Mon. Not. R. Astron. Soc. **363,** 49–56 (2005) doi:10.1111/j.1365-2966.2005.09378.x

Aligning spinning black holes and accretion discs

A. R. King,[1⋆] S. H. Lubow,[2,3] G. I. Ogilvie[2] and J. E. Pringle[1,2,3]

[1]*Theoretical Astrophysics Group, University of Leicester, Leicester LE1 7RH*
[2]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA*
[3]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*

Accepted 2005 June 28. Received 2005 June 19; in original form 2005 April 8

ABSTRACT

We consider the alignment torque between a spinning black hole and an accretion disc whose angular momenta are misaligned. This situation must hold initially in almost all gas accretion events on to supermassive black holes, and may occur in binaries where the black hole receives a natal supernova kick. We show that the torque always acts to align the hole's spin with the total angular momentum without changing its magnitude. The torque acts dissipatively on the disc, reducing its angular momentum, and aligning it with the hole if and only if the angle θ between the angular momenta $\boldsymbol{J}_{\rm d}$ of the disc and $\boldsymbol{J}_{\rm h}$ of the hole satisfy the inequality $\cos\theta > -J_{\rm d}/2J_{\rm h}$. If this condition fails, which requires both $\theta > \pi/2$ and $J_{\rm d} < 2J_{\rm h}$, the disc counteraligns.

Key words: accretion, accretion discs – black hole physics.

1 INTRODUCTION

In a recent paper, Volonteri et al. (2005, see also Madau 2004) consider how the spins of supermassive black holes in galaxies change as the holes grow through both mergers with other holes and gas accretion. Mergers occur at random angles, and when integrated over the mass distribution expected in hierarchical models lead neither to systematic spin-up nor spin-down. Gas accretion, driven for example by minor mergers with satellite galaxies, is likely also to occur at random angles, and thus to be initially retrograde wrt the current hole spin in half of all cases. However, Volonteri et al. argue that gas accretion nevertheless produces systematic spin-up, because a black hole tends to align with the angular momentum of an outer accretion disc on a time-scale typically much shorter than the accretion time-scale for mass and angular momentum (Scheuer & Feiler 1996, hereafter SF96; Natarajan & Pringle 1998). Volonteri et al. (2005) note that this conclusion holds only if most of the accretion takes place through a thin accretion disc. If, instead, accretion is largely via a geometrically thick disc (as happens if most mass accretes at super-Eddington rates), alignment occurs only on the mass accretion time-scale. In this case there would be no net long-term spin-up, assuming successive accretion events were randomly oriented.

Here we address the uncertainties in our current understanding of the evolution of warped accretion discs, and the resulting uncertainties in the alignment mechanisms and time-scales for discs and black holes.

We stress that throughout the paper we neglect the change of the black hole spin as it gains mass from the disc, i.e. we consider time-scales shorter than that for increasing the black hole mass significantly. Thus all the torques we consider arise from the Lense–Thirring effect on a misaligned disc. (These torques are dissipative, and can cause changes in the local accretion rate in the disc. However, these changes are at most by factors of $\sim$2.)

We find that under some conditions counteralignment occurs, contrary to what is usually thought. As mentioned above, initial misalignment must characterize most gas accretion events on to supermassive holes. Stellar-mass black holes accreting from a binary companion may also be misaligned in cases where the hole received a supernova kick at formation. We begin by summarizing briefly what is currently known in various cases.

2 DYNAMICS OF ALIGNMENT

2.1 High-viscosity discs

The dynamics of the process of alignment of black hole and accretion disc is not fully worked out. The best-understood case is when the viscosity is sufficiently high and/or the disc sufficiently thin that its tilt or warp diffuses through it in the way envisaged by Pringle (1992). This is also the most likely case for black hole discs in active galactic nuclei (AGN) and binary systems (Wijers & Pringle 1999; Pringle 1999). SF96 consider the case where a thin, high-viscosity disc is misaligned by a small fixed angle θ, and linearize in θ. In this approximation, the diffusion equation governing the time evolution of the disc surface density remains unchanged to first order, and the analysis assumes that the disc has reached a steady state. In practice this requires that we consider time-scales longer than the inflow time-scale at (and within) some relevant radius $R_{\rm warp}$. SF96 (see also Rees 1978) find that

$$R_{\rm warp} \sim \omega_{\rm p}/\nu. \quad (1)$$

⋆E-mail: ark@star.le.ac.uk

Here $\omega_p = 2GJ/c^2$, where $J = acM(GM/c^2)$ is the angular momentum of the hole (with $-1 < a < 1$), and ν is the kinematic viscosity in the disc.[1]

At radii $R \lesssim R_{\rm warp}$ the spins of the disc elements are aligned with the spin of the hole $\boldsymbol{J}_{\rm h}$. At radii $R \gtrsim R_{\rm warp}$, the spins of the disc elements make an angle $\theta \ll 1$ to $\boldsymbol{J}_{\rm h}$. A spinning black hole induces Lense–Thirring precession in the misaligned disc elements, and the precession rate falls off rapidly with radius ($\propto R^{-3}$). Thus this induced precessional torque acts mainly in the region around radius $R_{\rm warp}$. In Cartesian coordinates, let $\boldsymbol{J}_{\rm h}$ define the z-axis and consider an elemental annulus of the disc with spin $\Delta\boldsymbol{J}_{\rm d}$, which is not parallel to $\boldsymbol{J}_{\rm h}$. Then each such disc annulus feels a torque that tries to induce precession about the z-axis. Integrating over all these torques to get the net torque on the disc (and by Newton's third law the net torque on the hole), we conclude that the net torque can only have components in the (x, y)-plane. In the analysis of SF96, they assume that the disc extends to infinite radius, and thus that the angular momentum of the disc, $\boldsymbol{J}_{\rm d}$, dominates that of the hole. If we assume that the disc is tilted such that $\boldsymbol{J}_{\rm d}$ lies in the (x, z)-plane, then the conclusion of SF96 is that the x- and y-components of the torque are equal in magnitude. The sign of the y-component, which gives rise to precession, depends on the sign of the z-component of $\boldsymbol{J}_{\rm d}$. The x-component affects the degree of misalignment of the disc. It is negative both when the disc and hole are nearly aligned (as above) and also when the disc and hole are nearly counteraligned (i.e. when $\boldsymbol{J}_{\rm h} \cdot \boldsymbol{J}_{\rm d} < 0$). This implies that the net result is to try to align the spins of the disc and the hole. Thus the prediction of the SF96 analysis appears to be that eventually $\boldsymbol{J}_{\rm h}$ and $\boldsymbol{J}_{\rm d}$ should end up parallel. We see below that this conclusion does not hold in general.

