Mon. Not. R. astr. Soc. (1976) **177**, 65–71.

THERMAL INSTABILITIES IN ACCRETION DISCS

J. E. Pringle★

Department of Physics, University of Illinois at Urbana-Champaign,
Urbana, Illinois 61801, USA

(Received 1976 April 29)

SUMMARY

A criterion for the thermal stability of models of accretion discs is presented. It is shown to agree in the appropriate limit with the stability analysis of Shakura & Sunyaev. The two-temperature accretion disc proposed by Shapiro, Lightman & Eardley is shown to be thermally unstable.

1. INTRODUCTION

In a recent publication, Shakura & Sunyaev (1976, henceforth SS2) have presented a comprehensive discussion of the instabilities of the regions of the 'standard' accretion disc (defined by Shakura & Sunyaev 1973, henceforth SS1) in which the dominant opacity is electron scattering. They find that the fastest growing modes are those which correspond to a thermal instability, although the secular instability, discovered by Lightman & Eardley (1974) is also present.

Following Pringle, Rees & Pacholczyk (1973), who discussed the thermal stability of an optically-thin accretion disc radiating by thermal bremsstrahlung, we develop in Section 2 a general criterion for the thermal stability of an accretion disc. In Section 3 we apply our analysis to the 'standard' accretion disc. We consider the effect of radiative transfer in a radial direction neglected in the analysis of SS2. This effect should strictly not be neglected, but its inclusion makes little change to the findings of SS2. Otherwise our results are in full agreement with SS2. We find also that when electron scattering does not provide the dominant opacity, the disc is thermally stable even when the dominant contribution to the pressure is from radiation. In Section 4 we show that the two-temperature disc proposed by Shapiro *et al.* (1976) to explain the spectrum of Cygnus X-1 is thermally unstable. Non-axisymmetric perturbation are discussed in Section 5. Our conclusions are summarized in Section 6.

2. THERMAL INSTABILITY

If the rate at which energy is deposited in the disc per unit area by dissipative processes, Q^+, is not equal to the rate at which the disc can radiate energy per unit area Q^-, then the approximate timescale t_0 on which the disc adjusts to this imbalance at a given radius R is given by

$$t_0(R) = (3/2)\, U c_s^2 / Q^+ \qquad (2.1)$$

★ Permanent address: Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA.

where U is the surface density of the disc, and $c_s \equiv (P/\rho)^{1/2}$. P is the mean (z-averaged) pressure and ρ the mean density. When the disc is stationary,

$$Q^+ = \frac{3GM\dot{M}}{4\pi R^3}\left[1-\left(\frac{R_0}{R}\right)^{1/2}\right]. \tag{2.2}$$

We shall use cylindrical coordinates R, ϕ, z with the disc in the plane $z = 0$. Here (see SS1) $\dot{M}$ is the accretion rate, M is the mass of the central object and R_0 is the inner radius at which the azimuthal stresses may be taken to vanish. By continuity the mean inward radial velocity in the disc V_R is given by

$$V_R = \dot{M}(2\pi RU)^{-1}. \tag{2.3}$$

Defining an inflow timescale by $t_R = R/V_R$ we find

$$t_0 = t_R\left(\frac{c_s}{V_\phi}\right)^2\left[1-\left(\frac{R_0}{R}\right)^{1/2}\right]^{-1} \tag{2.4}$$

where $V_\phi = (GM/R)^{1/2}$ is the azimuthal velocity in the disc.

Throughout these calculations we shall assume that the disc is homogeneous in the z-direction and define the half-width H of the disc in the z-direction by $H \equiv U/2\rho$. This is a good approximation when the disc is supported in the z-direction predominantly by radiation pressure (see SS1) and will suffice in general for our present purposes. Hydrostatic equilibrium in the z-direction implies $H/R \simeq c_s/V_\phi$. Therefore, when the disc is thin and $R \gg R_0$, $t_0 \ll t_R$.

When the disc is stationary ($\dot{M}$ = constant) the azimuthal momentum equation and the continuity equation yield an expression for $W_{R\phi}$, the z-integrated stress,

$$W_{R\phi} = UV_\phi V_R\left[1-\left(\frac{R_0}{R}\right)^{1/2}\right]. \tag{2.5}$$

Making the usual assumption about the stresses,

$$W_{R\phi} = 2\alpha PH \tag{2.6}$$

where α is a constant, we find

$$V_R = \frac{\alpha c_s^2}{V_\phi}\left[1-\left(\frac{R_0}{R}\right)^{1/2}\right]^{-1}. \tag{2.7}$$

Substituting into equation (2.4) we obtain

$$t_0 = (\alpha\omega)^{-1} \tag{2.8}$$

where $\omega = (GM/R^3)^{1/2}$.

