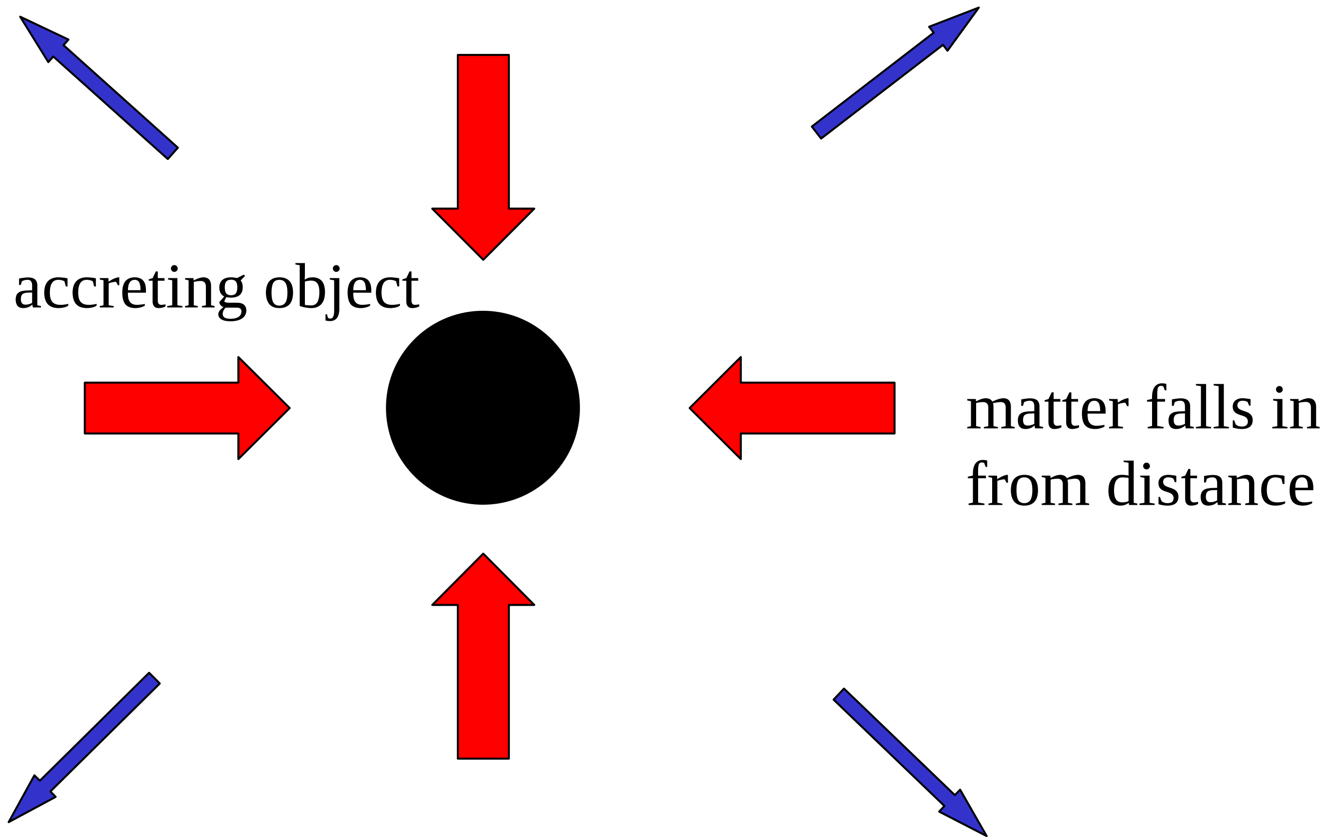
Accretion Power in Astrophysics

Andrew King

Theoretical Astrophysics Group, University of Leicester, UK

accretion = release of gravitational energy from infalling matter



energy released as electromagnetic
(or other) radiation

If accretor has mass M and radius R, gravitational energy release/mass is

$$\Delta E_{acc} = \frac{GM}{R}$$

this *accretion yield* increases with *compactness* *M/R:* for a given *M* the yield is greatest for the smallest accretor radius R

e.g. for accretion on to a neutron star $(M = M_{sun}, R = 10km)$

$$\Delta E_{acc} = 10^{20} erg / gm$$

compare with nuclear fusion yield (mainly H→ He)

$$\Delta E_{nuc} = 0.007c^2 = 6\times10^{18} erg / gm$$

Accretion on to a black hole releases significant fraction of rest—mass energy:

$$R \approx 2GM / c^2 \Rightarrow \Delta E_{acc} \approx c^2 / 2$$

(in reality use GR to compute binding energy/mass:
typical accretion yield is roughly 10% of rest mass)

This is the most efficient known way of using mass to get energy:

accretion on to a black hole must power the most luminous phenomena in the universe

$$\boxed{L_{acc} = \frac{GM}{R}\dot{M} = \eta c^2 \dot{M}}$$

Quasars: $L \approx 10^{46} erg/s$ requires $\dot{M} = 1M_{sun}/yr$

X—ray binaries: $L \approx 10^{39} erg/s$ $10^{-7} M_{sun}/yr$

Gamma—ray bursters: $L \approx 10^{52} erg/s$ $0.1 M_{sun}/\sec$

NB a gamma—ray burst is (briefly!) as bright as the rest of the universe

Accretion produces radiation: radiation makes pressure – can this inhibit further accretion?

Radiation pressure acts on electrons; but electrons and ions (protons) cannot separate because of Coulomb force. Radiation pressure force on an electron is

$$F_{rad} = \frac{L\sigma_T}{4\pi c r^2}$$

(in spherical symmetry).
Gravitational force on electron—proton pair is

$$F_{grav} = \frac{GM(m_p + m_e)}{r^2}$$

$$(m_p >> m_e)$$

thus accretion is inhibited once $F_{rad} \geq F_{grav}$, i.e. once

$$\boxed{L \geq L_{Edd} = \frac{4\pi G M m_p c}{\sigma_T} = 10^{38} \frac{M}{M_{sun}} erg/s}$$

Eddington limit: similar if no spherical symmetry: luminosity requires minimum mass

bright quasars must have $M > 10^8 M_{sun}$

brightest X—ray binaries $M > 10 M_{sun}$

In practice Eddington limit can be broken by factors ~ few, at most.

Eddington implies limit on *growth rate of mass*: since

$$\dot{M} = \frac{L_{acc}}{\eta c^2} < \frac{4\pi G M m_p}{\eta c \sigma_T}$$

we must have

$$M \leq M_0 e^{t/\tau}$$

where

$$\boxed{\tau = \frac{\eta c \sigma_T}{4\pi G m_p} \approx 5\times10^7\, yr}$$

is the *Salpeter timescale*

Emitted spectrum of an accreting object

Accretion turns gravitational energy into electromagnetic radiation. Two extreme possibilities:

(a) all energy thermalized, radiation emerges as a blackbody. Characteristic temperature T_b , where

$$T_b = \left(\frac{L_{acc}}{4\pi R^2 \sigma} \right)^{1/4}$$

i.e. significant fraction of the accretor surface radiates the accretion luminosity. For a neutron star near the Eddington limit

$$L \approx 10^{38} erg / s, R = 10 km \Rightarrow T_b \approx 10^7 K$$

(b) Gravitational energy of each accreted electron-proton pair turned directly into heat at (shock) temperature T_s . Thus

$$3kT_s = \frac{GMm_p}{R}$$

For a neutron star $T_s \approx 5 \times 10^{11} K$

Hence typical photon energies must lie between

$$kT_b = 1keV \leq h\nu \leq kT_s \approx 50MeV$$

i.e. we expect accreting neutron stars to be X—ray or gamma—ray sources: similarly stellar—mass black holes

Good fit to gross properties of X—ray binaries

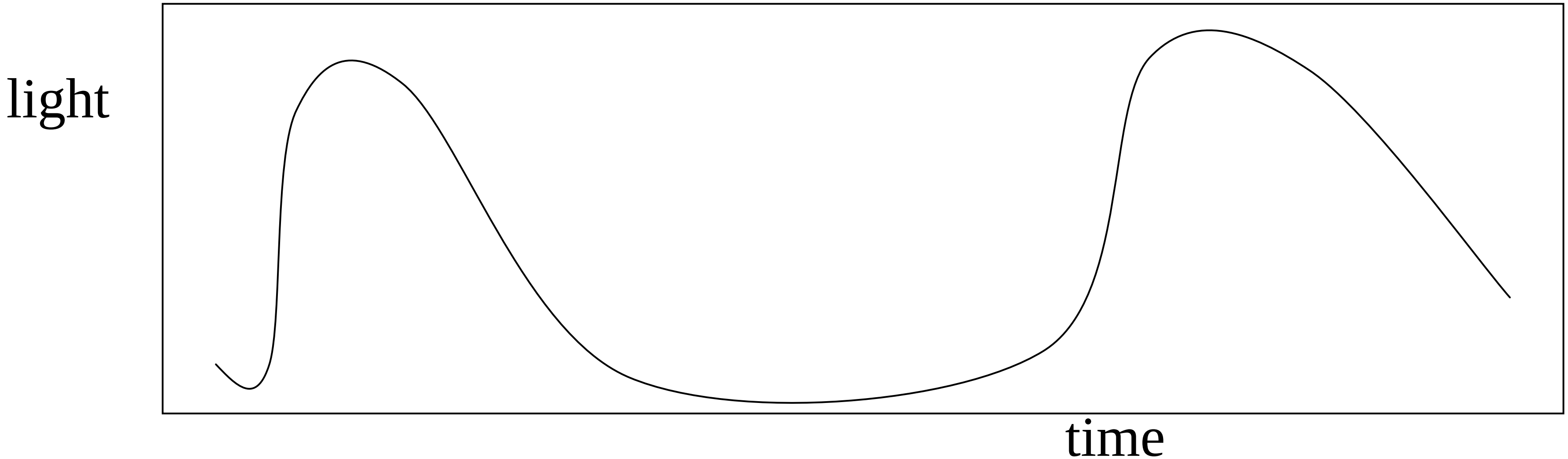
For a *white dwarf* accretor, mass = solar, radius = $10^9\, cm$
Find

$$T_b \sim 10^5 K, T_s = 10^8 K$$

so UV – X—ray sources.

Gross fit to gross properties of *cataclysmic variables* (CVs).

Many of these show *outbursts* of a few days at intervals of a few weeks – *dwarf novae.* See later….



For supermassive black holes we have

$$M \approx 10^8 M_{sun}, R = 2GM / c^2 \approx 3 \times 10^{13} cm$$

so

$$T_b \approx 10^7 (M_{sun} / M)^{1/4} K \approx 10^5 K$$

and T_s is unchanged. So we expect supermassive BH to be *ultraviolet*, *X—ray and possibly gamma—ray emitters.*

Good fit to gross properties of quasars

Modelling accreting sources

To model an accreting source we need to

(a) choose nature of compact object – black hole, neutron star, … to agree with observed radiation components

(b) choose minimum mass M of compact object to agree with luminosity via Eddington limit

Then we have two problems:

(1) we must arrange accretion rate $\dot{M}$ ***to provide observed luminosity, (the feeding problem) and***

(2) we must arrange to grow or otherwise create an accretor of the right mass M ***within the available time (the growth problem)***

Examine both problems in the following, for accreting binaries and active galactic nuclei (AGN)

for binaries
feeding: binary companion star
growth: accretor results from stellar evolution

for AGN
feeding: galaxy mergers?
growth: accretion on to `seed’ black hole?

Both problems better understood for binaries, so develop ideas and theory here first.