2.2 Low-viscosity discs with small-a black holes

If the disc is thick and/or its viscosity low, so that $\alpha < H/R$, then its tilt or warp propagates as a wave rather than diffusing (Papaloizou & Lin 1995). If such a disc is nearly aligned, relativistic precession effects do not align its inner regions with the symmetry plane of the black hole, in contrast to the viscous disc ($\alpha > H/R$). Instead, the disc tilt oscillates (Ivanov & Illarionov 1997; Lubow, Ogilvie & Pringle 2002) with an amplitude proportional to $R^{1/8}/(\Sigma H)^{1/2}$. In the inner regions of black hole accretion discs, this quantity typically increases with decreasing radius (Shakura & Sunyaev 1973; Collin-Souffrin & Dumont 1990). Thus even if the degree of misalignment is small in the outer disc, it can be large in the inner disc, and the angular momentum vector of the matter actually accreted at the horizon can make a large angle to the hole's spin. If the disc and hole are close to being counteraligned, Lubow et al. (2002) show that the inner disc and the hole align.

What interests us here are the disc torques on the hole. In Appendix A we summarize the particular analytic solution for the zero-viscosity case presented in Lubow et al. (2002) and give explicit expressions for the torques both when the disc and hole spins are nearly aligned, and when they are nearly counteraligned. In the absence of viscosity, the torques result in mutual precession of the disc and the hole. In Appendix B we show how this analysis can be modified to take account of a small viscosity, and we investigate the modified torques when this has been introduced. In the nearly aligned case, we find that the magnitudes of the torques are sensitive functions of the exact disc parameters, because of the oscillatory nature of the disc tilt in the inner regions. However, while the effect of this is to introduce great uncertainty into the direction of the component of the torque that gives rise to precession, we find that the sign of the component of the torque that affects the angle between the spins of the hole and the disc is exactly the same as in the viscous case analysis by SF96. We note that in practice this small-viscosity case is most unlikely to hold for the alignment process in most black hole discs in AGN or X-ray binaries (Wijers & Pringle 1999; Pringle 1999) but that warp waves may be important in the centres of these discs, where $H/R \sim 1$.

[1] SF96 assume that the effective viscosities in the disc, ν_1 and ν_2 (Pringle 1992), are comparable.

3 GENERALIZATION

Both sets of analyses reported above assumed that the degree of misalignment was small (that is, the hole and disc were either nearly aligned or nearly counteraligned) and that the disc tilt remained fixed at large radius (that is, the angular momentum of the disc dominates that of the hole). We now consider the physics of the general case where the angle of misalignment is not assumed to be small, and the disc tilt is not assumed to be fixed. The angular momentum of the hole $\boldsymbol{J}_{\rm h}$ is well defined. We shall denote the angular momentum of the disc as $\boldsymbol{J}_{\rm d}$, but note that this is not a well-defined quantity. We discuss the exact meaning of $\boldsymbol{J}_{\rm d}$ in this context below (Section 4.1). From these we construct a third vector representing the total angular momentum, $\boldsymbol{J}_{\rm t} = \boldsymbol{J}_{\rm h} + \boldsymbol{J}_{\rm d}$, which is therefore a constant vector. The torques in which we are interested come about solely because there is a misalignment. We now define the misalignment angle θ by

$$\cos\theta = \widehat{\boldsymbol{J}}_{\rm d} \cdot \widehat{\boldsymbol{J}}_{\rm h}, \tag{2}$$

where the 'hat' indicates a unit vector. We define θ so that $0 \leqslant \theta \leqslant \pi$, with $\theta = 0$ corresponding to full alignment and $\theta = \pi$ corresponding to full counteralignment. The degree of misalignment is measured by the vector $\boldsymbol{J}_{\rm d} \wedge \boldsymbol{J}_{\rm h}$, and so any torques (which are vectors) must depend on this quantity. Note that this vector is zero both for $\theta = 0$ and for $\theta = \pi$. Then in the above discussion (Section 2.1), the y-axis is in the direction of $\boldsymbol{J}_{\rm h} \wedge \boldsymbol{J}_{\rm d}$, and the x-axis is in the direction of $\boldsymbol{J}_{\rm h} \wedge (\boldsymbol{J}_{\rm h} \wedge \boldsymbol{J}_{\rm d})$.

We have argued above, and indeed the analyses of both the high- and low-viscosity cases confirm, that the torque on the hole cannot have a component in the direction of $\boldsymbol{J}_{\rm h}$. Thus the torque must have the form

$$\frac{{\rm d}\boldsymbol{J}_{\rm h}}{{\rm d}t} = -K_1[\boldsymbol{J}_{\rm h} \wedge \boldsymbol{J}_{\rm d}] - K_2[\boldsymbol{J}_{\rm h} \wedge (\boldsymbol{J}_{\rm h} \wedge \boldsymbol{J}_{\rm d})]. \tag{3}$$

Here the first term through the quantity K_1 gives the magnitude and sign of the torque that induces precession. It does not lead to a change in θ. The second term describes the torque that changes the alignment angle θ. Both sets of analyses in the high- and low-viscosity cases show that K_2 is a positive quantity whose magnitude is dependent on the properties of the disc and the hole. Indeed, we show below equation (10) quite generally that K_2 must be positive in the presence of dissipation. In general, of course, K_2 is likely to be a function of θ as well.

If we take the scalar product of this equation with $\boldsymbol{J}_{\rm h}$, we see that ${\rm d}J_{\rm h}^2/{\rm d}t = 0$, so that the magnitude of the spin of the hole remains constant, i.e. $J_{\rm h}$ = constant. Thus the tip of the $\boldsymbol{J}_{\rm h}$ vector moves on a sphere. The total angular momentum $\boldsymbol{J}_{\rm t} = \boldsymbol{J}_{\rm h} + \boldsymbol{J}_{\rm d}$ is of course a constant vector, representing a fixed direction in space. Using this, and the fact that $\boldsymbol{J}_{\rm h} \cdot {\rm d}\boldsymbol{J}_{\rm h}/{\rm d}t = 0$, we see that

$$\frac{\rm d}{{\rm d}t}(\boldsymbol{J}_{\rm h} \cdot \boldsymbol{J}_{\rm t}) = \boldsymbol{J}_{\rm t} \cdot \frac{{\rm d}\boldsymbol{J}_{\rm h}}{{\rm d}t} = \boldsymbol{J}_{\rm d} \cdot \frac{{\rm d}\boldsymbol{J}_{\rm h}}{{\rm d}t}. \tag{4}$$

Using (3) this leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{J}_\mathrm{h}\cdot\boldsymbol{J}_\mathrm{t}) = K_2\left[J_\mathrm{d}^2 J_\mathrm{h}^2 - (\boldsymbol{J}_\mathrm{d}\cdot\boldsymbol{J}_\mathrm{h})^2\right] \equiv A \geqslant 0. \tag{5}$$

Now since both J_h and J_t are constant, this means that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\cos\theta_\mathrm{h}) \geqslant 0, \tag{6}$$

where θ_h is the angle between $\boldsymbol{J}_\mathrm{h}$ and the fixed direction $\boldsymbol{J}_\mathrm{t}$. Thus $\cos\theta_\mathrm{h}$ always increases, implying that θ_h always decreases. This means that the angular momentum vector of the hole always aligns with the fixed direction corresponding to the total angular momentum vector $\boldsymbol{J}_\mathrm{t}$.

