Accretion disc theory since Shakura and Sunyaev

A. King

Theoretical Astrophysics Group, University of Leicester, Leicester LE1 7RH, U.K.
e-mail: ark@star.le.ac.uk

Abstract. I briefly review the progress of accretion disc theory since the seminal paper of Shakura and Sunyaev.

Key words. Accretion disc - Cataclysmic variables - Dwarf novae

1. Introduction

Discs are a natural occurrence in astrophysical systems whenever they have significant angular momentum. Astronomers had discussed discs of various kinds through the twentieth century. However the paper of Shakura and Sunyaev (1973) transformed the subject, partly because it unified concepts already discussed, and partly through technical innovation.

The most important results of the paper were

(a) the condition for an accretion disc be thin, i.e. to have scale height $H$ much less than disc radius $R$: Shakura & Sunyaev (1973) showed that the conditions ‘thin’, ‘efficiently cooled’ and ‘Keplerian’ are precisely equivalent: if one of them fails, so do the other two.

(b) it is perfectly possible for an accretion object to be supplied with mass at a rate that ultimately would produce a luminosity above the Eddington limit. Shakura & Sunyaev (1973) suggested that much of the excess would be blown away by radiation pressure at the radius where this luminosity was first reached.

We now know that this is probably what occurs in SS433, and most, if not all, ultraluminous X-ray sources (ULXs) (cf Begelman et al., 2006; Poutanen et al., 2007).

(c) the effective temperature profile of a steady thin disc goes as $T(R) \propto R^{-3/4}$, and this result is independent of the mechanism for gas to lose angular momentum and spiral inwards.

This effective temperature profile is now well attested, particularly by observations of CVs. The overall stretched–out blackbody continuum spectrum agrees with it, and more directly, the surface brightness distribution measured during accretion disc eclipses in CVs also agrees: eclipses are broad and shallow at long wavelengths and deep and narrow at short ones.

(d) the mechanism for angular momentum removal (Shakura & Sunyaev (1973) called it ‘viscosity’) may have a magnetic origin.

(e) in all cases one can parametrize the (kinematic) viscosity as $\nu = \alpha c_s H$, where $c_s$ is the local sound speed, and $\alpha$ a quantity of order unity.

This ‘alpha prescription’ has the great virtue of neatly separating the ‘vertical’ and
‘horizontal’ disc structure. However it is vital to realise that accretion disc theory is still incomplete, since we do not know \( v(x, t) \). Thus viscosity plays a similar role in accretion disc theory to that played by nuclear burning in stellar evolution theory in the early 20th century.

2. Progress

Just as astronomers were nevertheless able to make some progress with stellar structure theory despite not understanding nuclear burning (cf Eddington’s book *The Internal Constitution of the Stars*), theorists and observers have managed to understand how accretion discs behave in some situations. Much of this understanding has come because the disc diffusion equation

\[
\frac{\partial \Sigma}{\partial t} = \frac{3}{R} \frac{\partial}{\partial R} \left( \frac{R^{1/2}}{v \Sigma R^{1/2}} \right)
\]

(1)
defines a viscous timescale

\[ t_{\text{visc}} \sim \frac{R^2}{v} \]

(2)

which can be rewritten using the alpha-prescription as

\[ t_{\text{visc}} \sim \frac{1}{\alpha} \left( \frac{R}{H} \right)^2 t_{\text{dyn}} \]

(3)

where \( t_{\text{dyn}} = (R^3/GM)^{1/2} \) is the local dynamical time.

Thus for example we now know that superhumps result from the presence of the orbital 3:1 resonance within a sufficiently large accretion disc (corresponding to a fairly extreme mass ratio in a CV: cf Whitehurst & King, 1991; Lubow, 1991, 1992).

Here \( t_{\text{visc}} \) governs the overall timescale, but variations of it are unimportant for understanding superhumps. Similarly, the possibility of disc instabilities (see the talk by Lasota at this meeting) requires one to imagine only two things: (i) disc structure differs radically if hydrogen is predominantly ionized or not, and (ii) hotter regions have higher viscosity and evolve faster than cool ones. Perhaps suprisingly, the behaviour of the instability in a disc strongly irradiated by the central (X-ray) source is qualitatively much easier to understand (King & Ritter, 1998), since the irradiation traps the disc in the hot state, allowing a pure viscous-timescale decay.

These are particularly simple – pure exponentials – if the disc is small enough that the central source keeps it hot at all radii (the so-called FRED – fast rise, exponential decay systems).