Modelling X—ray binaries

Normal galaxies like Milky Way contain several 100 X—ray point sources with luminosities up to $10^{39} erg / s$

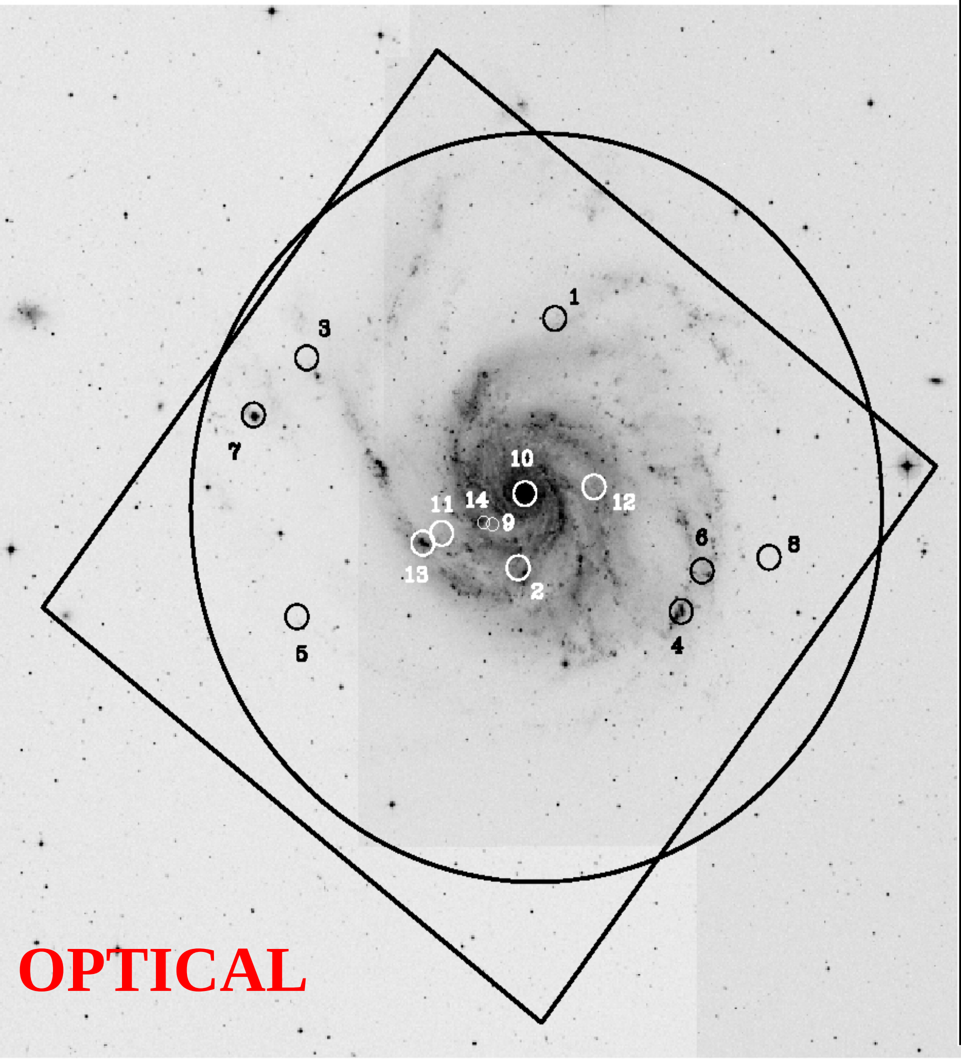
Typical spectral components ~ 1 keV and 10 – 100 keV

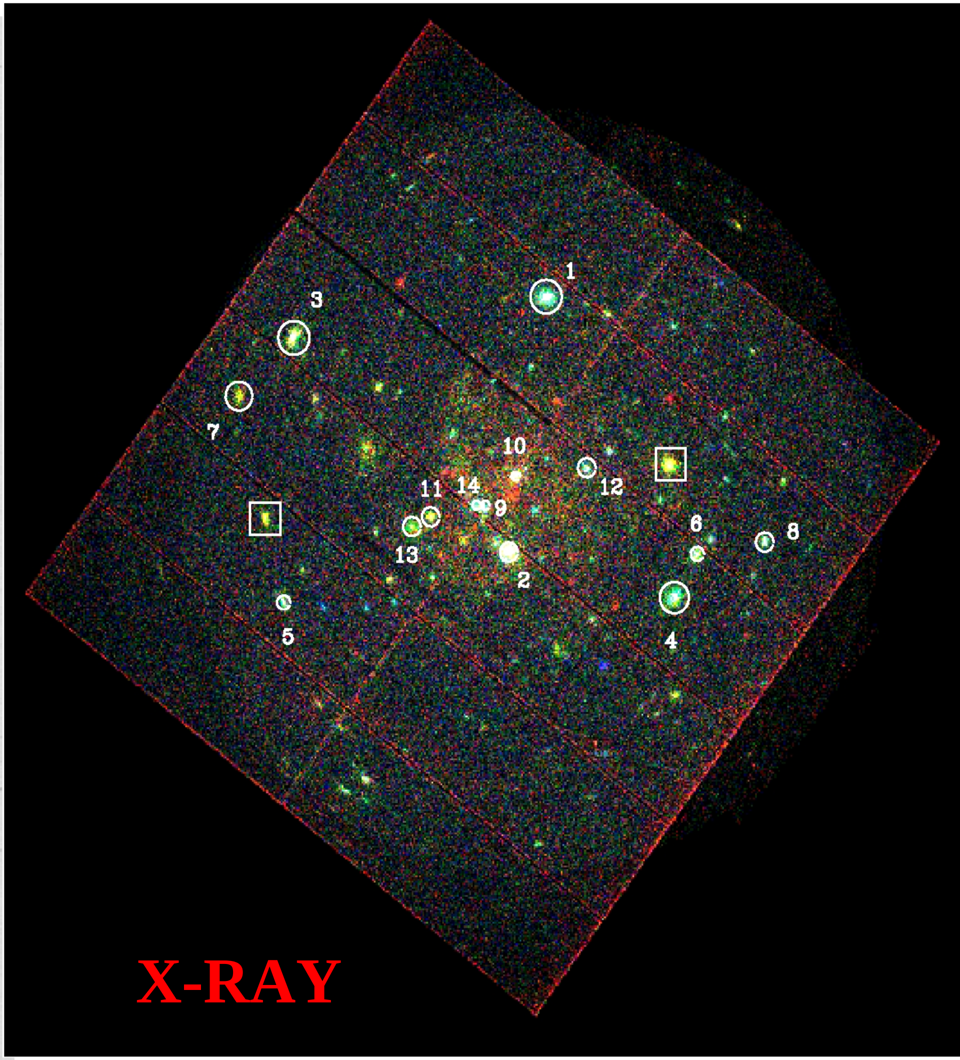
Previous arguments suggest *accreting neutron stars and black holes*
Brightest must be *black holes.*

Optical identifications: some systems are coincident with luminous hot stars: *high mass X—ray binaries* (HMXB).

But many do not have such companions: *low mass X—ray binaries* (LMXB).

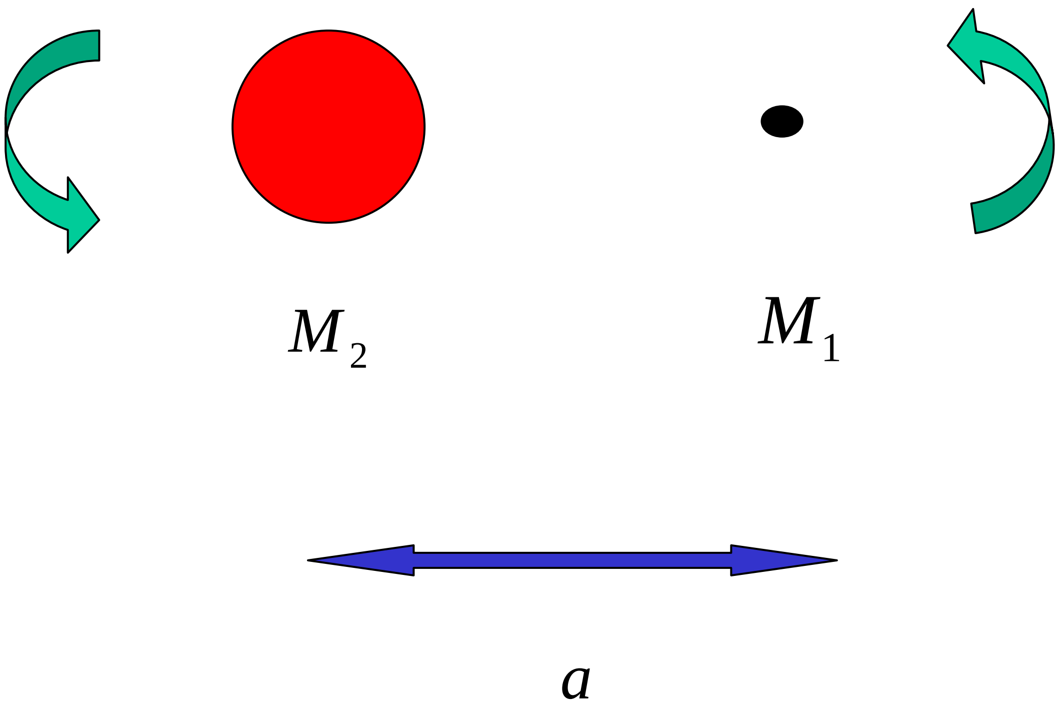
Accreting Black Holes in a Nearby Galaxy (M101)





Mass transfer in low mass X—ray binaries

Formation: starting from two newly—formed stars in a suitable binary orbit, a long chain of events can in a few rare cases lead to a BH or NS in a fairly close orbit with a low—mass main sequence star.



Two processes now compete to start mass transfer:

1. Binary loses angular momentum, to gravitational radiation or other processes. Binary separation a shrinks as

$$a \propto J^2$$

— full relation is

$$\boxed{J = M_1 M_2 \left(\frac{Ga}{M} \right)^{1/2}}$$

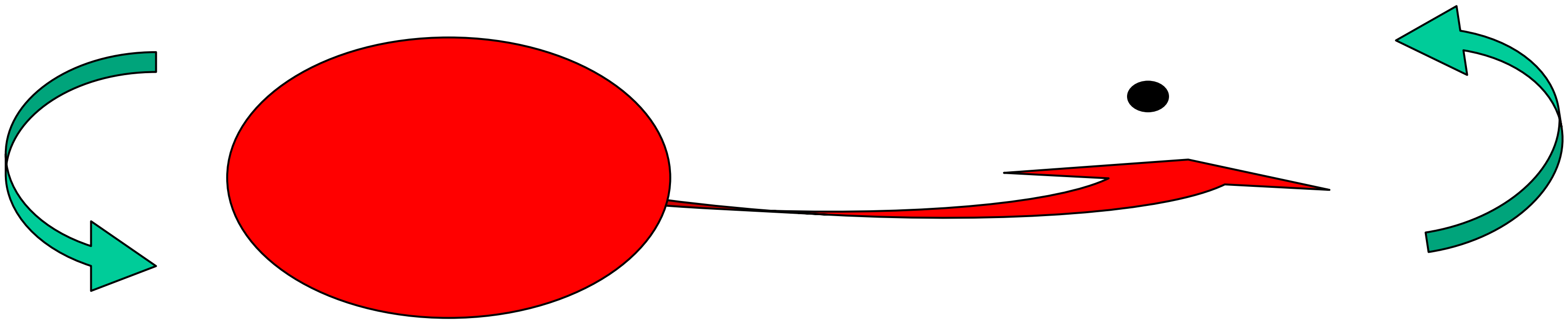
where
$$M = M_1 + M_2$$

2. Normal star evolves to become a giant, so radius increases to significant fraction of separation a

In both cases R_2/a *is continuously reduced.*

Combined gravitational—centrifugal (*Roche*) potential has two minima (`valleys') at the CM of each star, and a saddle point (`pass': *inner Lagrange point* L_1) between them.

Once R_2/a sufficiently reduced that the normal star reaches this point, mass flows towards the compact star and is controlled by its gravity – *mass transfer*



Mass transfer changes the binary separation itself:
orbital a.m. J and total binary mass M conserved, so logarithmic differentiation of J implies

$$0 = \frac{\dot{J}}{J} = \frac{\dot{M}_1}{M_1} + \frac{\dot{M}_2}{M_2} + \frac{\dot{a}}{2a}$$

and with

$$\dot{M}_1 = -\dot{M}_2 > 0$$

we have

$$\frac{\dot{a}}{2a} = (-\dot{M}_2)\frac{M_1 - M_2}{M_1 M_2}$$

Binary widens if accretor is (roughly) more massive than donor, shrinks if not.

In first case mass transfer proceeds on timescale of decrease of R_2 / a , i.e. a.m. loss or nuclear expansion:

$$\boxed{-\dot{M}_2 \approx -M_2 \frac{\dot{J}}{J}}$$

or

$$\boxed{-\dot{M}_2 = -\frac{M_2}{t_{nuc}}}$$

these processes can drive mass transfer rates up to ~ $10^{-6} M_{sun} / yr$ depending on binary parameters (masses, separation)

In this case –stable mass transfer — *star remains exactly same size as critical surface (Roche lobe):*

if lobe shrinks relative to star, excess mass transferred very rapidly (dynamical timescale)

if lobe expands wrt star, driving mechanism (a.m. loss or nuclear expansion) rapidly restores contact

Thus binary separation evolves to maintain this equality. *Orbital evolution follows stellar radius evolution.*

E.g. in some cases star expands on mass loss, even though a.m. loss drives evolution. Then orbit *expands*, and mass transferred to ensure that new *wider* binary has *lower angular momentum.*

If instead the donor is (roughly) more massive than accretor, binary shrinkage → mass transfer increases exponentially on dynamical timescale ~ few orbital periods.

Likely result of this *dynamical instability* is a binary *merger:* timescale so short that unobserved.