To see how $\boldsymbol{J}_\mathrm{d}$ behaves during this process, we consider the quantity A defined above (equation 5). We have, using the definition of $\boldsymbol{J}_\mathrm{t}$ and the fact that J_h = constant, that

$$A = \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{J}_\mathrm{h}\cdot\boldsymbol{J}_\mathrm{t}) = \frac{\mathrm{d}}{\mathrm{d}t}\left(J_\mathrm{h}^2 + \boldsymbol{J}_\mathrm{h}\cdot\boldsymbol{J}_\mathrm{d}\right) = \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{J}_\mathrm{h}\cdot\boldsymbol{J}_\mathrm{d}). \tag{7}$$

Thus

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{J}_\mathrm{h}\cdot\boldsymbol{J}_\mathrm{d}) = A \geqslant 0. \tag{8}$$

We have also that

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}J_\mathrm{t}^2 = \frac{\mathrm{d}}{\mathrm{d}t}\left(J_\mathrm{h}^2 + 2\boldsymbol{J}_\mathrm{h}\cdot\boldsymbol{J}_\mathrm{d} + J_\mathrm{d}^2\right), \tag{9}$$

which implies that

$$\frac{\mathrm{d}}{\mathrm{d}t}J_\mathrm{d}^2 = -2\boldsymbol{J}_\mathrm{d}\cdot\frac{\mathrm{d}\boldsymbol{J}_\mathrm{h}}{\mathrm{d}t} = -2A \leqslant 0. \tag{10}$$

From this we conclude that the magnitude of the disc angular momentum J_d^2 decreases as $\boldsymbol{J}_\mathrm{h}$ aligns with $\boldsymbol{J}_\mathrm{t}$. This is to be expected since, although of course the total angular momentum of the system (hole plus disc) is conserved, the alignment process requires dissipation. Since the magnitude of the spin of the hole remains unchanged, the dissipation must imply a reduction in (the magnitude of) the disc angular momentum. This justifies our statement above that $K_2 > 0$, as a negative K_2 would require energy fed into the disc rotation. Since we consider time-scales short compared with that for accretion, this is not possible.

It is now straightforward to discover when the dissipative torque leads to alignment or to counteralignment of $\boldsymbol{J}_\mathrm{h}$ and $\boldsymbol{J}_\mathrm{d}$. Since the angle between $\boldsymbol{J}_\mathrm{h}$ and $\boldsymbol{J}_\mathrm{d}$ is θ, the cosine theorem gives

$$J_\mathrm{t}^2 = J_\mathrm{h}^2 + J_\mathrm{d}^2 - 2J_\mathrm{h}J_\mathrm{d}\cos(\pi-\theta). \tag{11}$$

Evidently counteralignment ($\theta \to \pi$) occurs if and only if $J_\mathrm{h}^2 > J_\mathrm{t}^2$. This is equivalent to

$$\cos\theta < -\frac{J_\mathrm{d}}{2J_\mathrm{h}}. \tag{12}$$

Thus counteralignment is a possible outcome and requires

$$\theta > \pi/2, \qquad J_\mathrm{d} < 2J_\mathrm{h}. \tag{13}$$

So why did the analysis of SF96, and also that given in the Appendices, imply that the disc and hole always ended up aligned? The reason is that both sets of calculations made the assumption that the outer disc was fixed, that is, that $J_\mathrm{d} \gg J_\mathrm{h}$. In this case we see that counteralignment is forbidden, and alignment must result.

If, instead, $J_\mathrm{d} < 2J_\mathrm{h}$, then for $\boldsymbol{J}_\mathrm{h}$ and $\boldsymbol{J}_\mathrm{d}$ in random directions, counteralignment occurs in a fraction

$$f = \frac{1}{2}\left(1 - \frac{J_\mathrm{d}}{2J_\mathrm{h}}\right) \tag{14}$$

of cases. The disc spin is given by the relevant root of (11), i.e.

$$J_\mathrm{d} = -J_\mathrm{h}\cos\theta + \left(J_\mathrm{t}^2 - J_\mathrm{h}^2\sin^2\theta\right)^{1/2} \tag{15}$$

for alignment, and

$$J_\mathrm{d} = -J_\mathrm{h}\cos\theta - \left(J_\mathrm{t}^2 - J_\mathrm{h}^2\sin^2\theta\right)^{1/2} \tag{16}$$

for counteralignment. In both cases J_d decreases monotonically in time, reaching the final values $J_\mathrm{t} - J_\mathrm{h}$ and $J_\mathrm{h} - J_\mathrm{t}$, respectively.

We can now derive the equation governing the change of θ. From (8) we have

$$A = J_\mathrm{h}\frac{\mathrm{d}}{\mathrm{d}t}(J_\mathrm{d}\cos\theta) = J_\mathrm{h}J_\mathrm{d}\frac{\mathrm{d}}{\mathrm{d}t}(\cos\theta) - A\frac{J_\mathrm{h}}{J_\mathrm{d}}\cos\theta \tag{17}$$

where we have used (10) to write $\mathrm{d}J_\mathrm{d}/\mathrm{d}t = -A/J_\mathrm{d}$. Collecting terms and noting that $A = K_2 J_\mathrm{h}^2 J_\mathrm{d}^2 \sin^2\theta$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(\cos\theta) = K_2 J_\mathrm{h}\sin^2\theta(J_\mathrm{d} + J_\mathrm{h}\cos\theta) \tag{18}$$

and from (15) and (16) we get

$$\frac{\mathrm{d}}{\mathrm{d}t}(\cos\theta) = \pm K_2 J_\mathrm{h}\sin^2\theta\left(J_\mathrm{t}^2 - J_\mathrm{h}^2\sin^2\theta\right)^{1/2}, \tag{19}$$

where $+/-$ corresponds to alignment/counteralignment, respectively. Expanding these two equations about $\theta = 0, \pi$ respectively shows that these equilibria are stable in the two cases. Note that there is no contradiction between the global alignment criterion (12) and the local equation (18): θ does not always decrease monotonically for alignment or increase monotonically for counteralignment.

3.1 A geometrical picture

Although in terms of the algebra given above the alignment process looks somewhat complicated, in terms of the geometry of the situation it is quite simple.

