ACCRETION DISC INSTABILITIES

A.R. KING
Astronomy Group
University of Leicester
Leicester LE1 7RH, U.K.

ABSTRACT. I give a brief discussion of accretion disc instabilities, concentrating mainly on tidal instabilities caused by the presence of a binary companion. The superhumps observed in superoutbursts of SU UMa dwarf novae probably result from the excitation of a resonance in the accretion disc near the 3:1 commensurability with the binary orbit. This resonance can only appear for mass ratios $q = M_2/M_1 < q_{crit} \simeq 0.25 - 0.33$: for larger mass ratios the available resonances are considerably weaker. Application of this picture to other types of binary suggests that the condition $q < q_{crit}$ may be necessary but not sufficient. Further, some cataclysmic systems show phenomena which could be tidal in origin even though the condition evidently fails.

1. Introduction

The presence in certain binary systems of accretion discs conforming to simple theoretical ideas is by now well established. For example Horne and Cook (1985) use the eclipses by the companion to map the surface brightness distribution during outbursts of the dwarf nova Z Cha; they find a disclike object with surface temperature T varying with cylindrical radius as $R^{-3/4}$, exactly as expected for an optically thick $(T \simeq T_{eff})$ steady disc in Keplerian rotation. This result is important, as $T_{eff} \propto R^{-3/4}$ is simply a statement of energy conservation, and independent of the uncertainty over the nature of the viscosity which must drive the accretion within the disc (see e.g. Frank *et al.* 1992 for an introduction to disc theory). The viscosity is usually represented as

$$\nu = \alpha c_S H, \tag{1}$$

195

where c_S and H are the local disc sound speed and semithickness, and α is a dimension-less parameter; simple ideas about turbulent eddy viscosities suggest that α should be $\lesssim 1$. Our lack of understanding of viscosity means that α is an otherwise unconstrained function. Possibly as a result, there is now a vast literature on disc instabilities of the thermal–viscous type which I briefly describe in the next Section. In view of this I will devote most of this review to a rather different type of instability.

A class of dwarf novae, the SU UMa stars, from time to time undergo particularly long outbursts (superoutbursts). During these outbursts a photometric modulation of the optical light (superhumps) with a period a few percent longer than the orbital period always appears. Numerical studies of discs with a particle simulation code (Whitehurst, 1988) do give such a modulation provided that (a) the disc is large enough and (b) the ratio $q = M_2/M_1$ of secondary to primary (white dwarf) masses is less than $\sim 0.25 - 0.33$. The requirement (a) demands a high mass overflow rate from the secondary: this will hold in reality if superoutbursts are mass transfer events, as is widely believed. Indeed the impact of the denser gas stream from the secondary on the disc appears to stimulate the growth rate of the superhumps in the simulations (Whitehurst and King, 1991), suggesting that (c) some external stimulus, such as the impact of the gas stream, may be necessary for the disturbance to grow in a realistic time. The requirement (b), and indeed the physical origin of the superhumps, have up to now been unclear. I review recent progress in this area.

2. Thermal–viscous instabilities

This type of instability has been extensively treated in the literature (see e.g. Frank *et al.* 1992 for an introduction, and Verbunt 1986 for a review) and is discussed in this volume by Wheeler. The nature of the instability was first clarified by Bath and Pringle (1981, 1982). They showed that if a disc annulus with surface density Σ has $\nu\Sigma$ (or equivalently T^4, or the *local* mass flow rate $\dot{M}(R)$) decreasing with Σ, it is unstable on a local thermal timescale. For e.g. $\alpha = $ constant the form of the $T - \Sigma$ relation has two stable branches separated by an unstable branch, caused by the onset of hydrogen ionization at $\sim 10^4$ K (Meyer and Meyer–Hofmeister, 1981). For externally–imposed mass–flow rates corresponding to the unstable range the disc is forced to execute limit cycles, switching between a cool, low–$\dot{M}$ state and a hot, high–$\dot{M}$ state in a continually frustrated effort to reach equilibrium. It is possible to arrange that this local instability (fixed R) becomes a global one involving a large fraction of the disc mass.

It is obvious that this behaviour potentially offers an extremely elegant explanation for dwarf nova outbursts, and conceivably for those of soft X–ray transients also. In particular it naturally implies short outbursts separated by long quiescent intervals, as observed. However, the problem is sufficiently complex that a number of difficulties remain.

1. The detailed behaviour of the UV and optical during an outburst is still not fully reproduced.

2. Some dwarf novae (e.g. SS Cygni, Ricketts *et al.* 1979) show surprisingly high X–ray luminosities during quiescence, implying uncomfortably high central accretion rates $\gtrsim 10^{16}$ g s^{-1}.