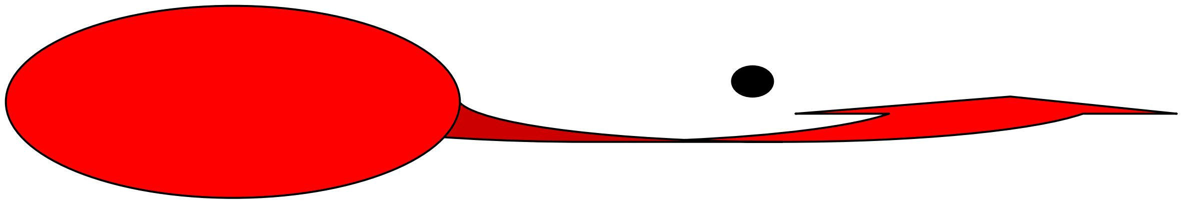
High mass X—ray binaries merge once donor fills Roche lobe: shortlived: accretion actually from *wind* of hot star. Many binaries pass through HMXB stage

Low mass X—ray binaries can have very long lifetimes, ~ a.m. or nuclear timescales

`Paradox': we observe *bright* LMXBs in *old* stellar populations! —see later……

Accretion disc formation

Transferred mass does not hit accretor in general, but must *orbit* it



— initial orbit is a rosette, but self—intersections → dissipation → energy loss, but no angular momentum loss

Kepler orbit with lowest energy for fixed a.m. is circle.

Thus orbit *circularizes* with radius such that specific a.m. j is the same as at L_1

Kepler's law for binary requires $GM/a^2 = a\omega^2$, or

$$\omega = (GM/a^3)^{1/2}$$

$$j = (GMa)^{1/2}$$

$\omega = 2\pi/P$, P = orbital period, j = specific a.m.,

Now L_1 roughly halfway across binary, and rotates with it, so specific a.m. *around accretor* of matter leaving it is

$$j \approx \left(\frac{a}{2}\right)^2 \omega \quad = \frac{(GMa)^{1/2}}{4}$$

So new circular orbit around accretor has radius r such that

$$j = (GM_1 r)^{1/2}$$, which gives

$$r_{circ} \approx \frac{M}{16\, M_1} a \approx \frac{a}{16} \approx \frac{R_2}{8}$$

In general *compact accretor radius is far smaller than* r_{circ}:

typically donor is at least as large as a main—sequence star, with

$$R_2 \sim R_{sun} = 7 \times 10^{10} cm \Rightarrow r_{circ} \sim 10^{10} cm$$

A neutron—star accretor has radius $10 km = 10^6 cm$

and a black hole has Schwarzschild radius

$$R = \frac{2GM}{c^2} = 3 \frac{M}{M_{sun}} km = 3 \times 10^5 \frac{M}{M_{sun}} cm$$

and last stable circular orbit is at most 3 times this

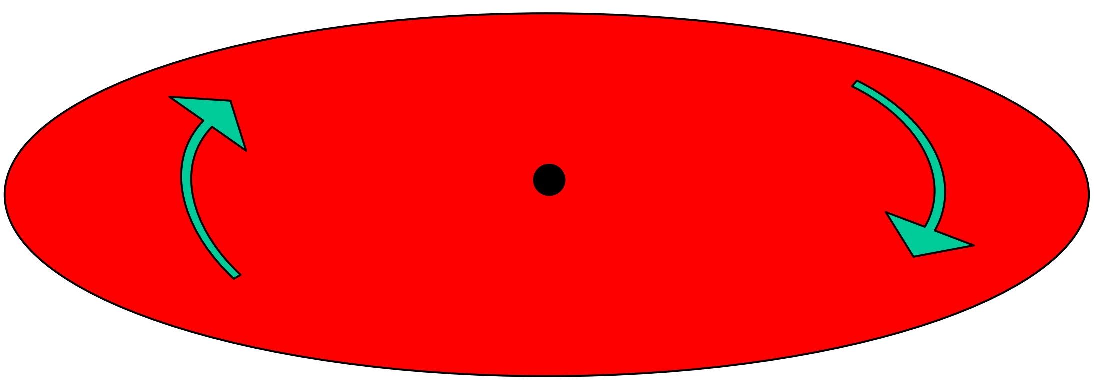
Thus in general matter orbits accretor. What happens?

Accretion requires angular momentum loss – see later: specific a.m. at accretor (last orbit) is smaller than initial by factor

$$(R / r_{circ})^{1/2} \geq 100$$

Energy loss through dissipation is quicker than angular momentum loss; matter spirals in through a sequence of circular Kepler orbits.

This is an *accretion disc*. At outer edge a.m. removed by tides from companion star



Accretion discs are *universal*:

matter usually has far too much a.m. to accrete directly – matter velocity not `aimed' precisely at the accretor!

in a galaxy, interstellar gas at radius R from central black hole has specific a.m. $\sim (GMR)^{1/2}$, where M is enclosed galaxy mass; *far* higher than can accrete to the hole, which is

$$\sim (GM_{bh}R_{bh})^{1/2} \sim (GM_{bh} \cdot GM_{bh} / c^2)^{1/2} = GM_{bh} / c$$

angular momentum increases in dynamical importance as matter gets close to accretor: matter may be captured gravitationally at large radius with `low' a.m. (e.g. from interstellar medium) but still has far too much a.m. to accrete

Capture rate is an upper limit to the accretion rate

- expect theory of accretion discs developed below to apply equally to supermassive black—hole accretors in AGN as well as close binaries

- *virtually all phenomena present in both cases*

Thin Accretion Discs

Assume disc is closely confined to the orbital plane with semithickness H, and surface density

$$\Sigma = \int_{-\infty}^{\infty} \rho dz \approx 2H \langle \rho \rangle$$

in cylindrical polars (R, ϕ, z). Assume also that

$$v_\phi = v_K = (GM/R)^{1/2}$$

These two assumptions are consistent: both require that *pressure forces are negligible*

Accretion requires angular momentum transport outwards. Mechanism is usually called `*viscosity*', but usual `molecular' process is too weak. Need torque *G(R)* between neighboring annuli

Discuss further later – but functional form must be

$$G(R) = 2\pi R \nu \Sigma R^2 \Omega'$$

with $$\Omega' = d\Omega / dR$$

reason: *G(R)* must vanish for rigid rotator $(\Omega' = 0)$

Coefficient $\nu \sim \lambda u$, where λ = typical lengthscale and u = typical velocity.

Net torque on disc ring between $R, R+\Delta R$ is

$$G(R+\Delta R)-G(R)=\frac{\partial G}{\partial R}\Delta R$$

Torque does work at rate

$$\Omega\frac{\partial G}{\partial R}\Delta R=\left[\frac{\partial}{\partial R}(G\Omega)-G\Omega'\right]\Delta R$$

but term

$$\frac{\partial}{\partial R}(G\Omega)\Delta\Omega$$

is transport of rotational energy – (a divergence, depending only on boundary conditions).

Remaining term represents dissipation: per unit area (two disc faces!) this is

$$D(R) = \frac{G\Omega'}{4\pi R} = \frac{1}{2}\nu\Sigma(R\Omega')^2$$

Note that this is positive, vanishing only for rigid rotation. For Keplerian rotation

$$\Omega = (GM / R^3)^{1/2}$$

and thus

$$D(R) = \frac{9}{8}\nu\Sigma\frac{GM}{R^3}$$

Assume now that disc matter has a small radial velocity v_R.

Then mass conservation requires (exercise!)

$$\boxed{R\frac{\partial \Sigma}{\partial t}+\frac{\partial}{\partial R}(R\Sigma v_R)=0}$$

Angular momentum conservation is similar, but we must take the `viscous' torque into account. The result is (exercise!)

$$R\frac{\partial}{\partial t}(\Sigma R^2\Omega)+\frac{\partial}{\partial R}(R\Sigma v_R R^2\Omega)=\frac{1}{2\pi}\frac{\partial G}{\partial R}$$

We can eliminate the radial velocity v_R, and using the Kepler assumption for Ω we get (exercise)

$$\boxed{\frac{\partial \Sigma}{\partial t} = \frac{3}{R}\frac{\partial}{\partial R}\left\{R^{1/2}\frac{\partial}{\partial R}\left[\nu\Sigma R^{1/2}\right]\right\}}$$

Diffusion equation for surface density: mass diffuses in, angular momentum out.

Diffusion timescale is *viscous timescale*

$$\boxed{t_{visc} \sim R^2 / \nu}$$

Steady thin discs

Setting $\partial / \partial t = 0$ we integrate the mass conservation equation as

$$R\Sigma v_R = const$$

Clearly constant related to (steady) accretion rate through disc as

$$\dot{M} = 2\pi R\Sigma(-v_R)$$

Angular momentum equation gives

$$R\Sigma v_R R^2\Omega = \frac{G}{2\pi} + \frac{C}{2\pi}$$

where *G(R)* is the viscous torque and *C* a constant.
Equation for *G(R)* gives

$$\nu \Sigma \Omega' = \Sigma(-v_R)\Omega + \frac{C}{2\pi R^3}$$

Constant *C* related to rate at which a.m. flows into accretor.
If this rotates with angular velocity << Kepler, there
is a point close to the inner edge R_* of the disc where

$$\Omega' = 0 \qquad \text{or equivalently} \qquad G(R_*) = 0$$

(sometimes called `no—stress' boundary condition). Then

$$C = -\dot{M}(GMR_*)^{1/2}$$

Putting this in the equation for Ω' and using the Kepler form of angular velocity we get

$$\boxed{\nu\Sigma = \frac{\dot{M}}{3\pi}\left(1-\left(\frac{R_*}{R}\right)^{1/2}\right)}$$

Using the form of D(R) we find the surface dissipation rate

$$\boxed{D(R) = \frac{3G\dot{M}M}{8\pi R^3}\left(1-\left(\frac{R_*}{R}\right)^{1/2}\right)}$$

Now if disc optically thick and radiates roughly as a blackbody,

$$D(R) = \sigma T_b^{\ 4}$$

so *effective temperature* T_b given by

$$\boxed{T_b^{\ 4} = \frac{3GM\dot{M}}{8\pi\sigma R^3}\left(1 - \left(\frac{R_*}{R}\right)^{1/2}\right)}$$

Note that T_b is *independent of viscosity*!

T_b is effectively observable, particularly in eclipsing binaries: this confirms simple theory.

Condition for a thin disc (H<<R)

Disc is almost hydrostatic in z-direction, so

$$\frac{1}{\rho}\frac{\partial P}{\partial z} = \frac{\partial}{\partial z}\left(\frac{GM}{(R^2 + z^2)^{1/2}}\right)$$

But if the disc is thin, $z<<R$, so this is

$$\frac{1}{\rho}\frac{\partial P}{\partial z} = -\left(\frac{GMz}{R^3}\right)$$

With $\partial P / \partial z \sim P / H, z \sim H$

and $P \sim \rho c_s^2$, where c_s is the sound speed, we find

$$H \sim c_s \left(\frac{R}{GM} \right)^{1/2} R \sim \frac{c_s}{v_K} R$$

Hence *for a thin disc we require that the local Kepler velocity should be highly supersonic*

Since $c_s \propto T^{1/2}$ *this requires that the disc can* ***cool.***

If this holds we can also show that *the azimuthal velocity is close to Kepler*

Thus for discs,

$$\boxed{\textbf{thin} \quad \Leftrightarrow \quad \textbf{Keplerian} \quad \Leftrightarrow \quad \textbf{efficiently cooled}}$$

*Either **all three** of these properties hold, or **none** do!*

Viscosity

Early parametrization: $\nu \sim \lambda u$ with typical length and velocity scales λ, u . Now argue that

$$\lambda < H, u < c_s$$

First relation obvious, second because supersonic random motions would shock. Thus set

$$\nu = \alpha c_s H$$

and argue that $\alpha < 1$. *But* no reason to suppose $\alpha = const$

`Alpha—prescription' useful because disc structure only depends on low powers of α . But *no predictive power*

Physical angular momentum transport

A disc has

$$\frac{\partial}{\partial R}(R^2\Omega) > 0, \qquad \text{but} \qquad \frac{\partial \Omega}{\partial R} < 0$$