The timescale on which the structure of the disc in the z-direction can adjust is $t_z = Hc_s^{-1} \simeq \omega^{-1}$. If $\alpha \ll 1$, $t_z \ll t_0$ and we may assume hydrostatic equilibrium in the z-direction. When $\alpha \sim 1$, however, this assumption is no longer strictly valid. We shall nevertheless assume hydrostatic equilibrium. This does not affect the criterion for thermal stability but may imply that the growth times we find for the instability are underestimates.

We assume initially that the disc is in equilibrium with $Q^+ = Q^-$ and mean temperature $T = T_0$. We investigate small, azimuthally symmetric perturbations of the quantities defining the disc structure about the equilibrium value, which have a sinusoidal radial structure with wavelength $2\pi\Lambda_R \ll R$. We write $T = T_0 + \delta T$. Since the thermal instability timescale is much shorter than the inflow timescale we may take $\delta U = 0$ and consider the self-consistency of this assumption below. This, together with the assumption of hydrostatic equilibrium in the z-direction, implies that we may express all the perturbed quantities in terms of T (or δT).

Let the internal energy per unit area of the disc be E at a given radius R. Then the equation for the rate of change of entropy per unit area may be written

$$\frac{dE}{dt}+2P\frac{dH}{dt} = Q^{+}-Q^{-}. \tag{2.9}$$

The first order equation for δT may then be written

$$\Omega \equiv \frac{1}{\delta T}\frac{d(\delta T)}{dt} = \frac{\dfrac{d\ln Q^{+}}{d\ln T}-\dfrac{d\ln Q^{-}}{d\ln T}}{\left[\dfrac{d\ln E}{d\ln T}+\dfrac{2PH}{E}\dfrac{d\ln H}{d\ln T}\right]}\frac{Q^{+}}{E} \tag{2.10}$$

where the right-hand side is evaluated for $T = T_0$ and full derivatives are made keeping U constant and the equation of hydrostatic equilibrium satisfied.

In general $Q^{+} = 2A(R)W_{r\phi}$ where A is the rate of shear at radius R. Given out assumptions we have $Q^{+} \propto PH \propto c_s^2$. We define

$$\mu = \frac{d\ln Q^{+}}{d\ln T}. \tag{2.11}$$

If β is the ratio of radiation pressure to total pressure in the disc we may write

$$E = (3/2)\,(1+\beta)P.2H. \tag{2.12}$$

Noting that $H \propto c_s$, that is, $d\ln H/d\ln T = \mu/2$, we find that the term in brackets in the denominator of equation (2.10) may be written

$$(1+\beta)^{-1}\left\{\frac{\mu}{6}(8+9\beta-3\beta^2)+3\beta(1-\beta)\right\}. \tag{2.13}$$

We must now evaluate $d\ln Q^{-}/d\ln T$. We may write (see SS)

$$Q_z^{-} = \frac{8}{3}\frac{cm_p}{\sigma}\cdot\frac{\epsilon_r}{U} \tag{2.14}$$

where σ is the relevant absorption or scattering cross-section, $\epsilon_r = aT^4$ the radiation density, c the velocity of light and m_p the mass of a proton. The subscript z denotes that the flux is in the z-direction. We write

$$\lambda \equiv \frac{d\ln Q_z^{-}}{d\ln T}. \tag{2.15}$$

When investigating the modes under consideration we must also take into account radiative transfer in the radial direction. By considering the transfer between two neighboring annuli at temperatures $T_0+\delta T$ and $T_0-\delta T$ a distance $\pi\Lambda_R$ apart, it can be shown approximately that

$$\frac{d\ln Q^{-}}{d\ln T} = \lambda\left[1+2\left(\frac{2H}{\pi\Lambda_R}\right)^2\right]. \tag{2.16}$$

Here, one factor of $2H/\pi\Lambda_R$ comes from consideration of the radiating areas, and the second from consideration of the temperature gradients. To be precise the correction to λ is always of order $(H/\Lambda_R)^2$ and the coefficient (here equal to $8/\pi^2$ and generally of order unity) depends on the structure of the disc and of the assumed perturbation in the z-direction.

We may now rewrite equation (2.10) as

$$\frac{\Omega}{\alpha\omega} = \frac{\mu - \lambda\left[1 + 2\left(\frac{2H}{\pi\Lambda_R}\right)^2\right]}{\frac{\mu}{6}\{8 + 9\beta - 3\beta^2\} + 3\beta(1-\beta)}. \tag{2.17}$$

We conclude that the disc is stable to all modes provided that $\lambda > \mu$. When $\mu > \lambda$, all modes with

$$\Lambda_R > \Lambda_{crit} = \frac{2}{\pi}\sqrt{\frac{2\lambda}{\mu - \lambda}}\,H$$

are (formally) unstable. In fact, when $\Lambda_R \simeq \Lambda_{crit}$ the growth time of the mode becomes comparable to the time for disc material to flow inwards through a distance Λ_R. The assumption $\delta U = 0$ is then no longer valid. The full analysis of such modes neglecting radial heat transfer and assuming electron scattering provides the dominant opacity has been presented by Shakura & Sunyaev (SS2). Inclusion of radial heat transfer in the analysis of SS2 introduces (see below) terms of order $(H/\Lambda_R)^2$ into the dispersion relation (equation 4.14 of SS2) which provide additional damping of importance for modes with $\Lambda_R \sim H$.