These cases and others give straightforward estimates of \( \alpha \) as \( \alpha \sim 0.1 - 0.4 \). By now one can arrive at some kind of understanding of how almost any pattern of light-curve behaviour can be understood using this picture of disc instability, and possibly allowing for some mild intrinsic variability (e.g. magnetic spots) on the secondary (King & Cannizzo, 1998).

In summary we can say that disc ‘theory’ works quite well provided that we assume these values of \( \alpha \).

However this apparent success comes at a double price. First, it is entirely ad hoc – it works (rather like the Old Quantum Theory) because of a series of fudges and empirical rules. Second, and more seriously, it means that we are entirely unable to predict, or confirm or deny, global changes in disc structure, such as whether thin disc accretion can make a transition to advection-dominated (ADAF) flow, or just how an accretion disc creates and powers a jet at its centre. This is analogous to the inability of pre-nuclear stellar structure theory to predict or explain supernovae.

3. The answer?

In searching for the true mechanism for angular momentum transport we need to remember the distinctive feature of accretion discs, that they simultaneously obey

\[
\frac{\partial}{\partial R} (R^2 \Omega) > 0, \quad \text{and} \quad \frac{\partial \Omega}{\partial R} < 0
\]

(4)

where \( \Omega \) is the local angular velocity (the Kepler value \( (GM/R^3)^{1/2} \) in a thin disc). That is, angular momentum increases outwards, but angular velocity decreases outwards. The first property tell us that discs are stable against axisymmetric perturbations (Rayleigh criterion), so removing a large number of candi-
date mechanisms. Indeed, most purely hydro-
dynamical mechanisms are sensitive to the gra-
dient of angular momentum rather than veloc-
ity, and so would if anything transport angular
momentum inwards.

The dragging of a magnetic fieldline an-
chored in an accretion disc on the other hand
is sensitive to the angular velocity gradient,
and does offer a promising candidate (Balbus
& Hawley 1991) for the transport mechanism.
But although this is clear, actually calculating
the full effect of this process is a formidable
challenge. In principle one has to solve the
full disc structure self–consistently, describing
gas motions in full 3D time–dependent MHD.
Most theoretical effort so far has gone into try-
ing us numerical simulations to quantify the
viscous transport (naively, estimate $\alpha$) in the
so–called ‘shearing box’ approximation. Here
one considers a corotating Cartesian box, plus
tidal gravity and Coriolis terms. This is a much
more tractable problem, but has inevitable lim-
itations. Most obviously, the scale of the box
is only $\sim H$, so the simulations are only sen-
sitive to high wavenumbers $\sim R/H$, and hence
small–scale magnetic fields.

The results of this procedure are mixed (see
King, Pringle & Livio, 2007 for a review).
Fully–ionized shearing–box simulations
tend to give rather small values $\alpha \sim 0.02$,
unless a vertical magnetic field is imposed
from the start. Worse, there is some indica-
tion that the value of $\alpha$ is resolution–dependent
(Fromang & Papaloizou, (2007) in the sense
that $\alpha$ decreases as the numerical resolution
of a simulation is increased. So although MHD
effects are probably the basis of accretion disc
viscosity, current simulations have not con-
vincingly shown this, still less that the effect
is large enough to account for observations.
Evidently this problem will require global disc
simulations, and so even more powerful com-
puters.

4. Conclusions

Our current ad hoc picture of accretion discs is
based on the semi–empirical ideas of Shakura
and Sunyaev (1973). It works reasonably well
in a number of areas, provided that we as-
sume $\alpha \sim 0.1 - 0.4$. However it is still
ad hoc, and is unable to predict or confirm
global changes of disc structure. The real ba-
sis of accretion disc viscosity is probably mag-
etic, as suggested by Balbus and Hawley
in 1991. Attempts to demonstrate this with
shearing–box simulations produce viscosities
which are too weak (i.e. $\alpha \lesssim 0.02$) com-
pared with observed constraints, and may even
be resolution–dependent. Balbus and Hawley’s
paper appeared less than 20 years after Shakura
& Sunyaev (1973), but 20 years further on the
huge complexity of the viscosity problem has
markedly slowed practical progress. We are
still a long way from a theory of accretion discs
with real predictive power.

Acknowledgements. I thank many colleagues, par-
ticularly Jim Pringle, for insight and illumination on
these subjects over the years. My thanks to the or-
organisers for the superb hospitality and organisation
of the meeting. Research in theoretical astrophysics
at Leicester is supported by an STFC Rolling Grant.

References

Begelman, M. C., King, A. R., Pringle, J. E.
King, A. R., Pringle, J. E., Livio, M. 2007,
MNRAS, 376, 1740
Poutanen, J., Lipunova, G., Fabrika, S.,
Butkevich, A. G., Abolmasov, P. 2007,
MNRAS, 377, 1187