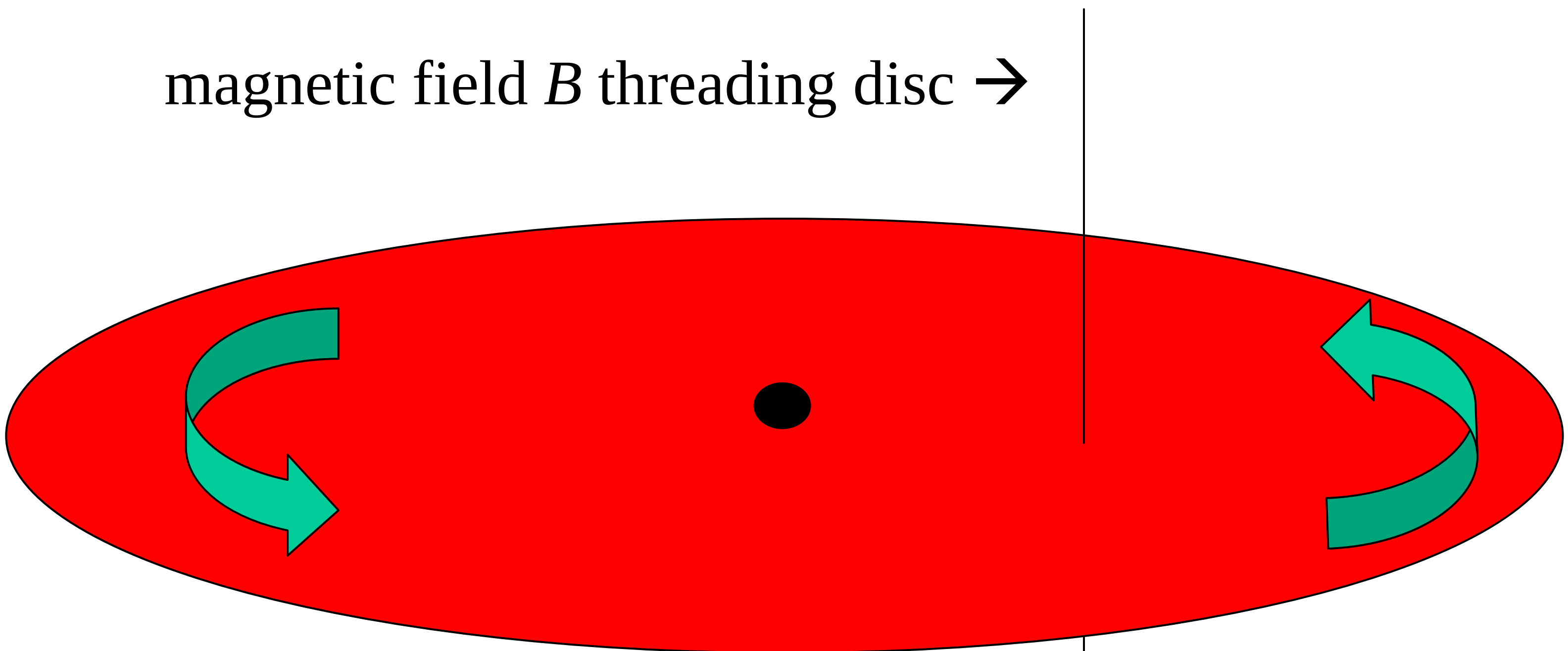
accretion requires a mechanism to transport a.m. outwards, but first relation → *stability* against axisymmetric perturbations (Rayleigh criterion).

Most potential mechanisms sensitive to a.m. gradient, so transport a.m. *inwards*!

need a mechanism sensitive to Ω or v_K

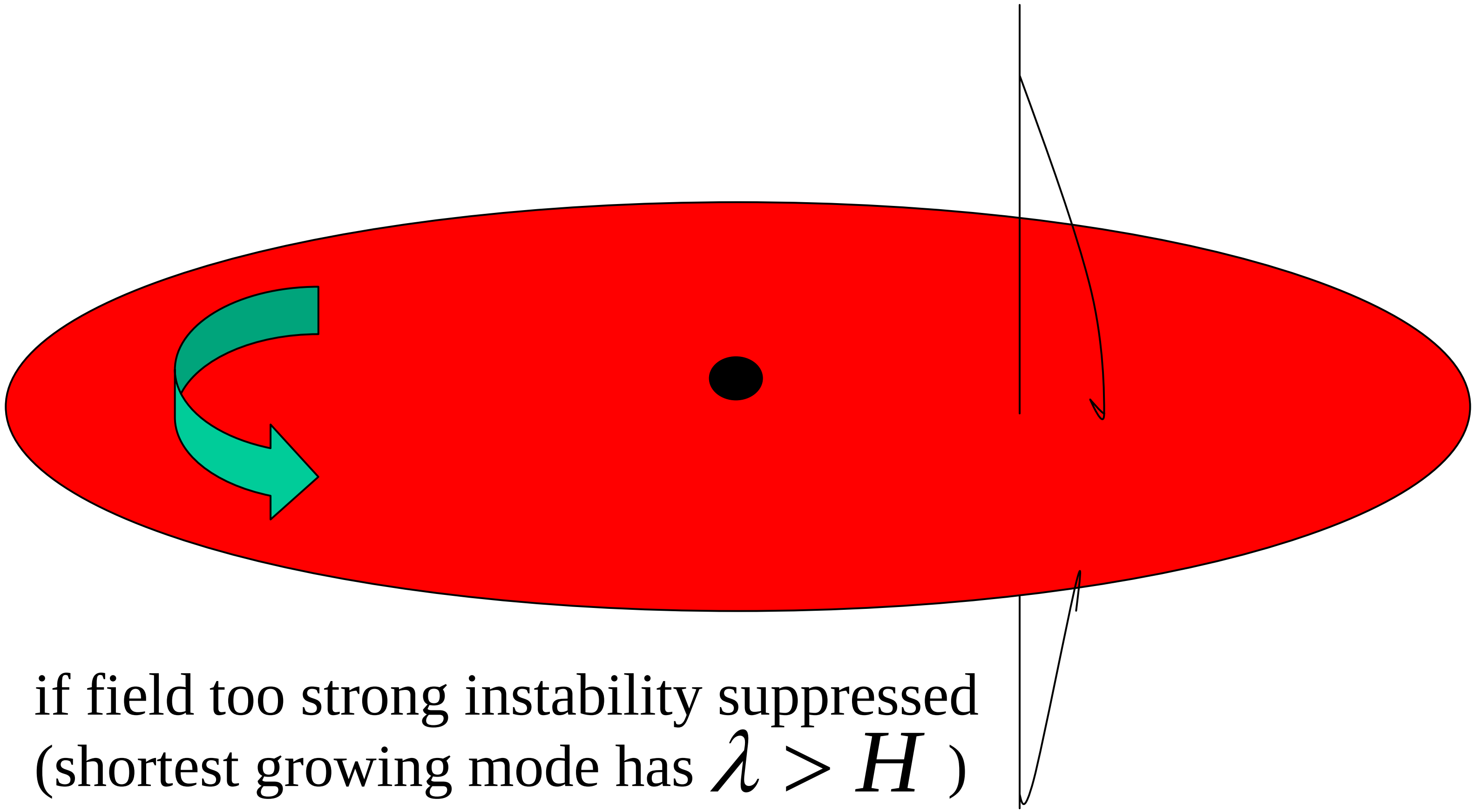
Balbus—Hawley (magnetorotational, MRI) instability

magnetic field *B* threading disc →



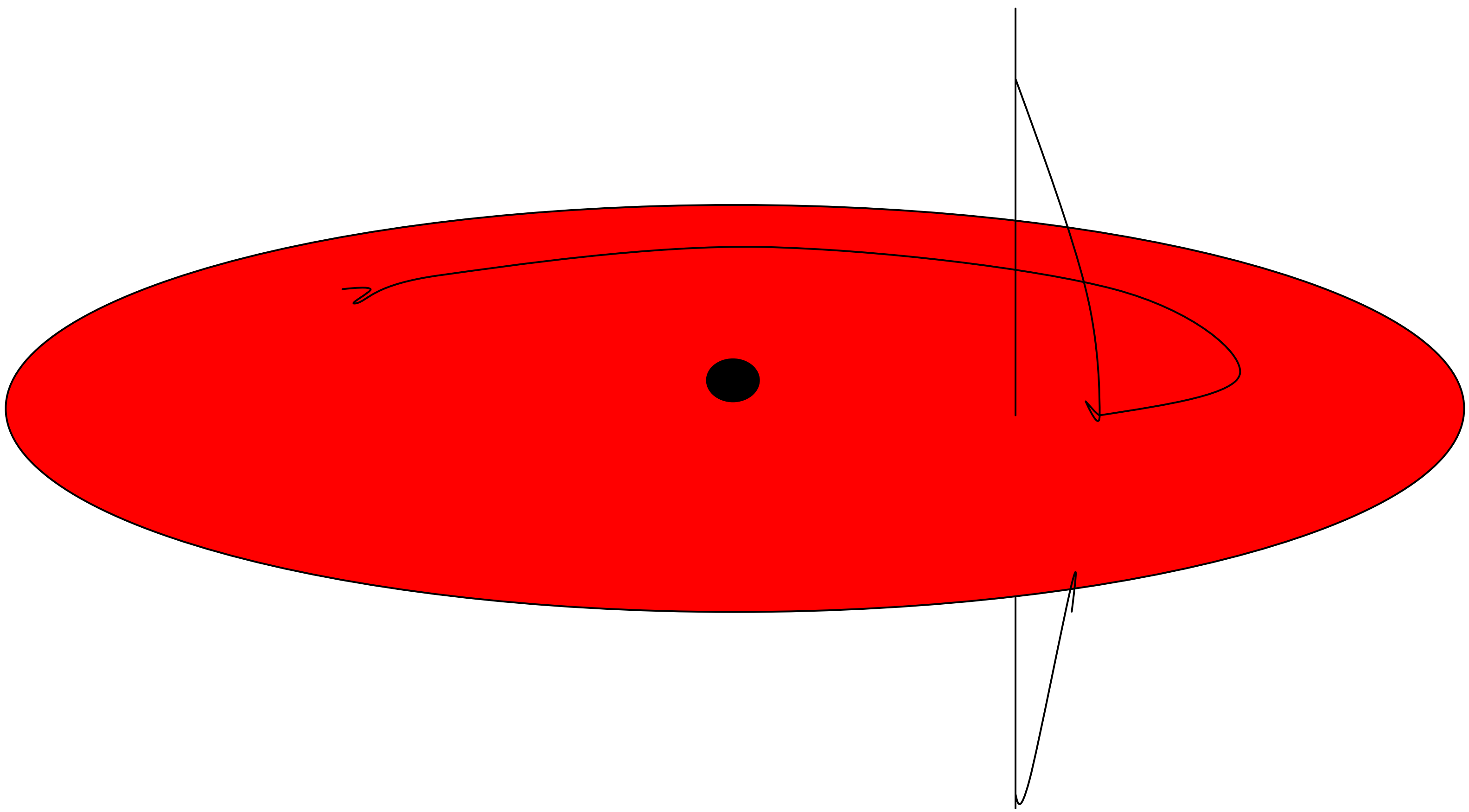
magnetic tension tries to *straighten* line
imbalance between gravity and rotation *bends* line

Vertical fieldline perturbed outwards, rotates faster than surroundings, so centrifugal force > gravity → *kink increases.* Line connects fast-moving (inner) matter with slower (outer) matter, and speeds latter up: ***outward a.m. transport***



if field too strong instability suppressed (shortest growing mode has $\lambda > H$)

distorted fieldline stretched azimuthally by differential rotation, *strength grows*