In Fig. 1(a) the initial vector $\boldsymbol{J}_\mathrm{d}$ is represented by the line OA, and the initial vector $\boldsymbol{J}_\mathrm{h}$ by the line OB. Then the total angular momentum $\boldsymbol{J}_\mathrm{t}$ is represented by the line OC, where OACB forms a parallelogram. Thus throughout the subsequent evolution the line OC remains fixed. Since we are just interested in the alignment process, rather than any precession around $\boldsymbol{J}_\mathrm{t}$, we need only consider what happens in the plane defined by OACB. In this plane we have seen that J_h remains constant, and that the effect of the evolution is to align $\boldsymbol{J}_\mathrm{h}$ with $\boldsymbol{J}_\mathrm{t}$. Thus, as shown in the figure, the tip B of the vector OB describes the arc of a circle centred on O and ending up on B′. Once full alignment has occurred, the final vector $\boldsymbol{J}_\mathrm{h}$ lies along OB′. Then in order that total angular momentum be conserved, the tip A of the vector $\boldsymbol{J}_\mathrm{d}$ must move along a corresponding arc, centred on C, and ending at A′. Note that, as this occurs, J_d decreases monotonically. We see that in this case (Fig. 1a) the final vector $\boldsymbol{J}_\mathrm{d}$ lies along OA′, and the disc and hole end up aligned.

In Fig. 1(b) we show exactly the same procedure, but with different initial values for $\boldsymbol{J}_\mathrm{d}$ and $\boldsymbol{J}_\mathrm{h}$. As before, the vector $\boldsymbol{J}_\mathrm{h}$ moves from the initial position OB along an arc centred on O to a final, aligned position, OB′. The total angular momentum $\boldsymbol{J}_\mathrm{t}$, represented by OC, remains fixed. The disc angular momentum vector OA moves along an arc, centred on C, to its final position OA′. However, now, because of the initial values of $\boldsymbol{J}_\mathrm{d}$ and $\boldsymbol{J}_\mathrm{h}$, the disc and the hole end up counteraligned.

In Figs 1(c) and 1(d) we show the same evolution, but in the case $J_\mathrm{d} \gg J_\mathrm{h}$ considered in the analytic calculations of SF96 and in the Appendix. We can see here that both disc and hole always end up aligned, independent of the initial alignment of the hole relative to the disc.

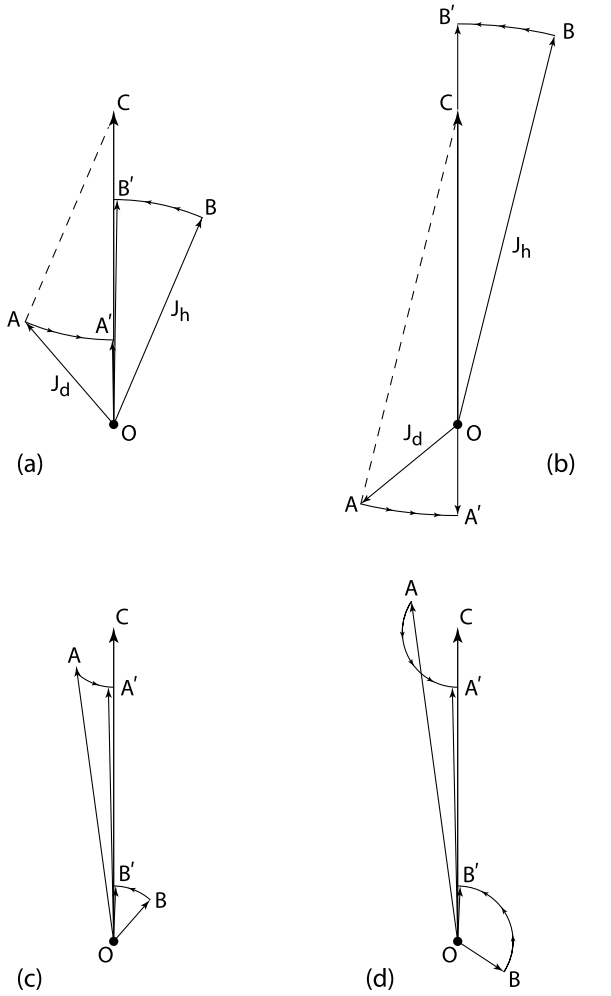

Figure 1. The evolution of hole and disc angular momenta $\boldsymbol{J}_{\rm h}$ (OB) and $\boldsymbol{J}_{\rm d}$ (OA) under the alignment torque, viewed in the plane they define. This plane precesses around the fixed total angular momentum vector $\boldsymbol{J}_{\rm t}$ (OC). (a) A case where the initial angle θ between $\boldsymbol{J}_{\rm h}$ and $\boldsymbol{J}_{\rm d}$ satisfies $\cos\theta > -J_{\rm d}/2J_{\rm h}$: the two angular momenta align. (b) A case where $\cos\theta < -J_{\rm d}/2J_{\rm h}$: the angular momenta counteralign. (c,d) Two cases where $J_{\rm d} \gg J_{\rm h}$ as considered by SF96 for which alignment always occurs.

3.2 Variation of θ with time

Fig. 2 illustrates the evolution of misalignment angle θ for a particular evolution rate, namely for K_2 constant in equation (18). The general behaviour of the solutions is independent of the detailed form of K_2, provided that K_2 is positive (i.e. $\mathrm{d}J_{\rm d}/\mathrm{d}t < 0$).

The highest and lowest curves in Fig. 2 show a monotonic approach to misalignment and alignment, respectively. The middle two curves are close to the bifurcation between the two end states. From equation (18), it follows that non-monotonic behaviour in time can occur for the case of a misaligned disc that begins an approach towards alignment, but in the end becomes misaligned (as seen in the second highest curve of Fig. 2). This situation is realized for an initial state having $J_{\rm d} < -2J_{\rm h}\cos\theta < 2J_{\rm d}$. In such cases, however, the disc never gets close to alignment; that is, θ never drops below $\pi/2$. A bifurcation between end states of alignment/misalignment occurs for cases where initially $J_{\rm d} = -2J_{\rm h}\cos\theta$. A misaligned disc that changes its sense of evolution (i.e. the sign of $\dot{\theta}$) from heading towards alignment to heading towards misalignment does so once $J_{\rm d}$ has been sufficiently reduced. The extent of the required reduction

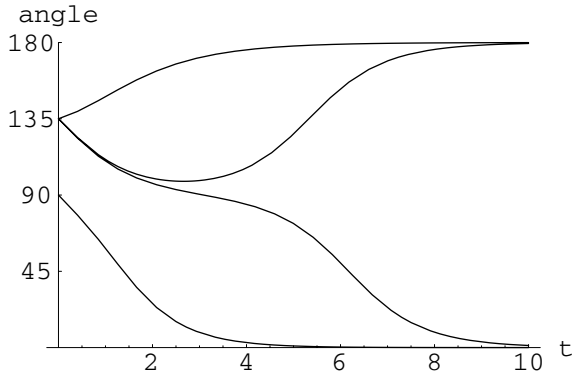

Figure 2. The time evolution of the disc–black hole misalignment angle θ in degrees as a function of dimensionless time, which is normalized by τ, the disc spin-down time-scale for $\theta = 90^\circ$. The evolution is determined by equation (18) with an assumed constant value of $K_2 = 1/(\tau J_{\rm h}^2)$. Initial misalignment angles are $\theta = 135^\circ$ for the uppermost three curves and 90° for the lowest curve. The initial angular momentum ratios from the highest to lowest curves are $J_{\rm d}/J_{\rm h} = 0.5$, 1.40, 1.42 and 0.5, respectively. The curves, from the highest down, correspond to the cases (a), (b), (d) and (c), respectively in Fig. 1. For initial misalignment angle 135°, equation (12) predicts that the transition between long-term alignment and counteralignment occurs when initially $J_{\rm d}/J_{\rm h} = \sqrt{2} \approx 1.414$, as displayed in the middle two curves, which are on opposite sides of the transition. Notice that the second highest curve shows non-monotonic behaviour in time.

is larger for initial values of $J_{\rm d}/J_{\rm h}$ closer to the bifurcation value $-2\cos\theta$, for which misalignment is achieved at a time when $J_{\rm d}(t)$ is zero.