3. APPLICATION TO THE 'STANDARD' ACCRETION DISC

We first evaluate μ and λ. Since $Q^+ \propto c_s^2 = c_{sg}^2 + c_{sr}^2$ where c_{sg} is the contribution due to gas pressure and c_{sr} the contribution due to radiation pressure, we find

$$\mu = \frac{2d\ln c_s}{d\ln T} = 2.\frac{3\beta + 1}{2 - \beta}. \tag{3.1}$$

Using equation (2.14) for Q_z^- and writing $\sigma = \sigma_T + \sigma_{ff}$ where σ_T is the Thomson cross-section and $\sigma_{ff} \propto \rho T^{-7/2}$ the free–free (or Kramers') absorption cross-section we obtain

$$\lambda = 4 + \frac{\gamma}{2}(\mu + 7) \tag{3.2}$$

where $\gamma = \sigma_{ff}/\sigma$ in the equilibrium solution. The criterion for stability, $\lambda > \mu$, can then be written

$$\beta < \frac{12 + 16\gamma}{20 + \gamma}. \tag{3.3}$$

In particular in the case $\gamma = 0$ considered by SS2 we confirm the stability criterion $\beta < 3/5$. When $\beta > 3/5$, and $\Lambda_R \gg H$, equation (2.17) becomes

$$\frac{\Omega}{\alpha\omega} = \frac{6(5\beta - 3)}{8 + 51\beta - 3\beta^2} \tag{3.4}$$

in agreement with the asymptotic behaviour of the thermally unstable modes discovered by SS2.

As we mentioned above, the analysis of SS2 is modified slightly by the inclusion of radiative transfer in the z-direction. Using the approximation of equation (2.16), the dispersion relation (equation 4.14 of SS2) becomes

$$\left(\frac{\Omega}{6\alpha\omega}\right)^2 A(\beta)+\left(\frac{\Omega}{6\alpha\omega}\right)\left[B'(\beta)\left(\frac{H}{3\Lambda_R}\right)^2+3-5\beta\right]+\left(\frac{H}{3\Lambda_R}\right)^2(5-3\beta)\left[1+2\left(\frac{2H}{\pi\Lambda_R}\right)\right]^2=0 \tag{3.5}$$

where $A(\beta) = 8+51\beta-3\beta^2$ and

$$B'(\beta) = B(\beta)+9[8(2-\beta)/\pi^2].$$

Note that $A(\beta)$ and $B(\beta)$ are as defined in SS2. Thus as expected the overall structure of the solutions found in SS2 remain unchanged but stability sets in at slightly larger values of Λ_R/H.

We note in addition that if $\gamma > (8/15)$ equation (3.3) implies that all modes are thermally stable, since $\beta \leqslant 1$. For this reason the cooler accretion discs to be expected around supermassive black holes are likely to be thermally stable even when their luminosities are close to the Eddington limit ($\beta \simeq 1$).

4. FURTHER APPLICATIONS

The conclusion of Pringle *et al.* (1973) that a disc emitting optically-thin bremsstrahlung is thermally unstable follows directly from the above. In this case $Q_z^- \propto \rho^2 T^{1/2} H \propto T^{1/2} c_s^{-1}$ and, therefore, $\lambda \leqslant 0$. Since $\mu > 0$ this situation is unstable. Moreover, an optically-thin disc radiating by a process which may be written in the form $Q_z^- \propto \rho^2 H f(T)$ is thermally unstable unless $d \ln f/d \ln T$ is sufficiently large. An optically-thin collisionally heated gas with $T \gtrsim 10^4$ K is always thermally unstable.

Lightman & Shapiro (1975) have argued that the 'cool disc model' of SS1 is incapable of producing the hard ($\sim$100 keV) X-rays from Cyg X-1. For this reason Shapiro *et al.* (1976) have constructed a two-temperature accretion disc model for Cyg X-1 in which the ions are at $T_i \sim 10^{11}$ K are heated by dissipative processes and are cooled by Coulomb interaction with the electrons. These are at $T_e \sim 10^9$ K and in turn cool by Comptonizing ambient photons which emerge at the required X-ray energies. We show here that such a disc is thermally unstable.