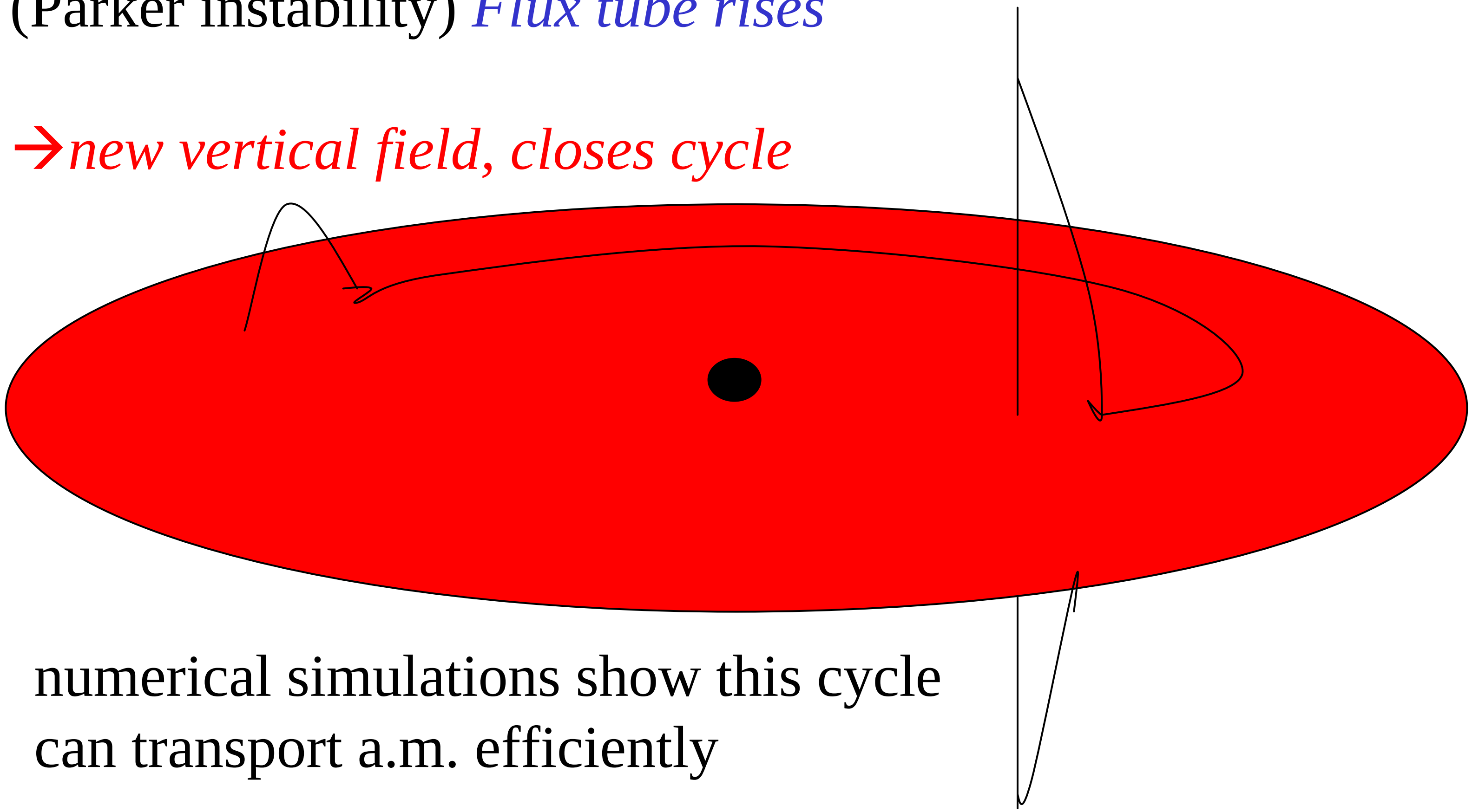


pressure balance between flux tube and surroundings requires

$$\frac{B^2}{8\pi} + P_{gas,in} = P_{gas,out}$$

so gas pressure and hence density lower inside tube → *buoyant* (Parker instability) *Flux tube rises*

→*new vertical field, closes cycle*



numerical simulations show this cycle can transport a.m. efficiently

Thin discs?

Thin disc conditions hold in many observed cases.

If not, disc is *thick, non—Keplerian, and does not cool efficiently*.
Pressure is important: disc ~ rapidly rotating `star'.

Progress in calculating structure slow: e.g. flow timescales far shorter at inner edge than further out.

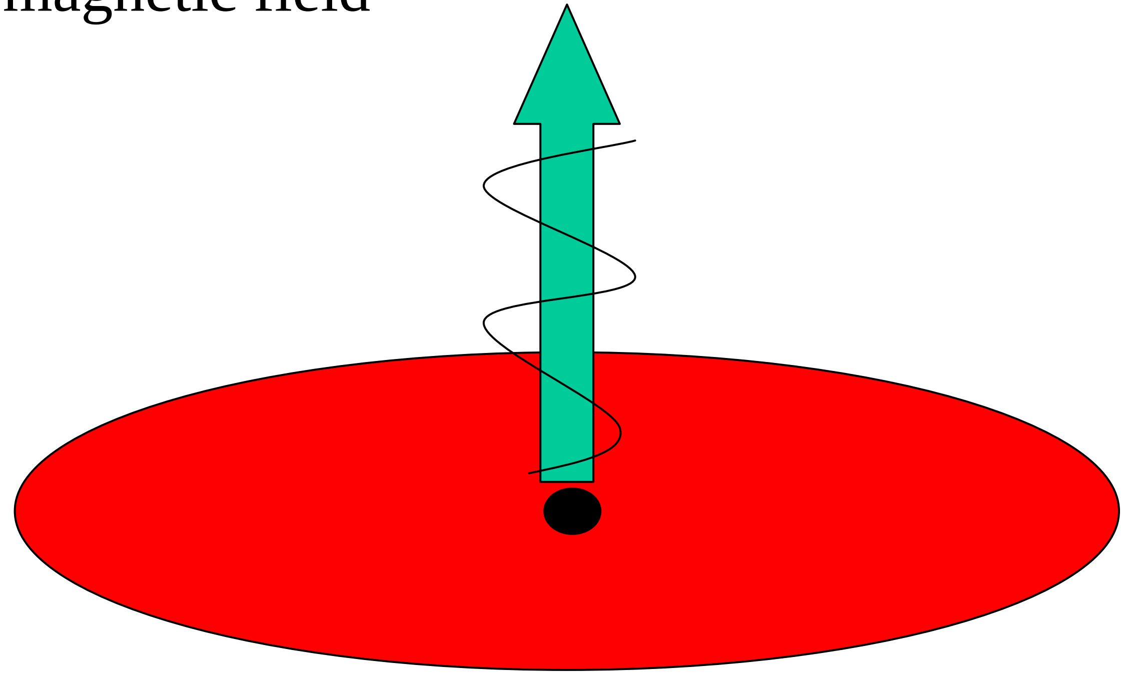
One possibility: matter flows inwards without radiating, and can accrete to a black hole `invisibly' (ADAF = advection dominated accretion flow). Most rotation laws → dynamical instability (PP).

Numerical calculations suggest indeed that most of matter flows out (ADIOS = adiabatic inflow—outflow solution)

Jets

One observed form of outflow: jets with ~ escape velocity from point of ejection, ~ *c* for black holes

Launching and collimation not understood – probably requires toroidal magnetic field

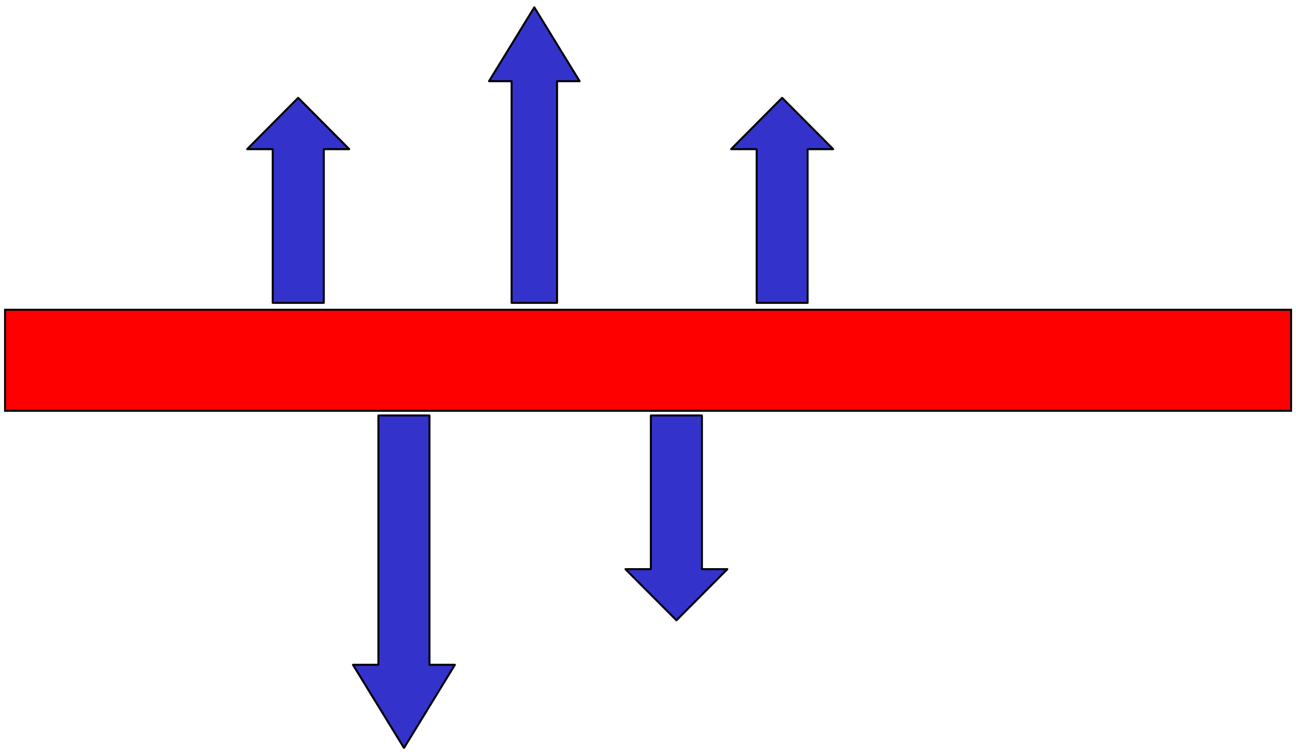


Disc may have *two* states:

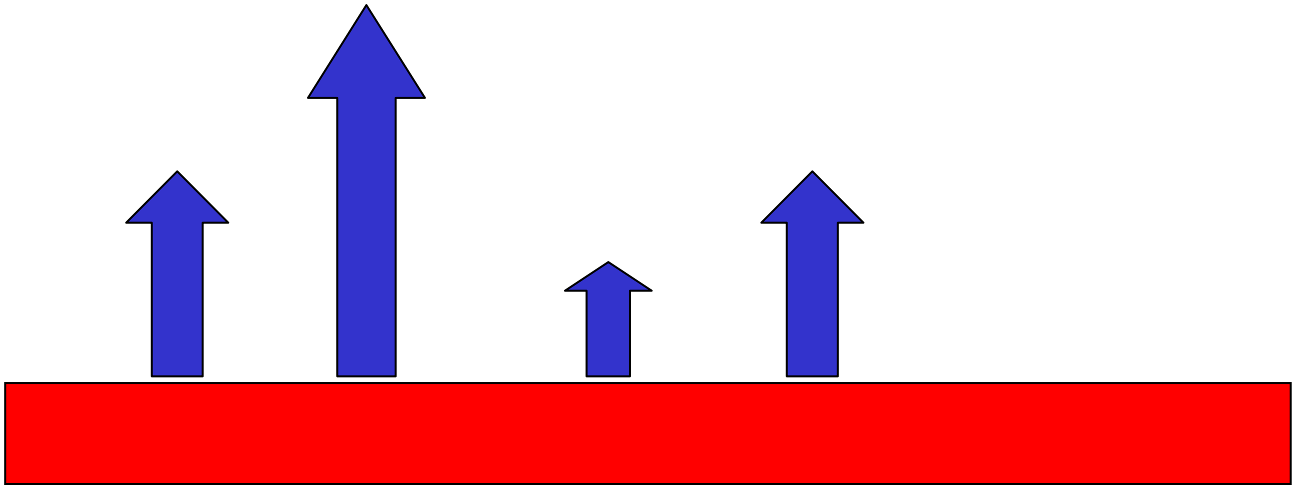
1. infall energy goes into radiation (standard)

2. infall energy goes into winding up internal disc field – thus

disc



generally vertical field directions uncorrelated in neighboring disc annuli (dynamo random); BUT



occasionally all fields line up → matter swept inwards, strengthens field → energy all goes into field → jet ???

(see King, Pringle, West, Livio, 2004)

jets seen (at times) in almost all accreting systems: AGN, LMXBs etc

Disc timescales

Have met dynamical timescale

$$t_{dyn} = R / v_K = (R^3 / GM)^{1/2}$$

and viscous timescale

$$t_{visc} = R^2 / \nu$$

We define also the thermal timescale

$$t_{th} = \Sigma c_s^2 / D(R) = \frac{R^3 c_s^2}{GM\nu} = \frac{c_s^2}{v_K^2} \frac{R^2}{\nu} = \frac{H^2}{R^2} t_{visc}$$

so

$$\boxed{t_{dyn} < t_{th} < t_{visc}}$$

Disc stability

Suppose a thin disc has steady—state surface density profile Σ_0

Investigate stability by setting $\Sigma = \Sigma_0 + \Delta\Sigma$
With $\mu = \nu\Sigma$ so that $\Delta\mu = (\partial\mu / \partial\Sigma)\Delta\Sigma$
diffusion equation gives (Exercise)

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

Thus diffusion (stability) if $\partial\mu / \partial\Sigma > 0$,
but
anti—diffusion (instability) if $\partial\mu / \partial\Sigma < 0$ — mass flows towards denser regions, disc breaks up into rings , on viscous timescale.

