Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 27 July 2011 (MN LaTeX style file v2.2)

The final parsec problem: aligning a binary with an external accretion disc

C.J. Nixon[1], A.R. King[1] and J.E. Pringle[1,2]

[1] *Theoretical Astrophysics Group, University of Leicester, Leicester LE1 7RH UK*
[2] *Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA*

27 July 2011

ABSTRACT

We consider the interaction between a binary system (e.g. two supermassive black holes or two stars) and an external accretion disc with misaligned angular momentum. This situation occurs in galaxy merger events involving supermassive black holes, and in the formation of stellar–mass binaries in star clusters. We work out the gravitational torque between the binary and disc, and show that their angular momenta $\boldsymbol{J}_{\rm b}, \boldsymbol{J}_{\rm d}$ stably counteralign if their initial orientation is sufficiently retrograde, specifically if the angle θ between them obeys $\cos\theta < -J_{\rm d}/2J_{\rm b}$, on a time short compared with the mass gain time of the central accretor(s). The magnitude $J_{\rm b}$ remains unchanged in this process. Counteralignment can promote the rapid merger of supermassive black hole binaries, and possibly the formation of coplanar but retrograde planets around stars in binary systems.

Key words: accretion, accretion discs – black hole physics – galaxies: evolution – stars: formation – planets and satellites: formation

1 INTRODUCTION

Galaxy mergers are commonly thought to be the main mechanism driving the coevolution of galaxies and their central supermassive black holes (SMBH). In such a merger we expect the formation of a SMBH binary in the centre of the merged galaxy. Gravitational waves quickly drive the binary to coalesce if the orbital separation can be shortened to $\lesssim 10^{-2}$pc. The binary may stall at a separation greater than this if the interaction with the merged galaxy is not efficient enough in extracting orbital angular momentum and energy. For the stellar component of the galaxies this occurs at approximately a parsec, creating "the final parsec problem" (Milosavljević & Merritt, 2001). There have been many papers exploring potential solutions to this problem, for example a sling-shot mechanism involving a triple SMBH system (Iwasawa et al., 2006), efficient refilling of the binary loss cone by angular momentum exchange between stellar orbits and a triaxial dark matter halo (Berczik et al., 2006) and also the evolution of the binary with a prograde accretion disc (circumbinary discs: Armitage & Natarajan 2005; MacFadyen & Milosavljević 2008; Lodato et al. 2009; Cuadra et al. 2009 and embedded discs: Escala et al. 2005; Dotti et al. 2007Dotti et al. 2009). In a recent paper (Nixon et al., 2011) we explored the evolution of a binary interacting with a *retrograde* circumbinary accretion disc. We showed that this is more efficient than a prograde disc in removing binary orbital angular momentum and energy. This is simply because there are no orbital resonances between the binary and the disc and thus there is direct accretion of retrograde gas onto the binary.

Here we consider the alignment process between a binary system and an external misaligned accretion disc. This situation can arise in at least two astronomical contexts. First, a merger event between galaxies can produce a SMBH binary in the centre of the merged galaxy, and this or a later accretion event may surround the hole with a disc of accreting gas. A similar situation arises during the formation of stars in a cluster. A binary system may form, but also capture gas into an external disc.

In both of these cases, there is no compelling reason to assume that the binary and disc rotation are initially parallel or even roughly coaligned (cf King & Pringle 2006). As we shall see, the gravitational interaction between the binary and the disc generates differential precession in the disc gas, and thus viscous dissipation. This gives a dissipative torque which vanishes only when the binary and disc angular momenta $\boldsymbol{J}_{\rm b}, \boldsymbol{J}_{\rm d}$ are either parallel or antiparallel. In all such cases, the torque diffuses the tilt or warp through the disc (cf Pringle 1992Pringle 1999; Wijers & Pringle 1999) driving the system to one of these equilibria. The existence of a warp makes the precise definition of disc angular momentum $\boldsymbol{J}_{\rm d}$ quite subtle and we return to this point in the Discussion.