Fig. 3 illustrates the critical case of bifurcation, according to the geometrical picture of the previous subsection. Since $J_{\rm h} = J_{\rm t}$ in this case, we see that, as $\boldsymbol{J}_{\rm h}$ aligns with $\boldsymbol{J}_{\rm t}$, $J_{\rm d}$ approaches zero.

Non-monotonic behaviour in the opposite sense from what is discussed above, i.e. starting with an evolution towards alignment and ending with misalignment, is only possible if $J_{\rm d}$ increases in time. As described in Section 4.2, this can happen because of the uncertainty of what is meant by $J_{\rm d}$ in a real disc. Examples of such cases are described in Section 4.2. These cases involve the complication that $J_{\rm t}$ also changes in time.

3.3 Comparison with numerical simulations

To date there are two numerical simulations in the literature that have $J_{\rm h} > J_{\rm d}/2$ and thus potentially allow the counteralignment

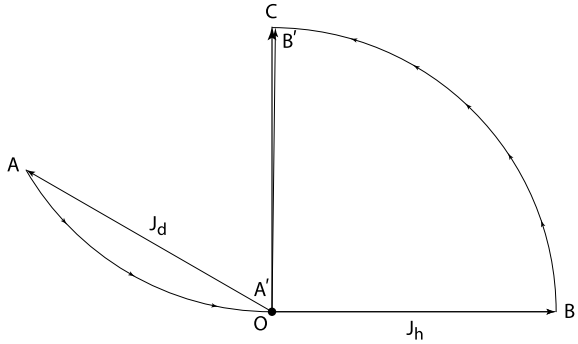

Figure 3. A geometrical picture of the critical case at the boundary between alignment and misalignment, following the notation of Fig. 1. In this case, we have $J_{\rm t} = J_{\rm h}$, with initially $J_{\rm d} = -2J_{\rm h}\cos\theta$. The end state has $\boldsymbol{J}_{\rm h} = \boldsymbol{J}_{\rm d}$ and $J_{\rm d} = 0$.

we predict. Nelson & Papaloizou (2000) perform a quasi-Newtonian three-dimensional smoothed particle hydrodynamic (3D SPH) simulation in which $J_h \gg J_d$. They show that an initially prograde disc does align with the hole, effectively also demonstrating again that $K_2 > 0$ in the presence of dissipation. However, they do not consider any cases with an initially retrograde spin ($\theta > \pi/2$).

Fragile & Anninos (2005) give fully relativistic 3D grid-based simulations, again with $J_h \gg J_d$. However, there is no explicit dissipation in their code ($K_2 \simeq 0$), so any (counter)alignment can only occur on a long time-scale associated with numerical dissipation. There is indeed a hint of counteralignment in the near-retrograde case they discuss.

4 DISCUSSION

We have considered the interaction between a misaligned accretion disc and a rotating black hole. We have argued that the torque between them must have the form (3). The net result is two-fold. First, there is a component (K_1) that causes the disc and hole to precess around the direction of the total angular momentum vector. The direction and rate of precession can depend in a complicated way on the properties of the disc. Secondly, there is a torque (K_2) that, since $K_2 > 0$, acts to align the hole with the total angular momentum without changing its spin rate. This torque acts dissipatively on the disc, and counteraligns or aligns it with the hole according as the conditions (13) hold.

In the high-viscosity case most relevant for black hole discs, the steady disc shape is relatively simple (SF96). In both the co- and counterrotating cases the disc is flat but inclined to the hole at large radii, and flat but aligned with the hole at small radii. The change between the two (the warp) occurs at a radius $R_{\rm warp}$ given by (1), where the rate at which the disc is twisted, i.e. the Lense–Thirring precession rate ω_p/R^3, is balanced by the rate $\sim\nu/R^2$ at which viscous torques can propagate the twist away. Here ν is the viscosity relevant to the process of smoothing disc warp. It corresponds to the viscosity ν_2, introduced by Pringle (1992), which measures the viscosity corresponding to the (R, z)-component of the stress tensor.

The actual dynamics of the various alignment processes are likely to be complicated and need further investigation. As we have seen in Section 3, the torque acts dissipatively on the disc, reducing its angular momentum. If the total disc angular momentum is $\boldsymbol{J}_d$, then the disc eventually aligns with the hole if and only if the angle θ between the angular momenta $\boldsymbol{J}_d$ of the disc and $\boldsymbol{J}_h$ of the hole satisfy $\cos\theta > -J_d/2J_h$. If this condition fails, which requires $\theta > \pi/2$ and $J_d < 2J_h$, the disc eventually counteraligns. This result just follows from the physical nature of the torques, together with the fact that the process is dissipative, so that $K_2 > 0$. Based on these simple physical ideas, we were able to sketch the evolution of the two vectors $\boldsymbol{J}_h$ and $\boldsymbol{J}_d$. However, what we are not able to do, without further consideration of the detailed properties of the disc in the form of the coefficient K_2, is to predict the time-scale on which this happens.

4.1 The meaning of J_d

So far we have been deliberately vague on the precise meaning of the disc angular momentum $\boldsymbol{J}_d$. For an accretion disc we may define the angular momentum vector $\boldsymbol{J}_d(R)$ of the material inside some radius, R. As an example, we consider the disc model for AGN discs given by Collin-Souffrin & Dumont (1990). For this disc model we are interested in the innermost region [called Regime A, which corresponds to region (b) in the disc models of Shakura & Sunyaev (1973)]. In this regime, if we define the radius inside which the angular momentum of the disc equals that of the hole as R_J, so that $J_d(R_J) = J_h$, then it is given in terms of the Schwarzschild radius of the hole, R_S, as

$$\frac{R_J}{R_S} = 3.9\times10^3\left(\frac{\epsilon}{0.1}\right)^{6/19}\left(\frac{L}{0.1L_E}\right)^{-6/19} \times M_8^{-12/19}\left(\frac{\alpha}{0.03}\right)^{8/19} a^{10/19}. \quad (20)$$

Here ϵ is the efficiency of the accretion process (i.e. $L = \epsilon\dot{M}c^2$), L is the accretion luminosity, L_E is the Eddington limit, M_8 is the mass of the black hole in units of 10^8 $M_\odot$, α is the Shakura & Sunyaev (1973) viscosity parameter, and a is the (dimensionless) spin of the black hole.