3. The value of α has to be manipulated (arbitrarily increased between quiescence and outburst) in order to obtain outbursts of the observed amplitude and duration (Smak, 1984).

4. Present calculations are essentially one–dimensional, although the problem is clearly three–dimensional; this is particularly so for the propagation of thermal fronts.

5. The outburst and quiescence timescales are quite different in dwarf novae (days, weeks) and soft X–ray transients (months, years), yet theoretical modelling seems to cope with both phenomena. It is unclear which of the external parameters (central mass, externally imposed average accretion rate, α) is crucial in accounting for these large differences.

6. The use of the α–parametrization is highly unsatisfactory: a proper theory would *predict* α and give instability behaviour without the need for *ad hoc* assumptions (cf 3. above).

3. Tidal instabilities

The first realisation that the superhumps observed in SU UMa systems were a manifestation of tidal phenomena came from the numerical simulations by Whitehurst (1988). These showed that the accretion disc becomes eccentric and precesses on the superhump period with respect to the binary, so that it and its dissipation pattern appear almost stationary in inertial space. Hirose and Osaki (1990) and Whitehurst and King (1991) independently suggested that this behaviour resulted from the presence in the disc, at small mass ratios, of the 3:1 resonance with the orbital frequency. Recently Lubow (1991a, b) has given a rigorous demonstration of how the eccentricity can grow in a stationary disc (i.e. with superhump period = orbital period). Here I shall follow the heuristic treatment of Whitehurst and King (1991); I begin by explaining the idea of a resonance.

Resonance occurs in a disc when the frequency of radial motion of a particle orbit in the disc is commensurate with the angular frequency of the secondary star as seen by the particle. The former (epicyclic) frequency is $\Omega - \omega$, where Ω is the particle's mean angular frequency around the primary and ω the apsidal precession frequency of its orbit, both measured in a non-rotating frame. The latter (synodic) frequency is $\Omega - \Omega_{\mathrm{orb}}$, where Ω_{orb} is the orbital frequency. Thus resonance requires

$$k(\Omega - \omega) = j(\Omega - \Omega_{\mathrm{orb}}) \tag{2}$$

where j and k are positive integers. (Cases with $k > 1$ arise because of the small but finite eccentricity of the binary orbit.) The apsidal precession rate is zero for periodic (self-closing) orbits in the disc and in any case much smaller than Ω and Ω_{orb} for aperiodic (rosette) orbits, so resonances appear close to commensurabilities of the form

$j : j-k$ between particle and binary frequencies. The strength of a resonance is measured by its growth rate, and goes as e^k, where e is the eccentricity of the particle orbit. Thus the strongest resonances arise at the $j : j - 1$ commensurabilities, followed by $j : j - 2$ etc. Near a given commensurability, the orbits with the largest eccentricity compatible with stability resonate most strongly. Generally, these are aperiodic orbits with non-vanishing precession rates ω. Resonances thus act as filters picking out these extreme orbits provided they can be populated by perturbations.

Even quite extreme aperiodic orbits do not differ greatly from Keplerian circles, and we may measure the typical radii of resonant orbits near the general $j : k$ commensurability as

$$R_{jk} = (GM_1/\Omega_{jk}^2)^{1/3} \tag{3}.$$

Using Kepler's law the binary separation is

$$a = [G(M_1 + M_2)/\Omega_{\rm orb}^2]^{1/3} \tag{4},$$

so we find

$$R_{jk}/a = [(j - k)/j]^{2/3}(1 + q)^{-1/3} \tag{5}.$$

To see what resonances are possible for a given mass ratio, we compare this with the primary's Roche lobe size $R_{\rm Roche}$, given by

$$R_{\rm Roche}/a = 0.49/[0.6 + q^{2/3}\ln(1 + q^{-1/3})] \tag{6}$$

(Eggleton, 1983), and ask when $R_{jk} < 0.9R_{\rm Roche}$, corresponding to the disc's tidal radius. It is readily apparent that the strongest ($k = 1$) resonances can only occur for very small mass ratios: $j = 2$, $k = 1$ requires $q \lesssim 0.025$, with still smaller ratios required for larger j. The 2:1 resonance therefore cannot be typical of SU UMa dwarf novae, which have $q \gtrsim 0.1$. For a possible explanation of the superhumps in SU UMa stars we must turn to the next strongest ($k = 2$) resonances. The smallest radius is for $j = 3$, and is inside $R_{\rm Roche}$ for $q \lesssim 0.3$. Larger values of j require ratios at least as extreme as for the 2:1 resonance. Since larger k-values give weaker resonances, this shows that the strongest resonance possible in the discs of SU UMa stars is likely to be at $j = 3$, $k = 2$, and requires $q \lesssim 0.3$. There is good reason to believe such ratios to be typical in SU UMa systems, and the implied limit on the orbital period agrees well with their absence at periods $\gtrsim 3$ hr (Whitehurst, 1988). Further, particles near the edge of the eccentric precessing discs producing superhumps in the numerical simulations (Whitehurst, 1988) have periods very close to the 3:1 commensurability with the binary orbit. Clearly this offers a promising explanation for the superhump phenomenon, so we examine the detailed behaviour near the resonance.