The binary–external disc interaction is very similar to the effect of the Lense–Thirring (LT) precession on an accretion disc around a spinning black hole (Bardeen & Petterson 1975; Pringle 1992; Scheuer & Feiler 1996; Natarajan & Pringle 1998; Armitage & Natarajan 1999; Natarajan & Armitage 1999;

⋆ E-mail: Chris.Nixon@astro.le.ac.uk

Nelson & Papaloizou 2000; Lodato & Pringle 2006 etc) if we replace $\boldsymbol{J}_{\rm b}$ by the hole spin angular momentum $\boldsymbol{J}_{\rm h}$. For some years it was thought that the LT interaction always led to co–alignment (i.e. $\boldsymbol{J}_{\rm b}$ and $\boldsymbol{J}_{\rm d}$ parallel). However King et al. (2005) (hereafter KLOP) showed on very general grounds that counteralignment does occur, if (and only if) the initial angle θ between $\boldsymbol{J}_{\rm d}$ and $\boldsymbol{J}_{\rm h}$ satisfies $\cos\theta < -J_{\rm d}/2J_{\rm h}$, where $J_{\rm d} = |\boldsymbol{J}_{\rm d}|$ and $J_{\rm h} = |\boldsymbol{J}_{\rm h}|$. Scheuer & Feiler (1996) had implicitly assumed $J_{\rm d} \gg J_{\rm h}$ and so enforced co–alignment. With this restriction lifted, King & Pringle (2006King & Pringle (2007) and King et al. (2008) showed that accretion from a succession of randomly–oriented discs leads to spindown of the supermassive black hole, allowing rapid mass growth.

In this paper we examine the alignment process for a binary and an external disc. We show that the argument of KLOP is generic, and that the disc and binary counteralign if and only if $\cos\theta < -J_{\rm d}/2J_{\rm b}$. As a result it is quite possible for SMBH binaries to be surrounded by a completely retrograde disc which strongly promotes coalescence (cf Nixon et al. 2011). In the case of a newly–formed stellar binary, the presence of a counteraligned disc can lead to the formation of planets with retrograde orbits.

2 THE BINARY–DISC TORQUE

We consider a binary system with masses M_1, M_2 and a circular orbit, with the binary angular momentum vector pointing along the z–axis of cylindrical polar coordinates (R, ϕ, z). For simplicity we assume $M_2 \ll M_1$ and place M_1 at the origin, with M_2 orbiting at radius a in the (R, ϕ) plane (our conclusions are not affected by this assumption). The orbit has angular velocity

$$\Omega_{\rm b} = \left[\frac{G(M_1 + M_2)}{a}\right]^{1/2}. \qquad (1)$$

Now we consider a disc particle in an orbit about the binary at radius $R \gg a$. If both the small quantities M_2/M_1 and a/R actually vanished, the particle's orbit would be a circle, with angular velocity $(GM_1/R)^{1/2}$. When these quantities are small but finite they induce various perturbations in the orbit. Some of these perturbations have (inertial–frame) frequency $2\Omega_{\rm b}$ and higher multiples. These are oscillatory, and have no long–term secular effect. Long–term effects on the orbit, and hence eventually on the disc, come from the zero–frequency (azimuthally symmetric $m = 0$) term in the binary potential. This point is considered in more detail in Bate et al. (2000), who considered the related problem of a disc around the primary mass M_1 (i.e. $M_2 \ll M_1$, but $R \ll a$).