origin of instability:

$$\mu = \nu\Sigma \propto \dot{M}$$

so

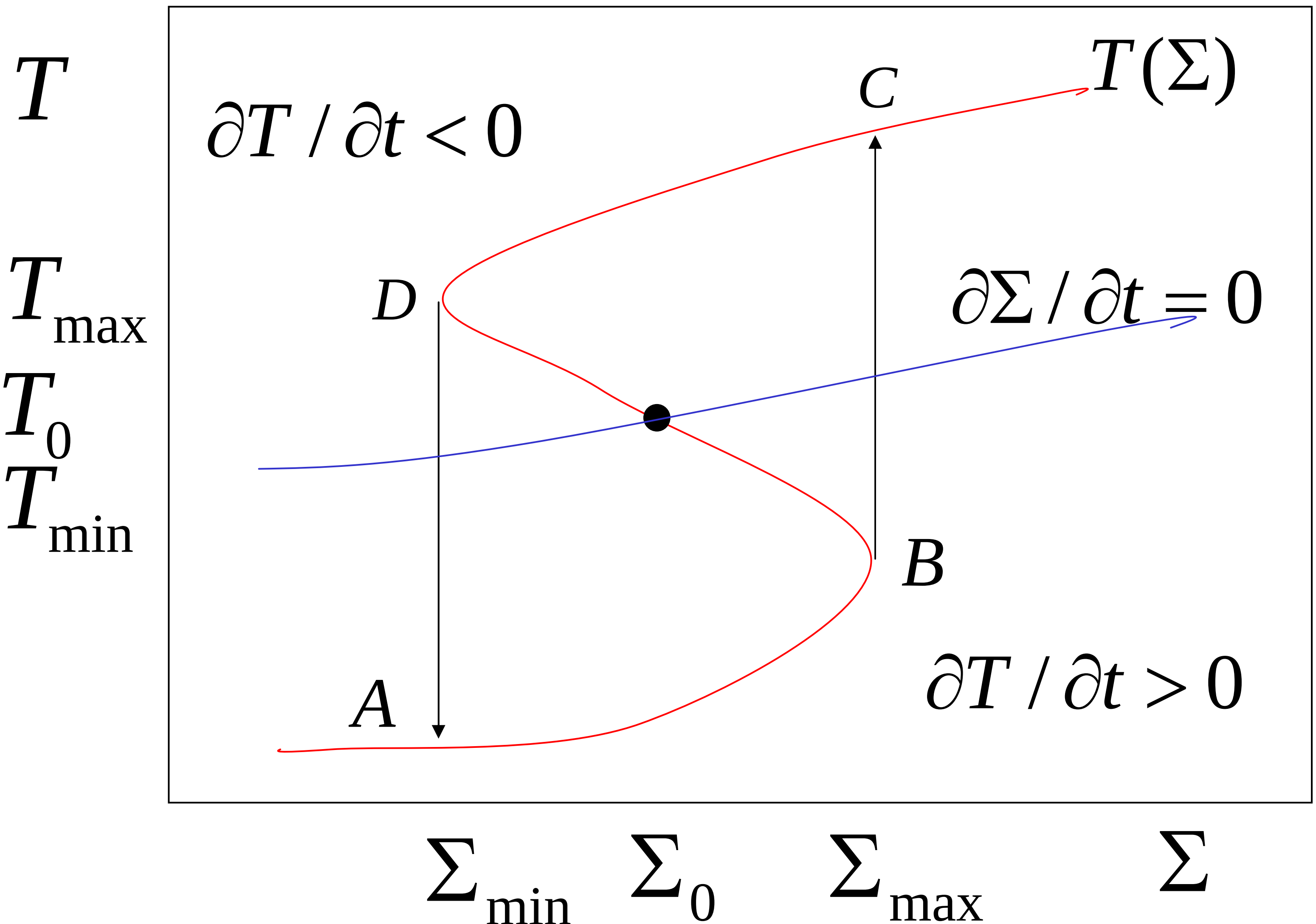
$$\partial\mu / \partial\Sigma < 0 \Rightarrow \partial \dot{M} / \partial\Sigma < 0$$

i.e. local accretion rate *increases* in response to a *decrease* in Σ (and vice versa), so local density drops (or rises).

To see condition for onset of instability, recall

$$\mu = \nu\Sigma \propto \dot{M} \propto T_b^4$$

and $T_b \propto$ internal temperature T. Thus stability/instability decided by sign of $\partial T / \partial \Sigma$ along the equilibrium curve $T(\Sigma)$ i.e. $\partial T / \partial t = 0$



Equilibrium

$$\partial T / \partial t = \partial \Sigma / \partial t = 0$$

here lies on unstable branch $\partial T / \partial \Sigma < 0$

System is forced to hunt around limit cycle ABCD, unable to reach equilibrium.

evolution A→B on long viscous timescale
evolution B→C on very short thermal timescale
evolution C→D on moderate viscous timescale
evolution C→A on very short thermal timescale

Thus get *long low states* alternating with *shorter high states*, with *rapid upwards and downward transitions* between them – **dwarf nova light curves.**

origin of wiggles in equilibrium $T(\Sigma)$ *curve* is hydrogen ionization threshold at $T \sim 10^4 K$

If all of disc is hotter than this, disc remains stably in the high state – no outbursts.

Thus *dwarf novae must have low mass transfer rates*:

$$T_b^{\,4} = \frac{3GM\dot{M}}{8\pi\sigma R_{out}^{\;3}} < 10^{16} K^4$$

where R_{out} is outer disc radius: requires $\dot{M} \sim 10^{-10} M_{sun} / yr$

Dwarf novae are *white dwarf* accretors: is there a *neutron—star or black—hole analogue?*

soft X—ray transients (*SXTs*) have outbursts, but much *brighter, longer and rarer*

why? observation →

discs are brighter than dwarf novae for same accretion rate

→ *X—ray irradiation* by central source: disc is *concave or warped (later)*

thus $T_b = T_{irr} \propto R^{-1/2}$ not $T_{visc} \propto R^{-3/4}$ *so dominates at large R* (where most disc mass is)

ionization/stability properties controlled by CENTRAL $\dot{M}$

Thus an SXT outburst cannot be ended by a cooling wave, as in DN.

outburst ends only when central accretion rate drops below a critical value such that $T_{irr}(R_{out}) = T_{ion} \approx 6500K$

→ mass of central disc drops significantly → *long*!

K & Ritter (1998): in outburst disc is roughly steady—state, with

$$\Sigma \approx \frac{\dot{M}_c}{3\pi\nu}$$

$\dot{M}_c$ the central accretion rate. Mass of hot disc is

$$M_h = 2\pi \int_0^{R_h} \Sigma R dR \approx \dot{M}_c \frac{R_h^2}{3\nu}$$

Now hot zone mass can change only through central accretion, so

$$\dot{M}_h = -\dot{M}_c$$

thus

$$-\dot{M}_h = \frac{3\nu}{R_h^2} M_h$$

i.e.

$$M_h = M_0 e^{-3\nu t / R_h^2}$$

so central accretion rate, X—rays, drop exponentially for small discs

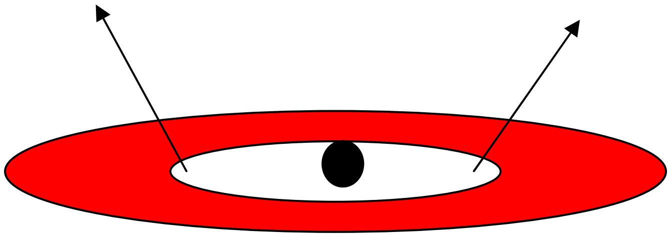
observation indeed shows that short—period (small disc) SXTs are exponential eventually central accretion rate low enough that disc edge is no longer ionized → *linear decay* rather than exponential

large discs (long period systems) always in this regime: linear decays sometime seen

however light curves complex since large mass at edge of disc not involved in outburst

main problem: why don't outbursts recur before disc mass reaches large values observed? $M_h \geq 10^{24} - 10^{28} g$

central mass loss?



condition for SXT outbursts – low disc edge temperature

→ low mass transfer rate/large disc

observable consequences:

ALL long—period LMXBs are transient

outbursts can last years and be separated by many centuries
e.g. GRS1915+105: outburst > 15 years

→ *outbursting systems may look persistent*

→ *quiescent systems not detectable*

`paradox' of bright X—ray sources in old stellar systems

elliptical galaxies have sources with $L_X \geq 10^{39} erg/s$:

this requires accretion rates $\dot{M} \geq 10^{-7} M_{sun}/yr$,

but galaxy has no stars younger than $\sim 10^{10} yr$,

so no extended stars with masses $\geq 1 M_{sun}$

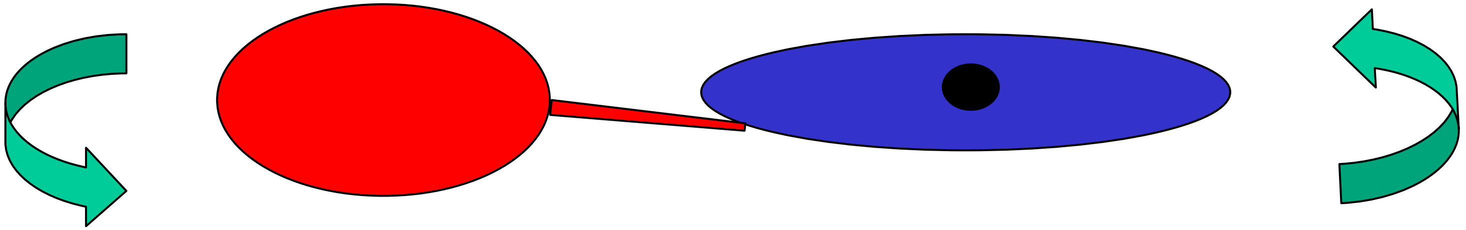
this would imply X—ray lifetimes $< 10^{7} yr$ i.e. we observe at a very special epoch!

resolution: sources *transient*, `duty cycle' $d \leq 10^{-2,-3}$

missing systems:

long—period LMXBs with neutron—star accretors

$$P \sim 20 days$$



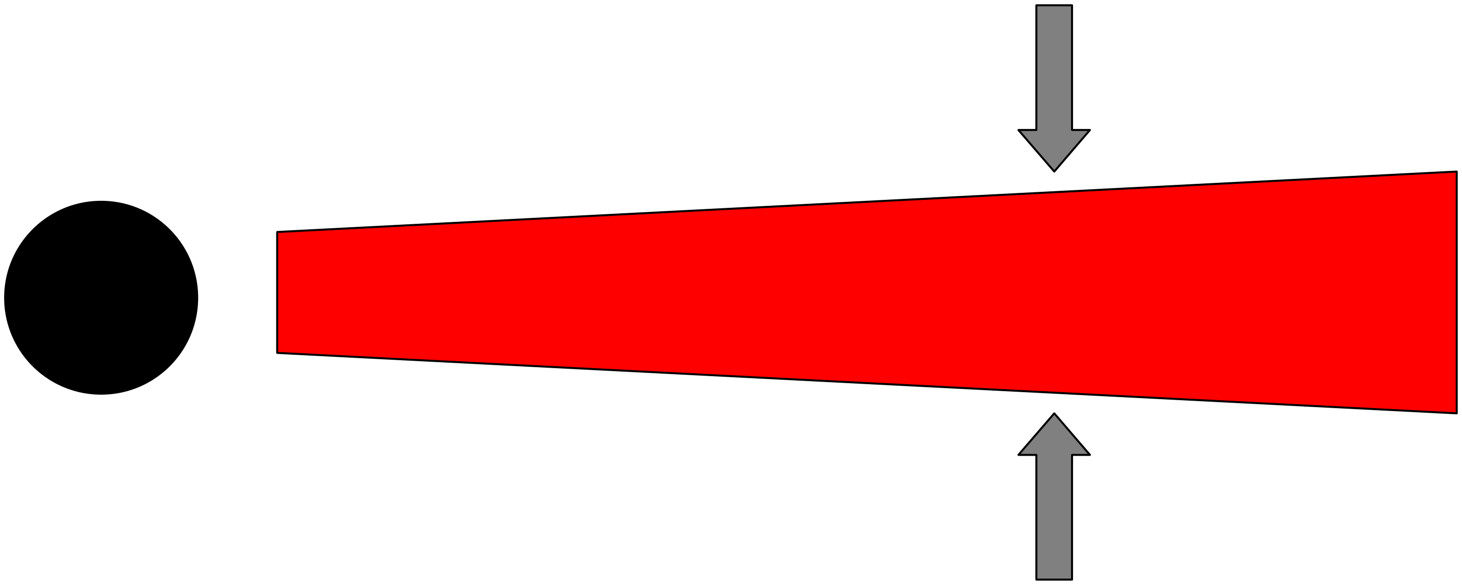
evolve into millisecond pulsars with white dwarf companions

$$P \sim 100 days$$

far too few of former cf latter → *transients* with $d \leq 10^{-2,-3}$

H—ionization (`thermal—viscous') instability so generic that probably occurs in supermassive black hole accretion too

main difference: *size of AGN disc* set by *self—gravity*



vertical component of gravity from central mass is $\sim GMH / R^3$

cf that from self—gravity of disc $\sim G\rho H^3 / H^2 \sim G\rho H$

Thus self—gravity takes over where $\rho \sim M / R^3$, or

$$\boxed{M_{disc} \sim R^2 H\rho \sim \frac{H}{R} M}$$

disc breaks up into stars outside this

almost all discs around SMBH have ionization zones, i.e.