4. Superhumps

The behaviour of *periodic* orbits ($\omega = 0$) near resonance has been thoroughly studied in the case of the restricted three–body problem by Hénon (1965). Viewed in the reference frame corotating with the binary, resonances cause period–doubling of such orbits, that is, the orbit *circulates* twice. In the sidereal frame, simply–periodic orbits are thus seen to circulate *three* times before closing on themselves, resulting in rosette orbits whose periods are three times those of the simply–periodic orbits just inside the resonance, i.e. $6\pi/\Omega$. A similar result must hold for the aperiodic (precessing) orbits with nonzero ω which we know are the most likely to resonate in accretion discs. Since these have frequencies $\Omega = 3\Omega_{\rm orb} - 2\omega$ just inside the resonance near the 3:1 commensurability (eq. 1), the circulating orbits must have radial periods

$$P_{\rm SH} = 6\pi/\Omega \approx P_{\rm orb}(1 + 2\omega/3\Omega_{\rm orb}) \tag{7}$$

where $P_{\rm orb} = 2\pi/\Omega_{\rm orb}$ is the binary period and we have used the fact that $\omega \ll \Omega_{\rm orb}$. The precession is prograde ($\omega > 0$), so the radial period of the resonant orbits is slightly longer than $P_{\rm orb}$. We cannot calculate by exactly how much as in any case discussion in terms of single–particle orbits becomes unreliable near resonances; since however ω must vanish in the limit of vanishing q we expect the fractional difference between $P_{\rm SH}$ and $P_{\rm orb}$ to be proportional to q. We see that $P_{\rm SH}$ has precisely the properties of the observed superhump period.

The arguments given above show compellingly that superhumps are probably a consequence of a near 3:1 resonance in the accretion discs of SU UMa systems as this produces a radial disturbance with the right period.

Just how the resonance causes the eccentricity of the disc to grow has recently been clarified by Lubow (1991,a, b). He considers perturbations of the form $e^{i(r\theta - s\Omega_{\rm orb}t)}$, where r, s are integers. Thus an eccentric perturbation stationary in inertial space has $r = 1, s = 0$. Neglecting the difference between the superhump and orbital periods this is a simple model of the disturbance giving the superhumps; one can now ask if this $(1,0)$ perturbation grows and produces an eccentric disc. The perturbation produced by the tide of the companion on the disc is a sum of terms of the form (j,j) $(j = 1,2,3,...)$. The interaction with the $(1,0)$ perturbation produces waves of the form $(j-1,j)$: these then interact with the (j,j) tide to produce a $(1,0)$ disturbance which strengthens the original eccentric perturbation. The dominant term in the growth of this perturbation comes from the $j = 3$ tide if the 3:1 resonance is inside the disc.

5. Tidal instabilities in other systems

The discussion above and the numerical simulations by Whitehurst suggest that two conditions are necessary for superhump–type disturbances to appear in an accretion disc in a binary system: (a) the disc must be large enough, i.e. have enough mass near the Roche lobe, and (b) $q < q_{\rm crit} \simeq 0.25 - 0.33$. In addition, (c) it may be that in practice the

disturbance grows too slowly unless stimulated by an external event such as the impact of a gas stream. This might explain why superhumps appear in superoutbursts but not normal outbursts, if the latter are say thermal–viscous disc instabilities as discussed in Section 2. It is interesting to ask first if these conditions are fulfilled in binaries other than the SU UMa dwarf novae, and also to consider possible tidal phenomena where they do not hold.