Physically this $m = 0$ term is given by replacing the orbiting mass M_2 with the same mass spread uniformly over its orbit, i.e. a ring of mass M_2 and radius a in the (R, ϕ) plane. Adding in the potential from the fixed point mass M_1 at the origin we find the effective gravitational potential experienced by a disc particle as

$$\Phi(R, z) = -\frac{GM_1}{(R^2 + z^2)^{1/2}} - \frac{GM_2}{2\pi}\int_0^{2\pi}\frac{{\rm d}\phi}{r} \qquad (2)$$

where r is the distance between the particle position and a point on the ring at (a, ϕ, z), i.e.

$$r^2 = R^2 + a^2 + z^2 - 2Ra\cos\phi. \qquad (3)$$

We now expand equation 2 in powers of a/R and z/R, keeping terms only up to second order. This gives

$$\Phi(R, z) = -\frac{G(M_1 + M_2)}{R} + \frac{GM_2 a^2}{4R^3} + \frac{G(M_1 + M_2)z^2}{2R^3} - \frac{9GM_2 a^2 z^2}{8R^5} + ... \qquad (4)$$

The orbital frequency Ω of the particle subject to this potential is given by

$$\Omega^2 = \frac{1}{R}\frac{\partial\Phi}{\partial R} \qquad (5)$$

and its vertical oscillation frequency ν by

$$\nu^2 = \frac{\partial^2\Phi}{\partial z^2}, \qquad (6)$$

both evaluated at $z = 0$. The nodal precession frequency is $\Omega_p = \Omega - \nu$ and we find

$$\Omega_{\rm p}(R) = \frac{3}{4}\left[\frac{G(M_1 + M_2)}{R^3}\right]^{1/2}\frac{M_2}{M_1 + M_2}\frac{a^2}{R^2}. \qquad (7)$$

This frequency is very similar to that for LT precession around a spinning black hole (e.g. Scheuer & Feiler 1996), which goes as R^{-3} rather than the $R^{-7/2}$ here. Equation 7 is formally almost identical to the precession frequency found by Bate et al. (2000) for a disc around the primary, although derived for $R \gg a$ rather than $R \ll a$. The same argument as in that paper shows that if the disc and binary axis are misaligned by an angle θ (called δ in Bate et al. 2000) with $0 \leq \theta \leq \pi/2$, the precession frequency is just multiplied by $\cos\theta$. The opposite case with the disc somewhat counteraligned (i.e. $\theta > \pi/2$) is equivalent to the $\theta < \pi/2$ case with the binary angular momentum reversed. But this reversal leaves the precession frequency unchanged, since we are dealing only with the $m = 0$ part of the potential. So for all θ with $0 < \theta < \pi$ the precession frequency is

$$\Omega_{\rm p}(\theta) = \Omega_{\rm p}\,|\cos\theta|\,. \qquad (8)$$

This result differs from the LT case, where the factor $\cos\theta$ appears without modulus signs.

3 CO– OR COUNTER–ALIGNMENT?

We have shown above that the effect of the binary potential on the disc is to induce precession of the disc orbits. This precession is strongly dependent on radius: rings of gas closer to the binary precess faster. The differential precession creates a dissipative torque between adjacent rings of gas tending to make $\theta \longrightarrow 0, \pi$ so that the precession ultimately vanishes.

The precession timescale in the disc increases with radius (cf. equation 7). The torque therefore acts faster at smaller radii to co– or counteralign disc orbits with the binary plane. This leads to the creation of a warp in the disc, where the inner parts are co– or counteraligned and the outer parts are still misaligned (cf. Fig. 1). This warp propagates outwards, eventually co– or counter aligning the entire disc with the binary plane. This effect was solved numerically for discs warped under the LT effect by Lodato & Pringle (2006).

Now we argue as in KLOP that since each ring feels a precession, the resultant back–reaction on the binary is a sum of precessions, which is just a precession. This argument is equivalent to that presented in Bate et al. (2000) who argue that because the binary potential is symmetric about the plane of the binary, the disc–binary torque cannot have a component in the direction of $\boldsymbol{J}_{\rm b}$. Accordingly $\boldsymbol{J}_{\rm b}$ can only precess. These arguments show that we can write the torque on the binary in the same form as the LT–induced torque on a spinning black hole considered in KLOP, i.e.

$$\frac{d\boldsymbol{J}_b}{dt} = -K_1[\boldsymbol{J}_b \wedge \boldsymbol{J}_d] - K_2[\boldsymbol{J}_b \wedge (\boldsymbol{J}_b \wedge \boldsymbol{J}_d)]. \quad (9)$$