The time-scale on which this disc radius can communicate with the central disc regions is the viscous time-scale at this radius and is given by

$$t_\nu(R_J) = 1.65\times10^8\left(\frac{\epsilon}{0.1}\right)^{16/19}\left(\frac{L}{0.1L_E}\right)^{-16/19} \times M_8^{6/19}\left(\frac{\alpha}{0.03}\right)^{58/65} a^{14/19}\ \text{yr}. \quad (21)$$

Thus on time-scales longer than this we expect the effective angular momentum of the disc to dominate that of the hole and therefore that on long time-scales the spin of the hole ultimately aligns with that of the disc as in Figs 1(c) and (d).

However, on time-scales less than this, the angular momentum of those parts of the disc which are able to interact with the hole is much less that that of the hole. On these shorter time-scales we might expect the disc evolution to resemble the evolution shown in Figs 1(a) and (b). Thus we have the apparently contradictory possible scenario in which on short time-scales the disc tries to counteralign with the hole, but on long time-scales $t \gg t_\nu(R_J)$ it ends up co-aligning with the hole. This means that the actual disc evolution depends crucially on how the warp is propagated radially by the disc. In other words, we need to be able to predict the nature of the (R, z)-stress denoted by the second viscosity ν_2.

4.2 Warp propagation

If the degree of warping is very small compared to the disc thickness H/R, then Papaloizou & Pringle (1983) showed that, because of resonant effects, the warp stress is much larger than the usual azimuthal stress. If the warp stress is parametrized by α_2 and the usual viscosity by α_1, then they found that $\alpha_2 = 1/(2\,\alpha_1)$. In this case the warp radius can be quite small, i.e. $R_{\rm warp}/R_S \sim$ 10–100 (Natarajan & Pringle 1998). However, once the warp becomes significant, the approximations made in this analysis break down. One possibility then is that the resonant flows become unstable (Gammie, Goodman & Ogilvie 2000), the flow becomes turbulent, and α_2 is reduced significantly until perhaps $\alpha_2 \sim \alpha_1$. If $\alpha_1 = \alpha_2$, which is the assumption made by SF96, we find that

$$\frac{R_{\rm warp}}{R_S} = 990\left(\frac{\epsilon}{0.1}\right)^{1/4}\left(\frac{L}{0.1L_E}\right)^{-1/4} M_8^{1/8} \times \left(\frac{\alpha_1}{0.03}\right)^{1/8}\left(\frac{\alpha_2}{0.03}\right)^{-5/8} a^{5/8}. \quad (22)$$

In reality, it is expected that disc viscosity is generated by magnetohydrodynamic (MHD) turbulence, instigated by the magnetorotational instability (Balbus & Hawley 1991). How this turbulence interacts with a rate of strain in the (R, z)-direction has yet to be fully determined (see, for example, Torkelsson et al. 2000). It is evident that, for finite-amplitude warps and misalignments, in order to estimate the time-scales and mechanisms for warp propagation it will be necessary to undertake numerical simulations.

One of the first goals for such simulations will be to determine whether the simple picture of two viscosities is adequate to a first approximation (cf. Ogilvie 2000). Even in this picture it is clear that a simple α prescription is inadequate. For example, Larwood et al. (1996) show that, when a disc is subject to strong forced precession (as is likely to occur in the inner regions of a tilted disc around a Kerr black hole), the disc may break in the sense that the disc tilt shows a sharp jump at some radius. This can only happen in the diffusive picture if the diffusion coefficient (ν_2) is a function of the (gradient of the) disc tilt angle. However, if this does happen, it enhances the possibility, discussed above, of the inner disc regions being able to counteralign on short time-scales, before eventually co-aligning at later times. If this happened, then accretion on to the hole would act initially [on time-scales $t \ll t_\nu(R_J)$] to spin the hole down, in contradiction of the assumption made by Volonteri et al. (2005).

4.3 Black holes in X-ray binaries

Maccarone (2002) reports that in at least two soft X-ray transient (SXT) binaries (GRO J 1655−40 and SAX J 1819−2525) the observed relativistic jets appear not to be perpendicular to the orbital plane. If the jet directions are indicative of the direction of the spin of the hole, then the most likely explanation is that the misalignment occurred during the formation process of the black hole, and that subsequent evolution has not had time to bring about alignment. This interpretation is interesting in that it points to black hole formation in a (presumably anisotropic) supernova explosion.

From (20) we see that for an $M \sim 10\,\mathrm{M}_\odot$ black hole relevant for such binary systems, the radius R_J is typically much larger than the binary separation. Thus in these systems $J_\mathrm{d} \ll J_\mathrm{h}$. However, the angular momentum in the binary orbit is much larger than that of the hole. Thus the crucial time-scale in these systems is the time-scale on which tidal effects can transfer angular momentum from the binary orbit to the disc. On time-scales shorter than this, the evolution of the disc tilt follows that shown in Figs 1(a) and (b), with the possibility that the disc can counteralign with the hole. Again, numerical simulations are required to provide estimates of the tidal torques for strongly misaligned discs.

Estimates of time-scales from stellar evolution theory can thus give lower limits to the alignment time-scales in SXT binaries. Maccarone (2002) concludes that current theoretical estimates indicate that alignment time-scales are likely to be at least a substantial fraction of the lifetimes of these systems. In any case the long quiescent intervals (10–50 yr or more) in SXT binaries strongly suggest that the inner regions of the disc are either absent or very tenuous. This means that virtually all of the disc mass is far outside the warp radius ($\sim$ a few Schwarzschild radii) and so the alignment torque must be very weak.

ACKNOWLEDGMENTS

ARK acknowledges a Royal Society–Wolfson Research Merit Award. JEP thanks STScI for continued support under their Visitor Program. We thank the referee for helpful remarks concerning numerical simulations.