their accretion discs should have outbursts

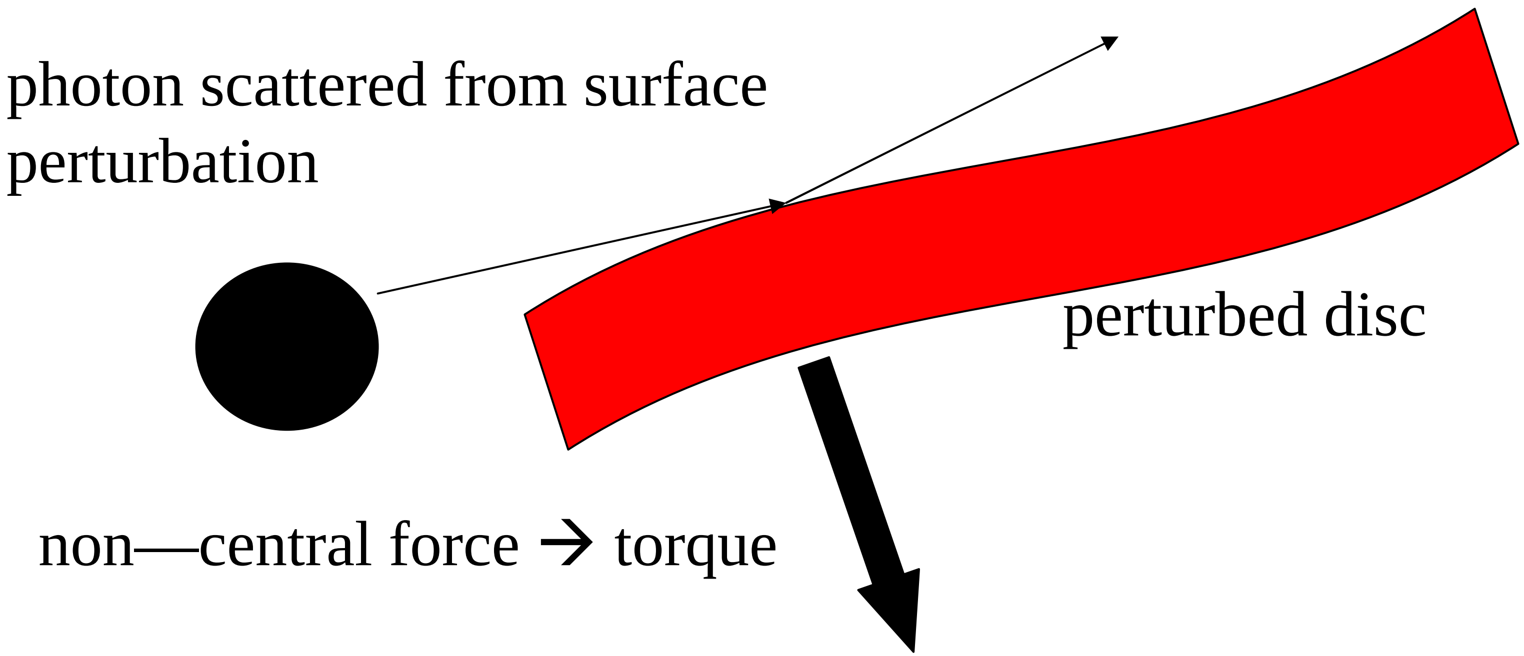
AGN = outburst state?

normal galaxies = quiescent state?

disc warping

gravitational potential of accretor ~ spherically symmetric: nothing special about orbital plane – other planes possible, i.e. disc can *warp*

radiation warping:



Pringle (1996) shows that resulting radiation torque makes perturbation grow at radii

$$R > R_{warp} \approx 8\pi^2 \psi^2 R_{in}^2 / R_{Schw}$$

where ψ is vertical/horizontal viscosity ratio, and R_{in}, R_{Schw} are inner disc and Schwarzschild radii. Once perturbation grows (on viscous time) it propagates inwards

Thus self—warping likely if accretor is a black hole or neutron star, i.e. LMXBs and AGN

Jets probably perpendicular to inner disc, so

jets can point anywhere

accretion to central object

central object gains a.m. and *spins up* at rate

$$\sim \dot{M}(GMR_{in})^{1/2}$$

reaches maximum spin rate (*a* ~ *M* for black hole) after accreting ~ *M* if starts from low spin. `Centrifugal' processes limit spin. For BH, photon emission limits *a/M <1*

thus *LMXBs and HMXBs do not significantly change BH spin*

magnetic neutron stars, WD *do* spin up, since accreted specific a.m.

$\sim (GMR_{mag})^{1/2}$ is much larger: needs only $\sim 0.1 M_{sun}$

in AGN, BH gains mass significantly

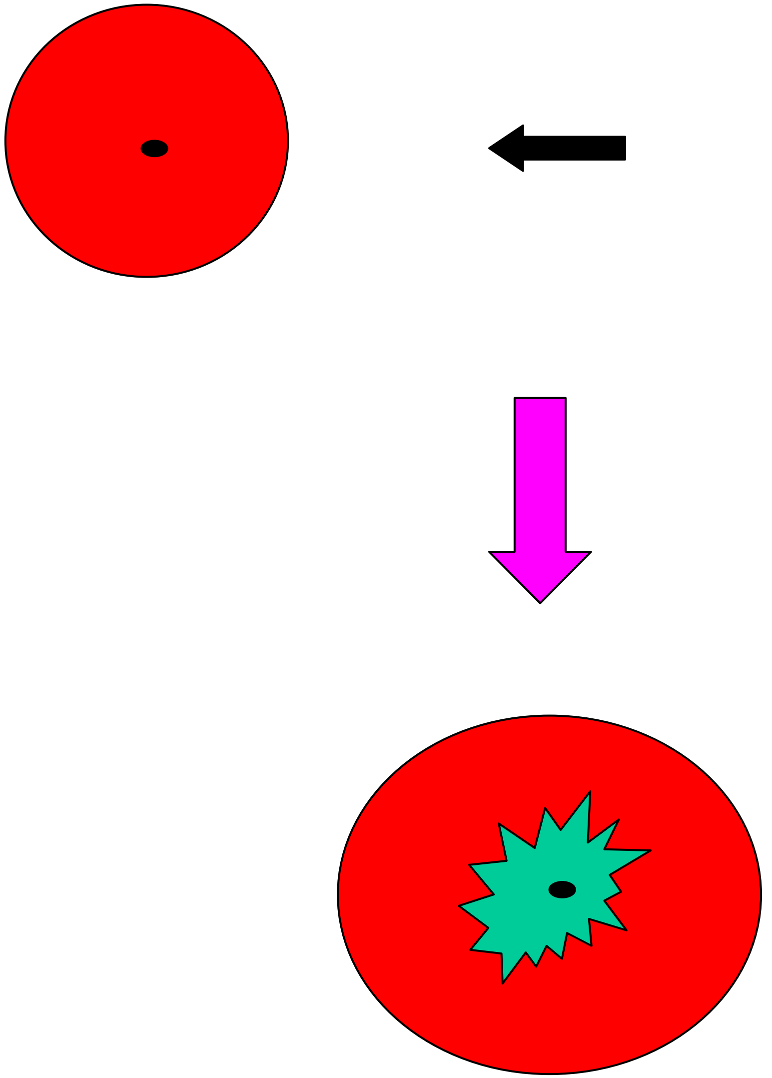
active galactic nuclei

supermassive BH $(10^6 - 10^9 M_{sun})$ in the centre of every galaxy

how did this huge mass grow?

cosmological picture:

big galaxy swallows small one



merger

galaxy mergers

two things happen:

1. *black holes coalesce*: motion of each is slowed by inertia of gravitational `wake' – *dynamical friction*. Sink to bottom of potential and orbit each other. GR emission → *coalescence*

2. *accretion*: disturbed potential → gas near nuclei destabilized, a.m. loss → accretion: *merged black hole grows*: radiation → *AGN*

black hole coalescence

black hole *event horizon area*

$$A = \frac{8\pi G^2}{c^4}[M^2 + (M^4 - c^2 J^2 / G^2)^{1/2}]$$

or

$$\boxed{A \propto M^2 [1 + (1 - a_*^2)^{1/2}]}$$

where $J =$ a.m., $a_* = cJ/GM^2$, *can never decrease* thus can give up angular momentum and still increase area, i.e. *release rotational energy* – e.g. as gravitational radiation

then *mass M decreases!* – minimum is $M/\sqrt{2}$ (irreducible mass) – start from $a_* = 1$ and spin down to $a_* = 0$ keeping *A* fixed

coalescence can be both *prograde* and *retrograde* wrt spin of merged hole, i.e. orbital opposite to spin a.m.

Hughes & Blandford (2003): long—term effect of coalescences is *spindown* since last stable circular orbit has larger a.m. in retrograde case.

black hole accretion

Soltan (1982): total rest—mass energy of all SMBH

consistent with radiation energy of Universe

if masses grew by luminous accretion (efficiency ~10 %)

thus ADAFs etc unimportant in growing most massive black holes *merger picture of AGN: consequences for accretion*

- mergers do not know about *black hole mass* M, so accretion may be super—Eddington

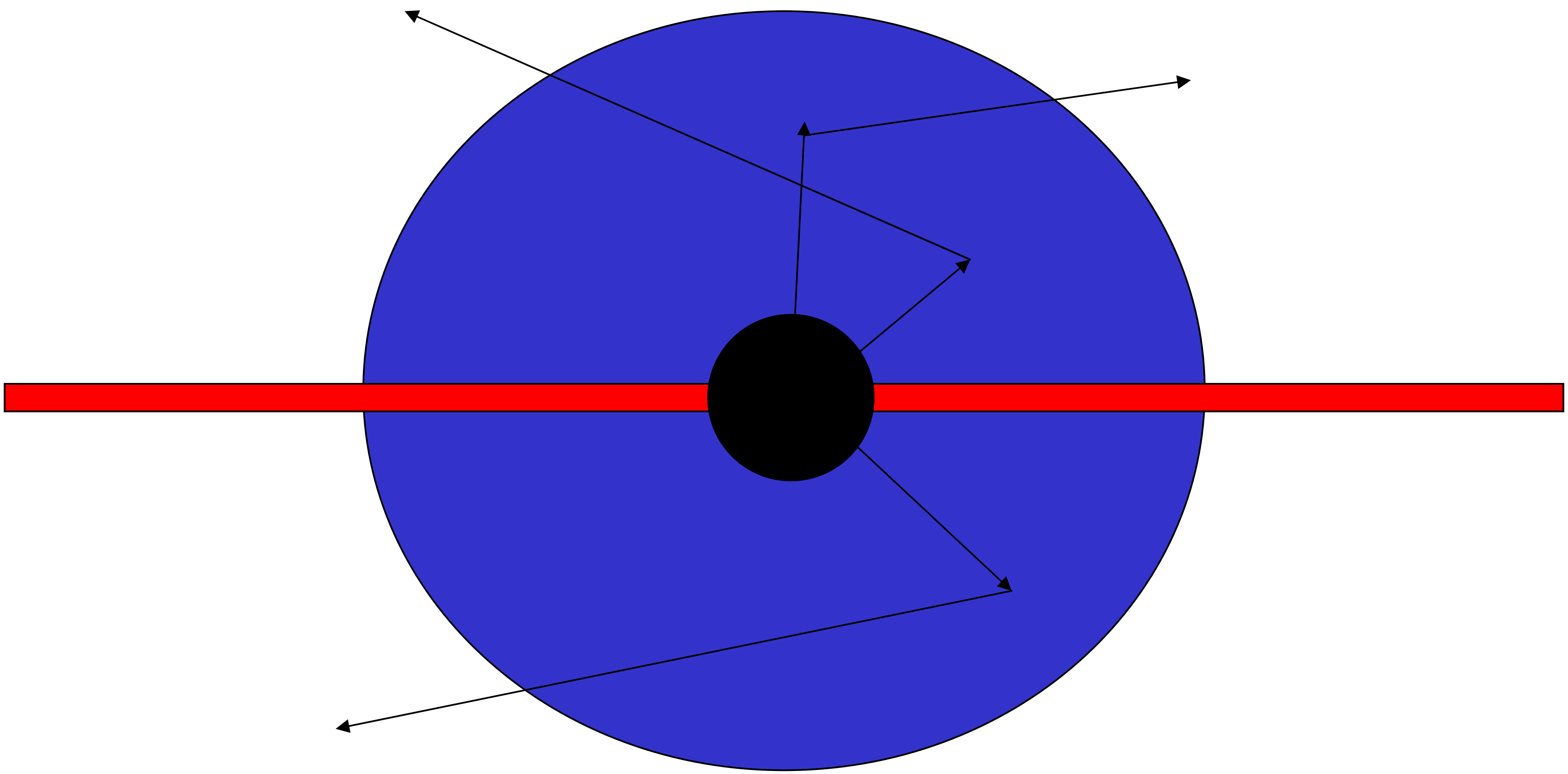
- mergers do not know about *hole spin* a, so accretion may be retrograde