Here K_1, K_2 are coefficients depending on disc properties. The first term gives the magnitude and sign of the torque inducing the precession. It does not change the alignment angle θ. The second term describes the torque which changes θ. The same arguments as in KLOP for the LT case, and Bate et al. (2000) for a disc around the primary, show that dissipation in the disc requires K_2 to be a positive quantity. Its magnitude depends on the properties of the disc and the binary. The one difference from the LT case is that the $|\cos\theta|$ dependence means that the sign of the coefficient K_1 can be either positive or negative. But this difference has no effect on the conditions under which the disc and binary co– or counteralign. These are formally identical with the ones for the LT case derived by KLOP, with the binary angular momentum $\boldsymbol{J}_b$ replacing the hole spin angular momentum $\boldsymbol{J}_h$. The process obviously has a different timescale specified by the different magnitude of K_2.

The same arguments as in KLOP now show that the magnitude J_b of the binary angular momentum remains constant, while the direction of $\boldsymbol{J}_b$ aligns with the total angular momentum $\boldsymbol{J}_t = \boldsymbol{J}_b + \boldsymbol{J}_d$, which is of course a constant vector. During this process the magnitude of J_d^2 decreases because of dissipation (KLOP). Counteralignment ($\theta \to \pi$) occurs if and only if $J_b^2 > J_t^2$. By the cosine theorem

$$J_t^2 = J_b^2 + J_d^2 - 2J_bJ_d\cos(\pi - \theta), \quad (10)$$

so this is equivalent to

$$\cos\theta < -\frac{J_d}{2J_b}. \quad (11)$$

Thus counteralignment of a binary and an external disc is possible, and requires

$$\theta > \pi/2, \quad J_d < 2J_b. \quad (12)$$

4 DISCUSSION

So far in this paper we have avoided fully spelling out the meaning of the disc angular momentum $\boldsymbol{J}_d$. This is complicated because the binary torque falls off very strongly with radius, and so a large contribution to the angular momentum in a distant part of the disc may be irrelevant to the alignment process, or affect this process in a time–dependent way (cf Lodato & Pringle 2006). Sections 3 and 4 of KLOP discuss these questions in more detail. Effectively $\boldsymbol{J}_d$ can be thought of as the disc angular momentum inside the warp radius, and therefore a time–dependent quantity.

At early times J_d is small, as only a fraction of the total gas interacts with the binary. Counter–alignment may occur if $\theta > \pi/2$, but at later times, as J_d grows and more gas is able to interact with the binary, alignment eventually happens (when $J_d > 2J_b$). So if $\theta > \pi/2$, even for $J_d > 2J_b$ we expect $\sim 2J_b|\cos\theta|$ of disc angular momentum to counteralign with the binary before the outer disc comes to dominate and enforce coalignment (cf. Lodato & Pringle 2006).

The typical timescale for co– or counter–alignment for a SMBH binary is

$$t_{\rm binary} \simeq \frac{J_b}{J_d(R_w)}\frac{R_w^2}{\nu_2} \quad (13)$$

where R_w is the warp radius, $J_d(R_w)$ is the disc angular momentum within R_w, and ν_2 is the vertical disc viscosity. This is identical to the formal expression for LT alignment of a spinning black hole if we replace the spin angular momentum J_h with J_b (cf

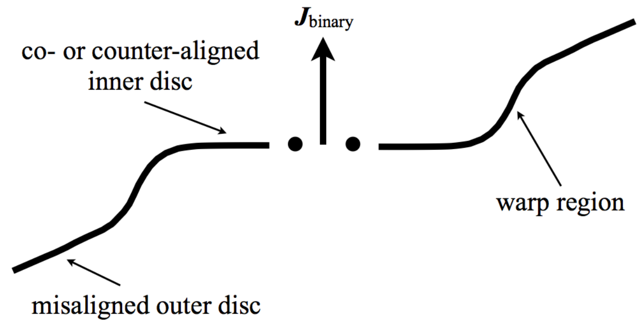

Figure 1. The warped disc shape expected after the inner disc co– or counter–aligns with the binary plane but the outer disc stays misaligned. Eventually the entire disc will co–or counter–align with the binary plane, depending on the global criterion (eqn 12). Note that in practice precession makes the warp non–axisymmetric.