REFERENCES

Balbus S. A., Hawley J. F., 1991, ApJ, 376, 214
Collin-Souffrin S., Dumont A. M., 1990, A&A, 229, 292
Ivanov P. B., Illarionov A. F., 1997, A&A, 285, 394
Fragile P. C., Anninos P., 2005, ApJ, 623, 347
Gammie C. F., Goodman J., Ogilvie G. I., 2002, MNRAS, 318, 1005
Larwood J. D., Nelson R. P., Papaloizou J. C. B., Terquem C., 1996, MNRAS, 282, 597
Lubow S. H., Ogilvie G. I., Pringle J. E., 2002, MNRAS, 337, 706
Maccarone T. J., 2002, MNRAS, 336, 1371
Madau P., 2004, in Merloni A., Nayakshin S., Sunyaev R., eds, ESO Astrophys. Symp., Growing Black Holes. ESO, Garching, p. 3
Natarajan P., Pringle J. E., 1998, ApJ, 506, L97
Nelson R. P., Papaloizou J. C. P., 2000, MNRAS, 315, 570
Ogilvie G. I., 2000, MNRAS, 317, 607
Papaloizou J. C. B., Pringle J. E., 1983, MNRAS, 202, 1181
Papaloizou J. C. B., Lin D. N. C., 1995, ApJ, 438, 841
Pringle J. E., 1992, MNRAS, 258, 811
Pringle J. E., 1999, in Sellwood J. A., Goodman J., eds, ASP Conf. Ser. Vol. 160, Astrophysical Discs. Astron. Soc. Pac., San Francisco, p. 53
Rees M. J., 1978, Nat, 275, 516
Scheuer P. A. G., Feiler R., 1996, MNRAS, 282, 291 (SF96)
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Torkelsson U., Ogilvie G. I., Brandenburg A., Pringle J. E., Nordlund A., Stein R. F., 2000, MNRAS, 318, 47
Volonteri M., Madau P., Quataert E., Rees M. J., 2005, ApJ, 620, 69
Wijers R. A. M. J., Pringle J. E., 1999, MNRAS, 308, 207

APPENDIX A: ZERO-VISCOSITY DISCS WITH SMALL-*a* BLACK HOLES

Lubow et al. (2002) consider the torque exerted between the central black hole and a disc for the case of zero viscosity. They use the same linearized approximation as SF96, with the disc again assumed steady. At each radius the disc angular momentum is in the direction of the unit vector $\boldsymbol{l}(R) \approx (l_x, l_y, 1)$, where $l_x, l_y \ll 1$. They then use the complex quantity $W(R, t) = l_x + \mathrm{i}l_y$ to describe the shape of the disc. They consider explicitly a simple example in which the amplitude of oscillations is independent of radius, and for which there is a simple analytic solution for $a^2 \ll r = R/(GM/c^2)$. They take the disc thickness to vary as

$$H/R = \epsilon r^{h-1}, \tag{A1}$$

and the disc surface density to vary as

$$\Sigma = \Sigma_0 r^{\frac{1}{4}-h}, \tag{A2}$$

where we require $h > -(1/8)$.

In this case, we set

$$x = \left(\frac{24|a|}{\epsilon^2}\right)^{1/2} \frac{r^{-[h+(1/4)]}}{h+(1/4)}, \tag{A3}$$

where a is the black hole spin parameter with positive (negative) a corresponding to alignment (counteralignment).

If $a > 0$, Lubow et al. (2002) find that

$$W = W_\infty \frac{\cos(x_\mathrm{in} - x)}{\cos x_\mathrm{in}}, \tag{A4}$$

where W_∞ gives the tilt at large radius, x_in corresponds to the inner boundary where the torque vanishes, i.e. $\mathrm{d}W/\mathrm{d}r = 0$, and we need the proviso that $\cos x_\mathrm{in} \neq 0$.

If $a < 0$, it is simple to show that the corresponding expression is

$$W = W_\infty \frac{\cosh(x_{\rm in} - x)}{\cosh x_{\rm in}}. \tag{A5}$$

If T_x and T_y are the components of the torque on the disc, assuming as above that the hole spin is aligned along the z-axis, and writing $T = T_x + {\rm i}T_y$, we find that

$$T = \frac{\sqrt{6}\,{\rm i}\pi}{3}|a|^{1/2}\epsilon\frac{G^2M^2\Sigma_0}{c^2}\int_0^{x_{\rm in}} W\,{\rm d}x. \tag{A6}$$

Then in the case we are considering, for $a > 0$

$$T \propto {\rm i}|a|^{1/2}W_\infty \tan x_{\rm in}, \tag{A7}$$

and for $a < 0$

$$T \propto -{\rm i}|a|^{1/2}W_\infty \tanh x_{\rm in}. \tag{A8}$$

In both cases the coefficient of proportionality is real and positive. Thus, in this case with zero disc viscosity the torque causes a mutual precession on the hole and the disc. However, unlike in the high-viscosity case, if $a > 0$ the sign of the precession is determined by the details of the disc properties.

APPENDIX B: LOW-VISCOSITY DISCS WITH SMALL-a BLACK HOLES

It is not possible to introduce a small constant-α viscosity as a small perturbation of the Lubow et al. (2002) analysis. This is because, at large radii, the Lense–Thirring effect goes to zero, and so the small viscous perturbation does not remain small. However, one can apply perturbation theory for a spatially varying α that takes the form

$$\alpha = \alpha_0 r^{-1} \tag{B1}$$

for the case of small viscosity.

If we write

$$\zeta = \sqrt{1 + (1/3){\rm i}\alpha_0}, \tag{B2}$$

where it is understood that we take the root with positive real part, then in the particular case considered above the solutions become for $a > 0$

$$W = W_\infty \frac{\cos[\zeta(x_{\rm in} - x)]}{\cos(\zeta x_{\rm in})}, \tag{B3}$$

and for $a < 0$

$$W = W_\infty \frac{\cosh[\zeta(x_{\rm in} - x)]}{\cosh(\zeta x_{\rm in})}. \tag{B4}$$

The warps acquire a y-component that is out of phase with respect to W_∞ as a consequence of the viscosity. This phaseshift leads to a net alignment torque (non-zero x-component of torque) on the disc. In the $a > 0$ case, radially oscillatory warped waves can occur provided that $x_{\rm in} > 2\pi$. Such waves are possible for discs that are sufficiently thin ($\sqrt{a}/\epsilon$ sufficiently large). In Fig. B1, we plot W as a function of r for two sets of disc parameters that differ in the value of ϵ. Notice that oscillatory behaviour occurs for the thinner disc.

The torque on the disc is given by

$$T = \frac{\sqrt{6}\,{\rm i}\pi}{3}a^{1/2}\epsilon W_\infty \frac{G^2M^2\Sigma_0}{c^2}\frac{\tan(\zeta x_{\rm in})}{\zeta}. \tag{B5}$$

This torque applies to both positive and negative values of a through analytic continuation. It is straightforward to show that the x-component (real part) of the torque is negative, independently of the sign of a. This result implies that the degree of misalignment decreases in time, as will be discussed more fully in the next subsection.