The orbital periods of SU UMa systems are $\lesssim 3$hr, down to the CV minimum period ~ 80 min. With the usual period–secondary mass relation for lower main sequence stars filling the Roche lobe

$$M_2 \simeq 0.11 P_{hr} M_\odot \tag{8}$$

where P_{hr} is the orbital period measured in hours (e.g. Warner, 1976) requirement (b) implies

$$M_1 > 0.11 P_{hr} M_\odot / q_{crit} \sim 0.36 P_{hr} M_\odot \tag{9}$$

All cataclysmic variables with $P_{hr} < 2$ are SU UMa systems unless they contain a magnetic white dwarf, suggesting that

$$M_1 \gtrsim 0.7 M_\odot \tag{10}$$

for this group. This is probably satisfied by all observed SU UMa systems. (The longest period SU UMa, TU Men, has $P_{hr} = 2.8$ and thus requires $M_1 \gtrsim 1.0 M_\odot$.) It is no surprise that the strongly magnetic AM Her systems in the same period range do not show superhumps, as the field prevents a disc from forming at all.

The only other cataclysmic system in this range is the intermediate polar EX Hya ($P_{hr} = 1.64$). The white dwarfs in intermediate polars are usually assumed to have rather weaker ($\lesssim 10^6$ G) magnetic fields, and there have been explicit claims that EX Hya has an accretion disc (e.g. Rosen et al. 1988). Indeed the system has rather small and widely–separated outbursts, possibly caused by thermal–viscous disc instabilities of the type discussed in Section 2. However there have been no claims of superhump behaviour, even though the condition (10) is probably satisfied. This may be either because the absence of mass transfer events means that even the long quiescent intervals are to short for the disturbance to grow before a minor outburst disrupts the disc, or that even quite weak fields are able to prevent the growth of the instability, or that the field is rather stronger than usually thought (we note that Ferrario et al. 1991 find no kinematical evidence of accretion from a disc from the emission–line behaviour of EX Hya).

The systems closest in many respects to the cataclysmics are the low–mass X–ray binaries (LMXBs). There are no LMXBs in the period range 80 min – 2 hr occupied by the SU UMa systems, but the mass ratio condition (b) is probably satisfied for any period $\lesssim 3.9$ hr (from (9) with a typical neutron star mass $1.4 M_\odot$). This range includes about 7 confirmed LMXBs, 3 of them ultrashort systems with periods less than 1 hr. The mass ratio condition is certainly satisfied in the soft X–ray transient systems

A0620–00, where $M_2 \lesssim 0.8M_\odot$, $M_1 > 3.2M_\odot$ (McClintock and Remillard, 1986), and Cen X–4, where $M_2 \sim 0.1M_\odot$ (McClintock and Remillard, 1990). There have been only two suggestions of superhump–like behaviour in LMXBs (see below). This may be because of the relatively sparse observational coverage of these optically rather faint systems, or possibly because requirements (a) or (c) are not met in steady systems. Models of soft X–ray transients both as disc instabilities (see Wheeler, this volume) and mass transfer events (Hameury *et al.* , 1990, and references therein) have been proposed.

White (1989) proposed a superhump–like modulation to explain the difference between the optical (50.46 min) and X–ray (50.0 min) periods in XB1916-05. As the secondary here must be degenerate and have $M_2 < 0.1M_\odot$ the mass ratio condition (b) is easily satisfied for any reasonable neutron star mass, lending support to the idea. Bailyn (1991) has suggested that an optical quasi–period at about 10.5 hr found in the recent outburst of the soft X–ray transient Nova Muscae 1991 is a superhump modulation. If correct, this would tend to support the idea that such outbursts are mass transfer events analogous to superoutbursts of SU UMa systems.

A competely unexplored field is the possibility of finding superhumps in other semidetached binaries where the mass ratio condition (b) is satisfied. Perhaps surprisingly, there are no obvious candidates: in many cases (e.g symbiotics) the mass ratios q are $\gtrsim 1$, or too close to unity (β Lyrae systems). The best possibilities are probably some of the Algols; the mass ratios satisfy (b) in at least some cases, and the main problem seems to be that discs either cannot form (the accreting star is too large) or are outshone by the component stars.

6. Other possible tidal instabilities

At least 3 cataclysmics not in the SU UMa class show photometric periods which it is tempting to interpret as some kind of tidal disturbance. V795 Her ($P_{hr} = 2.60$, Shafter *et al.* 1990) and 1H0709-36 ($P_{hr} = 2.44$, Tuohy *et al.* 1990) are both squarely in the CV period gap and were both tentatively classified as intermediate polars since the photometric periods differed from the orbital ones. However in both cases the photometric period is the longer of the two. This would be unprecedented for an intermediate polar (it implies a white dwarf spinning more slowly than the orbital rotation) but would be quite normal for a superhump modulation. Neither system shows outbursts, so we would have to postulate that the superhump disturbance can grow if the disc is left undisturbed for long enough; this does appear to happen in the numerical simulations. The superhumps would then be "permanent". The only problem with this rather elegant interpretation (by Dr. H. Ritter, personal communication) is that the eclipses subsequently found in H0709-36 are rather long, and require $q \gtrsim 0.6$ (Bailey, 1990) except for very special combinations of system geometry and inclination. The 3:1 resonance is clearly well outside the disc, and one would have to appeal to the much slower growth rates of other possible resonances (e.g. 4:1, with growth rate $\propto e^3$ in the notation of Section 3). A similar problem arises in interpreting the intermediate polar TV Col: this

