ApJ*, 511, 660  
 Shakura N. I., Sunyaev R. A., 1973, *A&A*, 24, 337  
 Shapiro S. L., Teukolsky S. A., 1985, *ApJ*, 292, L41  
 Shlosman I., Begelman M. C., 1987, *Nat*, 329, 810  
 Shlosman I., Begelman M. C., 1989, *ApJ*, 341, 685  
 Shu F. H., Adams F. C., Lizano S., 1987, *ARA&A*, 25, 23  
 Sirko E., Goodman J., 2003, *MNRAS*, in press  
 Spitzer L., 1978, *Physical Processes in the Interstellar Medium*. Wiley-Interscience, New York  
 Spitzer L. J., Saslaw W. C., 1966, *ApJ*, 143, 400  
 Stone J. M., Hawley J. F., Gammie C. F., Balbus S. A., 1996, *ApJ*, 463, 656  
 Zel'dovich Y. B., Podurets M. A., 1966, *Sov. Astron.*, 9, 742

## APPENDIX A: VISCOUSLY HEATED DISCS

For completeness, in this appendix we give formulae for the mid-plane temperature ( $T$ ) and other characteristics of a steady optically thick disc heated by viscous dissipation only.

Combining equations (2), (3) and (5), and writing  $\beta c_s^2 = k_B T / m$ , where  $m \approx m_H$  is the mean mass per gas particle, we have the radial dependence of  $\Sigma$  and  $T$ :

$$T = \left( \frac{\kappa m}{16\pi^2 \alpha \beta^{b-1} k_B \sigma} \right)^{1/5} \dot{M}^{2/5} \Omega^{3/5} \\ \approx 1.0 \times 10^5 \left( \frac{l_E^2 \hat{\kappa}}{\varepsilon_{0.1}^2 \alpha_{0.01} \beta^{b-1}} \right)^{1/5} M_8^{-1/5} \left( \frac{10^3 R_S}{r} \right)^{9/10} \text{ K}, \quad (\text{A1})$$

$$\Sigma = \frac{2^{4/5}}{3\pi^{3/5}} \left( \frac{m^4 \sigma}{k_B^4} \right)^{1/5} (\alpha \beta^{b-1})^{-4/5} \kappa^{-1/5} \dot{M}^{3/5} \Omega^{2/5} \\ \approx 3.9 \times 10^6 (\alpha_{0.01} \beta^{b-1})^{-4/5} l_E^{3/5} \varepsilon_{0.1}^{-3/5} \hat{\kappa}^{-1/5} M_8^{1/5} \\ \times \left( \frac{10^3 R_S}{r} \right)^{3/5} \text{ g cm}^{-2}. \quad (\text{A2})$$

If viscosity scales with gas pressure ( $b = 1$ ) then equations (A1) and (A2) do not depend on  $\beta$ , which in any case is a known function of density and temperature:

$$\frac{\beta}{1-\beta} = \frac{p_{\text{gas}}}{p_{\text{rad}}} = \frac{3ck_{\text{B}}}{4\sigma m} \frac{\rho}{T^3} = \frac{3c}{8\sigma} \left(\frac{k_{\text{B}}}{m}\right)^{1/2} \beta^{1/2} \frac{\Sigma\Omega}{T^{7/2}};$$

using equations (A1) and (A2),

$$\begin{aligned} \frac{\beta^{(1/2)+(b-1)/10}}{1-\beta} &= (2^3\pi^4)^{1/5} \alpha^{-1/10} c(k_{\text{B}}/m)^{2/5} \sigma^{-1/10} \\ &\quad \times \kappa^{-9/10} \Omega^{-7/10} \dot{M}^{-4/5} \\ &\approx 0.44 \alpha_{0.01}^{-1/10} \varepsilon_{0.1}^{4/5} l_{\text{E}}^{-4/5} \hat{\kappa}^{-9/10} M_8^{-1/10} \left(\frac{r}{10^3 R_{\text{S}}}\right)^{21/20}. \end{aligned} \quad (\text{A3})$$

Using equation (A1) to eliminate  $c_s$  from equation (7) yields the gravitational stability parameter

$$\begin{aligned} Q &= 3(4\pi)^{-3/5} \alpha^{7/10} \beta^{(7b-12)/10} (k_{\text{B}}/m)^{6/5} \sigma^{-3/10} G^{-1} \\ &\quad \times \kappa^{3/10} \dot{M}^{-2/5} \Omega^{9/10} \\ &\approx 0.094 \alpha_{0.01}^{7/10} \beta^{(7b-12)/10} \varepsilon_{0.1}^{2/5} l_{\text{E}}^{-2/5} \hat{\kappa}^{3/10} M_8^{-13/10} \left(\frac{10^3 R_{\text{S}}}{r}\right)^{27/20}. \end{aligned} \quad (\text{A4})$$

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.