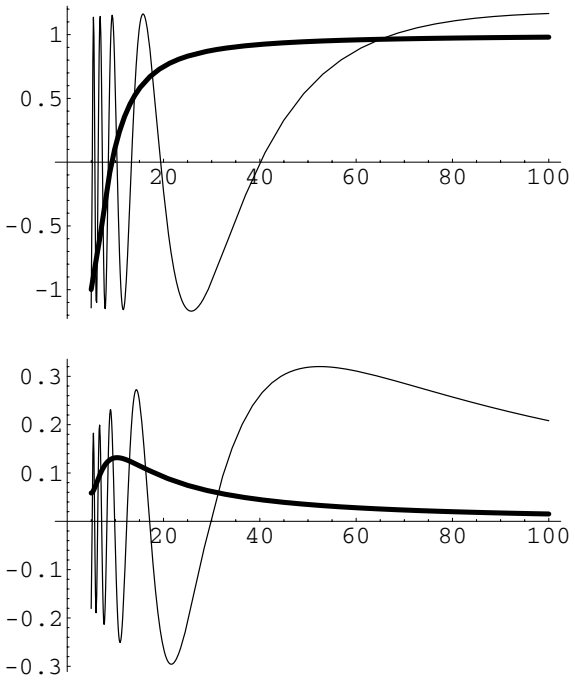

Figure B1. The real (top) and imaginary (bottom) parts of $W(r)/W_\infty$ as a function of radius r (in units of GM/c^2) for two disc–black hole cases. In both cases, we adopt $a = 0.3$ and $h = 1$. The inner radius $r_{\rm in}$ occurs at the marginally stable orbit. The oscillatory (wave-like) solutions are for $\epsilon = 0.01$, while the non-oscillatory solutions (plotted with the heavier lines) occur for $\epsilon = 0.1$. The viscosity parameter $\alpha = \epsilon r_{\rm in}/r$.

We consider discs with inner truncation at the radii of the marginally stable orbits. In the prograde spin case ($a > 0$), the alignment torque Re(T) can undergo large variations as a function of parameters a and ϵ. However, there is a well-defined average value.

We determine the a-averaged value of the alignment torque in the case of $a > 0$ with a fixed value of ϵ. For $a \gg \epsilon^2$ (i.e. $x_{\rm in} \gg 1$), the torque undergoes multiple local peaks in value where $x_{\rm in}(a) = n\pi/2$ for positive integer n. Near such points,

$$\tan(\zeta x_{\rm in}) \simeq \left(c_1 a^{1/2}/\epsilon - (1/2)n\pi + {\rm i}c_2 n\alpha_0\right)^{-1}, \tag{B6}$$

where c_i are real constants of the order of unity and we have ignored variations in the inner disc radius (marginally stable orbit) as a function of a. Since these peaks are spaced in a by an amount $c_3 n\,\epsilon^2$, the a-averaged value of $q(a) = {\rm Im}(\tan[\zeta x_{\rm in}(a)])$ can then be expressed as

$$\langle q(a)\rangle = -\int_{-w}^{w}\frac{{\rm d}z}{z^2+1}, \tag{B7}$$

where

$$z = \left(c_1 a^{1/2}/\epsilon - \frac{1}{2}n\pi\right)\Big/(n\alpha_0 c_2).$$

Here w is the peak width expressed in terms of z, which is inversely proportional to α_0. For small α_0, we can take the integral limits to infinity and we find that q is independent of ϵ and α_0. Consequently, we can approximate the average alignment torque for $a > 0$ by taking

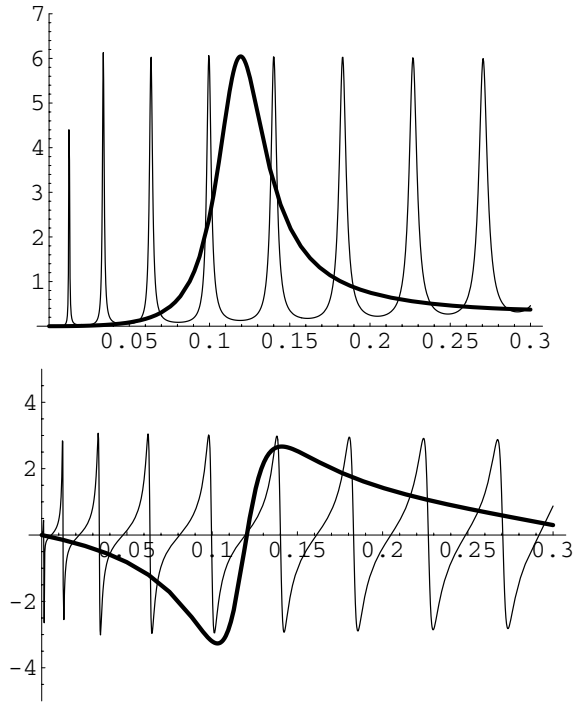

Figure B2. The negative dimensionless alignment torque (top) and the dimensionless precession torque (bottom) on the disc as a function of $a > 0$. The dimensional torques are recovered by multiplying by $\epsilon W_\infty G^2 M^2 \Sigma_0 / c^2$. The disc inner radius $r_{\rm in}$ occurs at the marginally stable orbit for each a. Parameter h is unity. Two cases of disc–black hole systems are plotted, corresponding to the two cases in Fig. 3 (but with varying a). The more strongly fluctuating torques are for $\epsilon = 0.01$, while the smoother torques (plotted with the heavier lines) occur for $\epsilon = 0.1$. The viscosity parameter $\alpha = \epsilon r_{\rm in}/r$.

the average value of $\tan(\zeta x_{\rm in})/\zeta$ in equation (B5) to be a constant of the order of unity that is independent of ϵ and α_0.

For a fixed value of ϵ such that $x_{\rm in} \gg 1$ (so that wave-like behaviour occurs), the a-averaged value of the torque from 0 to a is approximately

$$\langle {\rm Re}(T) \rangle \approx -a^{1/2} \epsilon W_\infty \frac{G^2 M^2 \Sigma_0}{c^2}. \tag{B8}$$

This average torque is then independent of viscosity parameter α_0. The precessional torque is of similar order.

In Fig. B2, we plot the dimensionless torques on the black hole as a function of $a > 0$. Notice that the disc alignment torque is negative, indicating that alignment occurs. The precession rates undergo changes in sign as the spin rate a changes.

For the retrograde spin case ($a < 0$) with $x_{\rm in} \gg 1$, the torque follows from equation (B5),

$$T = -\frac{\sqrt{6}\,\pi}{3}\left({\rm i} + \tfrac{1}{6}\alpha_0\right)\epsilon |a|^{1/2} W_\infty \frac{G^2 M^2 \Sigma_0}{c^2}. \tag{B9}$$

In this case, we see that the ratio of the alignment torque to the precessional torque is $\alpha_0/6$. Furthermore, the precession does not change direction as a function of a, as was found in the case for $a > 0$ (see Fig. B2). Consequently, for low values of the turbulent viscosity parameter, the alignment time-scale can be much longer than the precession time-scale. This situation is unlike the case for $a > 0$, where the two time-scales are comparable.

For both the prograde and retrograde cases, the torque on the disc is exerted where $x \simeq 1$. The torque radius in units of GM/c^2 is then given by

$$r_{\rm T} = \left(\frac{\sqrt{24|a|}}{\epsilon}\right)^{1/[h+(1/4)]}. \tag{B10}$$

Consequently, for the purposes of computing torques, such as in equation (B9), the viscosity parameter α_0 is related to $\alpha = \alpha_0/r$ by

$$\alpha_0 \approx \frac{\alpha\sqrt{|a|}}{\epsilon}, \tag{B11}$$

for $h \approx 1$.

This paper has been typeset from a TEX/LATEX file prepared by the author.