has *four* periods of various types (e.g. Ritter, 1990). In addition to the orbital period (5.5 hr) and white dwarf spin period (1911 s) and their beat period (~ 4 d) there is a dominant optical photometric period at 5.2 hr. An interpretation of this last period as some kind of permanent tidal disturbance would explain why it is completely unrelated to the other three periods. Clearly this would be a new type of disturbance, as it is *shorter* than the orbital period. At an orbital period of 5.5 hr TV Col probably has $q \simeq 1$. An examination of the ballistic orbits for $q = 1$ (see Fig. 2 of Whitehurst and King, 1991) suggests that period–doubling can occur at radii lying within the white dwarf's Roche lobe. However the relevant family of orbits lies outside the usual tidal disc radius. A fascinating possibility is that the white dwarf's magnetic field somehow forces matter on to these orbits, which would normally remain unpopulated.

7. Conclusion

The work described here strongly suggests that superhumps are a result of the 3:1 resonance in systems with mass–ratio $q \lesssim 0.3$, provided that the accretion disc grows out to fill most of the primary's Roche lobe. Given these simple criteria, it may be possible to use the presence or absence of superhumps as a way of discovering more about the underlying binary systems. It would be particularly interesting to find superhump–like disturbances in other semidetached binaries.

More work is needed to see if tidal instabilities can grow in discs which do not contain the 3:1 resonance. There are a number of unexplained periodicities in cataclysmic systems for which some kind of tidal explanation appears very attractive, even though the mass–ratio condition for superhumps is not met.

8. Acknowledgments

References

Bailey, J., 1990 *Mon. Not. R. astr. Soc.* **243**, 57.
Bailyn, C.D., 1991, Yale University preprint.
Bath, G.T. and Pringle, J.E., 1981, *Mon. Not. R. astr. Soc.* **194**, 967.
Bath, G.T. and Pringle, J.E., 1982, *Mon. Not. R. astr. Soc.* **199**, 267
Eggleton, P. P., 1983, *Astrophys. J.* **268**, 368.
Ferrario, L., Wickramasinghe, D.T. and King, A.R., 1991, preprint
Frank, J., King, A.R., and Raine, D.J., 1992, *Accretion Power in Astrophysics,* 2nd Ed., Cambridge University Press.

King, A. R., 1988, *Quart. J. R. astr. Soc.* **29**, 1.

Hameury, J.M, King, A.R., and Lasota, J.P., 1990 *Astrophys. J.* **353**, 585.

Hénon, M., 1965, *Ann. d'Astr.* **28**, 992.

Lubow, S.H., 1991a, *Astrophys. J.* **381**, to appear.

Lubow, S.H., 1991b, *Astrophys. J.* **381**, to appear.

Meyer, F. and Meyer–Hofmeister, E., 1981, *Astr. Astrophys.* **104** L10.

McClintock, J.E. and Remillard, R.A., 1986, *Astrophys. J.* **308**, 110.

McClintock, J.E. and Remillard, R.A., 1990, *Astrophys. J.* **350**, 386.

Rosen, S.R., Mason, K.O., and Córdova, F.A., 1988, *Mon. Not. R. astr. Soc* **231**, 549.

Ricketts, M.J., King, A.R. and Raine, D.J., 1979 *Mon. Not. R. astr. Soc.* **186**, 233.

Ritter, H., 1990, *Astr. Astrophys Supp.* **85**, 1179.

Shafter, A.W., Robinson, E.L., Crampton, D., Warner, B. and Prestage, R.M., 1990, in *Accretion–Powered Compact Binaries*, ed C.W. Mauche, Cambridge University Press.

Smak, J., 1984 *Acta Astr.* **34** 161.

Tuohy, I.R., Remillard, R.A., Brissenden, R.J.V. and Bradt, H.V., 1990 *Astrophys. J* **359**, 204.

White, N. E., 1989, *Astr. Astrophys. Rev.* **1**.

Whitehurst, R., 1988, *Mon. Not. R. astr. Soc.* **232**, 35.

Whitehurst, R. and King, A.R., 1991 *Mon. Not. R. astr. Soc.* **249**, 25.