In the two-temperature disc of Shapiro *et al.* (1976) ion pressure dominates and hence $c_s \propto T_i^{1/2}$, $\mu = 1$. The rate at which the ions lose energy to the electrons (per unit area) is

$$Q_i^- = (3/2)H\rho\nu_E k(T_i - T_e)/m_p \tag{4.1}$$

where $\nu_E \propto \rho T_e^{-7/2}$ (Spitzer 1962). We equate this to the rate Q_e^- at which electrons lose energy which for unsaturated Comptonization, required by Shapiro *et al.*, is given by

$$Q_e^- \propto \rho H T_e U_r. \tag{4.2}$$

Here U_r is the radiation density of a copious supply of soft (X-ray) photons, which we assume does not change on timescales short compared to Ω^{-1}. Taking

$T_e \ll T_i$, we find $T_i \propto T_e^5$ and hence $\lambda = 1/5$. Since $\mu = 1 > \lambda$ the two-temperature disc proposed by Shapiro *et al.* is thermally unstable.

5. NON-AXISYMMETRIC PERTURBATIONS

We now consider perturbations with wavelengths $2\pi\Lambda_\phi$ in the azimuthal direction, ($\Lambda_\phi \ll R$) and $2\pi\Lambda_R$ in the radial direction (as above). If such modes are convected by the azimuthal flow $V_\phi(R)$, they are destroyed on a shearing time

$$t_s = \frac{\Lambda_\phi}{\Lambda_R}\,\omega^{-1}. \tag{5.1}$$

Therefore, for the thermal instability to occur before the mode is mixed azimuthally by differential rotation, we require $t_s \gg t_0$, or

$$\Lambda_\phi \gg \alpha^{-1}\Lambda_R. \tag{5.2}$$

At first sight it seems possible that azimuthal temperature—and hence pressure—gradients might be able to stabilize some modes against shear. The relative azimuthal velocity of a node and of its neighbouring antinode in a radial direction is $(1/2)\,\pi\Lambda_R\Omega$. The greatest azimuthal velocity, relative to V_ϕ, that pressure gradients can give rise to is c_s. For this reason stabilization with regard to shear can only occur for modes with $\Lambda_R \simeq c_s\Omega^{-1} \simeq H$. We have, however, already noted (see SS2, Fig. 1) that such modes are stable.

We conclude that spiral modes ($\Lambda_\phi \gg \Lambda_R$) can display thermal instability but that, hot spots (that is, modes with $\Lambda_\phi \sim \Lambda_R$) are unlikely to be caused by the thermal-type instabilities we have discussed. Modes with $\Lambda_\phi \sim \Lambda_R$ are mixed azimuthally before the thermal instability can grow.

6. SUMMARY

We have derived a simple criterion for the thermal stability of an accretion disc in the approximation that an instability grows sufficiently rapidly that radial flow can be neglected. The criterion is in complete agreement with the full analysis of the central regions of the standard accretion disc presented by Shakura & Sunyaev (SS2). We have extended their analysis by including the effects of radiative transfer in the radial direction and show that this modifies their analysis to a small degree. We have also shown that when the dominant opacity is not electron scattering, for example in the relatively cool accretion discs around massive objects, the standard accretion disc is thermally stable even when radiation pressure provides the dominant opacity.

Application of our stability criterion to the two-temperature disc model of Shapiro *et al.* (1976) shows that it is thermally unstable. We note that the calculations of Shakura & Sunyaev (SS1) indicated that a spectrum similar to that of Cyg X-1 could indeed be obtained by Comptonization processes when $\alpha \sim 1$ and when the luminosity is close to the Eddington limit. In addition these authors (SS2) argue that such processes could become important in the non-linear regime of the thermal instability, so that a Cyg X-1 type spectrum could be produced at lower values of α and of the luminosity. Thus, the model proposed by Shapiro *et al.* could be correct in a time-averaged sense.

ACKNOWLEDGMENTS

It is a pleasure to thank D. Q. Lamb, C. J. Pethick and D. Pines for valuable discussions. Research was supported by National Science Foundation Grant PHY75-08790.

REFERENCES

Lightman, A. P. & Eardley, D. M., 1974. *Astrophys. J. Lett.*, **187,** L1.
Lightman, A. P. & Shapiro, S. L., 1975. *Astrophys. J. Lett.*, **198,** L1.
Pringle, J. E., Rees, M. J. & Pacholczyk, A. G., 1973. *Astr. Astrophys.*, **29,** 179.
Shakura, N. I. & Sunyaev, R. A., 1973. *Astr. Astrophys.*, **24,** 337.
Shakura, N. I. & Sunyaev, R. A., 1976. *Mon. Not. R. astr. Soc.*, **175,** 613.
Shapiro, S. L., Lightman, A. P. & Eardley, D. M., 1976. *Astrophys. J.*, **204,** 187.
Spitzer, L., 1962. *Physics of fully ionized gases*, Wiley, New York.