Scheuer & Feiler 1996). The warp radius is given by equating the precession time $1/\Omega_p(R)$ to the vertical viscous time R^2/ν_2 . Inside this radius the precession timescale is short and the disc dissipates and co– or counter–aligns with the binary plane. Outside this radius the disc is not dominated by the precession and so maintains its misaligned plane. The connecting region therefore takes on a warped shape shown in Fig. 1. As time passes the warp propagates outwards and co– or counter–aligns the entire disc with the binary plane.

Approximating the disc angular momentum as

$$J_d(R_w) \sim \pi R_w^2 \Sigma (GMR_w)^{1/2} \quad (14)$$

with Σ the disc surface density and $M = M_1 + M_2$, and using the steady–state disc relation $\dot{M} = 3\pi\nu\Sigma$ we find

$$t_{\rm binary} \sim 3\frac{M_2}{M_1}\left(\frac{a}{R_w}\right)^{1/2}\frac{\nu_1}{\nu_2}\frac{M}{\dot{M}}, \quad (15)$$

where we have also used

$$J_b = M_1M_2\left(\frac{Ga}{M}\right)^{1/2}. \quad (16)$$

Since $\nu_1 < \nu_2$ (Papaloizou & Pringle, 1983), $a \ll R_w$ and $M_2 < M_1$, we see that alignment takes place on a timescale shorter than the mass growth of the central accretor(s).

The timescale (15) is directly analogous to the expression

$$t_{\rm LT} \sim 3a_*\left(\frac{R_s}{R_w}\right)^{1/2}\frac{\nu_1}{\nu_2}\frac{M}{\dot{M}} \quad (17)$$

for alignment under the LT precession, where $a_* < 1$ is the Kerr spin parameter and R_s the Schwarzschild radius of the spinning hole. Evaluating R_w in the two cases we find

$$\frac{t_{\rm LT}}{t_{\rm binary}} \sim \frac{3^{1/2}}{2}\left(\frac{a_*}{M_2/M_1}\right)^{1/2}\left(\frac{a}{R_s}\right)^{1/4}. \quad (18)$$

Thus in general, provided we assume that the ratio ν_1/ν_2 is similar in the two cases and that the hole spin is not rather small ($a_* < (R_s/a)^{1/2}(M_2/M_1)$), then the binary–disc alignment is rather faster than the corresponding process for spinning black holes.

Our result has significant consequences for SMBH binaries. For random orientations, equation 12 shows that initial disc angles leading to alignment occur significantly more frequently than those giving counteralignment only if $J_d >$

$2J_b$. (In the LT case this fact leads to a slow spindown of the hole, because retrograde accretion has a larger effect on the spin, King et al. 2008.) A number of studies (Armitage & Natarajan 2005; MacFadyen & Milosavljević 2008; Cuadra et al. 2009; Lodato et al. 2009) have shown that prograde external discs are rather inefficient in shrinking SMBH binaries and solving the last parsec problem. This is essentially because of resonances within the disc. In contrast, the slightly rarer retrograde events have a much stronger effect on the binary. These rapidly produce a counterrotating but coplanar accretion disc external to the binary, which has no resonances. We note that Nixon et al. (2011) show that the binary gradually increases its eccentricity as it captures negative angular momentum from the disc, ultimately coalescing once this cancels its own. A non–zero binary eccentricity changes the detailed form of the perturbing potential from that in equation 4, but cannot change the precessional character leading to the torque equation (9). Our results remain unchanged, particularly the counteralignment condition (12), apart from minor modifications of the timescale (15).

Thus in a random sequence of accretion events producing external discs, the prograde events have little effect, and the retrograde ones shrink the binary. In particular, a sequence of minor retrograde events with $J_d < J_b$ has a cumulative effect and must ultimately cause the binary to coalesce once the total retrograde $\sum J_d = J_b$. This is important, since the disc mass is limited by the onset of self–gravity to $M_d \lesssim (H/R)\, M_1$ (cf King, Pringle and Hofmann 2008). Coalescence will then occur once the retrograde discs have brought in a total mass M_2, i.e. once a sequence of $\gtrsim (M_2/M_1)(R/H)$ retrograde discs have accreted. For minor mergers this requires at most a few randomly oriented accretion disc events, rising to a few hundred for major mergers ($q > 0.1$).

We note finally that similar considerations apply in planet–forming discs around stellar binary systems, which can also be initially misaligned (Bate et al., 2010). This may offer a way of making retrograde planets in binaries, as recently suggested for ν Octantis (Eberle & Cuntz, 2010).

ACKNOWLEDGMENTS

We would like to thank the anonymous referee for helping to improve the clarity of this work. CJN holds an STFC postgraduate studentship. Research in theoretical astrophysics at Leicester is supported by an STFC Rolling Grant.

REFERENCES

Armitage P. J., Natarajan P., 1999, ApJ, 525, 909
Armitage P. J., Natarajan P., 2005, ApJ, 634, 921
Bardeen J. M., Petterson J. A., 1975, ApJL, 195, L65+
Bate M. R., Bonnell I. A., Clarke C. J., Lubow S. H., Ogilvie G. I., Pringle J. E., Tout C. A., 2000, MNRAS, 317, 773
Bate M. R., Lodato G., Pringle J. E., 2010, MNRAS, 401, 1505
Berczik P., Merritt D., Spurzem R., Bischof H.-P., 2006, ApJL, 642, L21
Cuadra J., Armitage P. J., Alexander R. D., Begelman M. C., 2009, MNRAS, 393, 1423
Dotti M., Colpi M., Haardt F., Mayer L., 2007, MNRAS, 379, 956
Dotti M., Ruszkowski M., Paredi L., Colpi M., Volonteri M., Haardt F., 2009, MNRAS, 396, 1640
Eberle J., Cuntz M., 2010, ApJL, 721, L168
Escala A., Larson R. B., Coppi P. S., Mardones D., 2005, ApJ, 630, 152
Iwasawa M., Funato Y., Makino J., 2006, ApJ, 651, 1059
King A. R., Lubow S. H., Ogilvie G. I., Pringle J. E., 2005, MNRAS, 363, 49
King A. R., Pringle J. E., 2006, MNRAS, 373, L90
King A. R., Pringle J. E., 2007, MNRAS, 377, L25
King A. R., Pringle J. E., Hofmann J. A., 2008, MNRAS, 385, 1621
Lodato G., Nayakshin S., King A. R., Pringle J. E., 2009, MNRAS, 398, 1392
Lodato G., Pringle J. E., 2006, MNRAS, 368, 1196
MacFadyen A. I., Milosavljević M., 2008, ApJ, 672, 83
Milosavljević M., Merritt D., 2001, ApJ, 563, 34
Natarajan P., Armitage P. J., 1999, MNRAS, 309, 961
Natarajan P., Pringle J. E., 1998, ApJL, 506, L97
Nelson R. P., Papaloizou J. C. B., 2000, MNRAS, 315, 570
Nixon C. J., Cossins P. J., King A. R., Pringle J. E., 2011, MNRAS, 412, 1591
Papaloizou J. C. B., Pringle J. E., 1983, MNRAS, 202, 1181
Pringle J. E., 1992, MNRAS, 258, 811
Pringle J. E., 1999, in J. A. Sellwood & J. Goodman ed., Astrophysical Discs - an EC Summer School Vol. 160 of Astronomical Society of the Pacific Conference Series, The General Physics of WARPS. pp 53–+
Scheuer P. A. G., Feiler R., 1996, MNRAS, 282, 291
Wijers R. A. M. J., Pringle J. E., 1999, MNRAS, 308, 207