• ***super—Eddington accretion***:

must have been common as most SMBH grew (z ~2), so

outflows

outflow is optically thick to scattering: radiation field L » L_{Edd} transfers » *all* its momentum to it



- *response to super—Eddington accretion*: expel excess accretion as an outflow with ***thrust*** given purely by L_{Edd} , i.e.

$$\boxed{\dot{M}_{out} v \approx \frac{L_{Edd}}{c}}$$

- *outflows with Eddington thrust must have been common as SMBH grew*

- NB mechanical energy flux $\frac{1}{2}\dot{M}_{out} v^2 \approx \frac{L_{Edd} v}{c}$ requires knowledge of v or $\dot{M}_{out}$

• *effect on host galaxy large*: must absorb most of the outflow momentum and energy – galaxies not `optically thin’ to matter – unlike radiation

• e.g. could have accreted at » 1M– yr^{-1} for » 5£10^{7} yr

• *mechanical energy* deposited in this time » 10^{60} erg

• cf *binding energy* » 10^{59} erg of galactic bulge with *M* » *10^{11} M*– and velocity dispersion σ » 300 km s^{-1}

• examine effect of super—Eddington accretion on growing SMBH (K 2003)

• model protogalaxy as an isothermal sphere of dark matter: gas density is

$$\rho(R) = \frac{f_g \sigma^2}{2\pi G r^2}$$

with $f_g = \Omega_{\mathrm{baryon}}/\Omega_{\mathrm{matter}}$ ' 0.16

• so gas mass inside radius R is

$$M(R) = 4\pi \int_0^R \rho r^2 dr = \frac{2 f_g \sigma^2 R}{G}$$

- dynamics depend on whether gas cools (`momentum—driven’) or not (`energy—driven’)

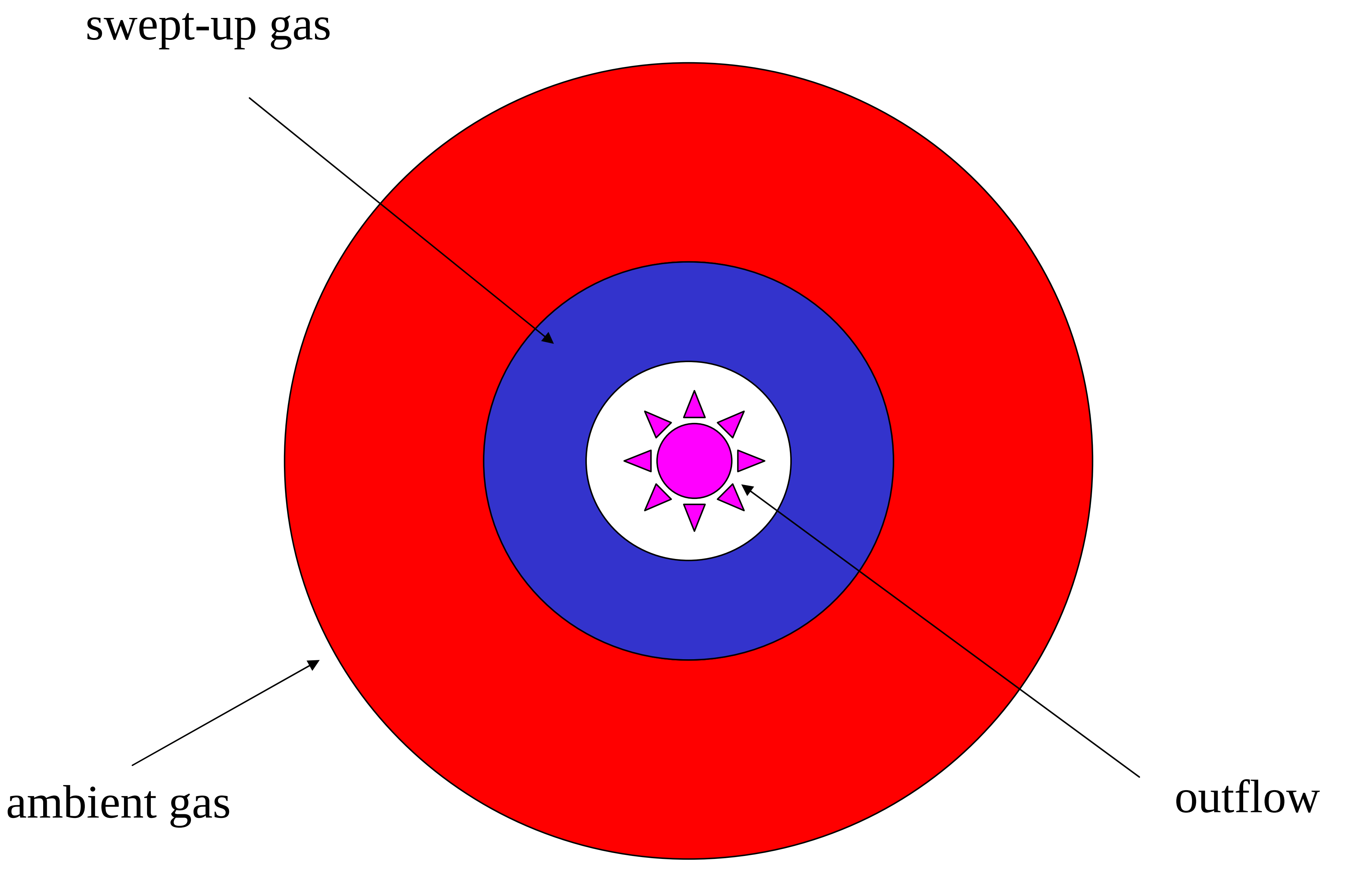
- Compton cooling is efficient out to radius R_c such that

$$M(R_c) \approx 2 \times 10^{11} \sigma^3_{200} M_8^{1/2} M_\odot$$

where $\sigma_{200} = \sigma/200\ km\ s^{-1}$, $M_8 = M/10^8 M_\odot$

- flow is momentum—driven (i.e. gas pressure is unimportant) out to $R = R_c$

for $R > R_c$ flow speeds up because of pressure driving


swept-up gas
ambient gas
outflow

ram pressure of outflow drives expansion of swept-up shell:

$$\frac{d}{dt}[M(R)\dot{R}] = 4\pi R^2 \rho v^2 = \dot{M}_{out} v - \frac{GM^2(R)}{R^2}$$

$$= \frac{L_{Edd}}{c} - 4 f_g \frac{\sigma^4}{G} = const$$

(using $M(R) = 2f_g\sigma^2 R/G$ etc)

thus

$$R^2 = \left[\frac{GL_{Edd}}{2 f_g \sigma^2 c} - 2\sigma^2\right] t^2 + 2R_0 v_0 t + R_0^2$$

for small L_{Edd} (i.e. small *M*), *R* reaches a maximum

$$R_{\max}^2 = \frac{R_0^2 v_0^2}{2\sigma^2 - GL_{Edd}/2 f_g \sigma^2 c} + R_0^2$$

in a dynamical time $\sim R_{\max}/\sigma$

R cannot grow beyond $R_{\max}$ until *M* grows: expelled matter is trapped inside bubble

M and *R* grow on Salpeter timescale $\sim 5\times10^7\ yr$

gas in shell recycled – star formation, chemical enrichment

- ***starbursts accompany black—hole growth***
- AGN accrete gas of *high metallicity*

ultimately shell too large to cool: drives off gas outside

- velocity large: *superwind*
- remaining gas makes bulge stars — *black—hole bulge mass relation*
- no fuel for BH after this, so *M* fixed: *M—sigma relation*

thus M grows until

$$\boxed{M = \frac{f_g \kappa}{\pi G^2} \sigma^4}$$

or

$$M = 2 \times 10^8 \sigma_{200}^4 M_\Theta$$

for a dispersion of 200 km/s

Note: predicted relation

$$\boxed{M = \frac{f_g \kappa}{\pi G^2} \sigma^4}$$

Note: predicted relation

$$\boxed{M = \frac{f_g \kappa}{\pi G^2} \sigma^4}$$

has no free parameter!

- M—sigma is very close to observed relation (Ferrarese & Merritt, 2000; Gebhardt et al., 2000; Tremaine et al, 2002)

- only mass inside cooling radius ends as bulge stars, giving

$$\boxed{M \sim 7\times10^{-4} M_8^{-1/4} M_{bul}}$$

- cooling radius is upper limit to galaxy size

$$\boxed{R_c < 80\sigma_{200} M_8^{1/2} kpc}$$

- above results in good agreement with observation

- argument via Soltan assumes standard accretion efficiency

- but ***mergers*** imply accretion flows initially *counter*aligned in half of all cases, i.e. low accretion efficiency, initial spindown

• how does SMBH react? i.e. what are torques on hole?

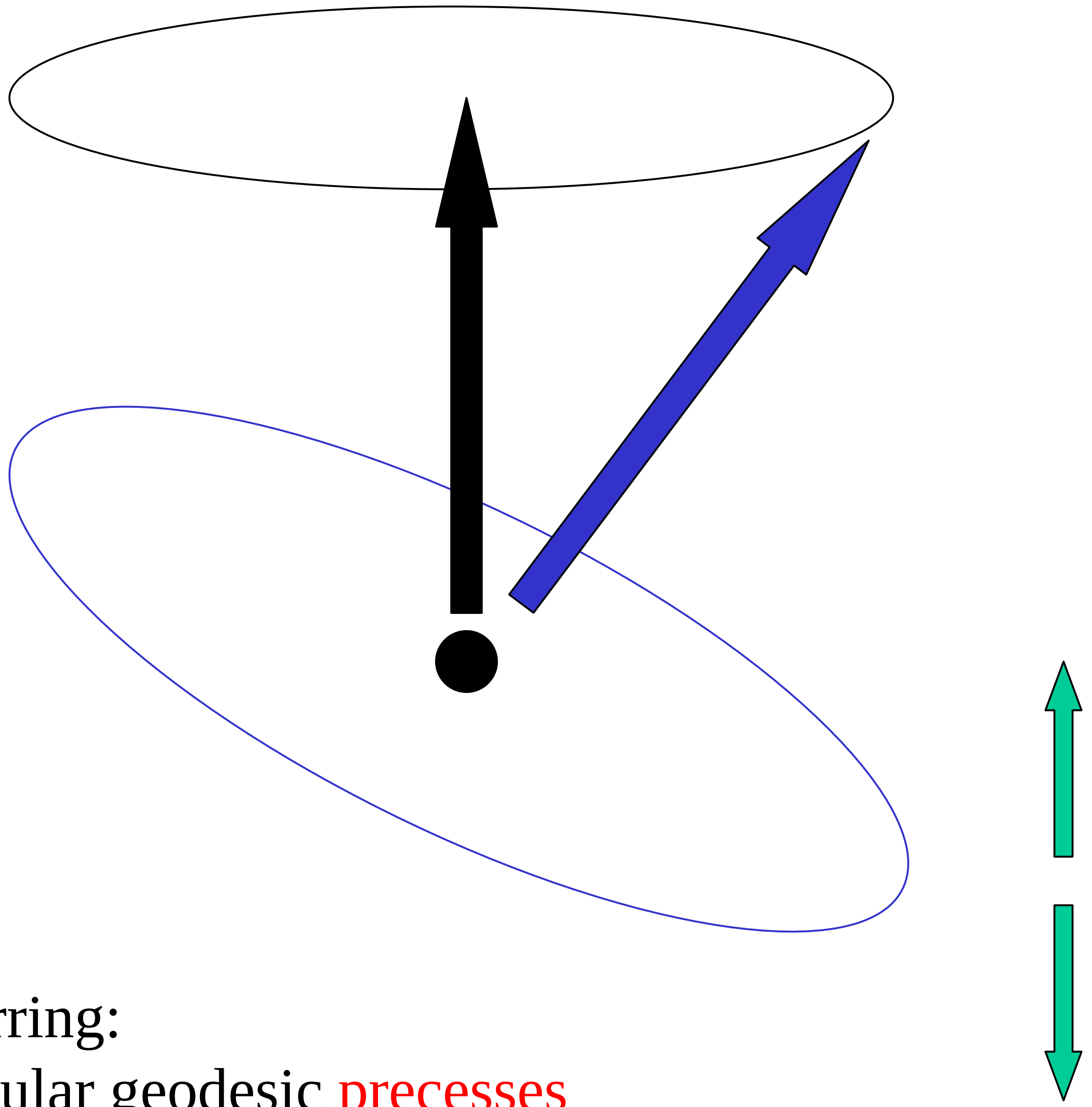
• two main types: 1. accretion – spinup or spindown – requires hole to accrete ~ its own mass to change *a/M* significantly — *slow*

2. Lense—Thirring from misaligned disc viscous timescale — *fast* in inner disc

• standard argument: *alignment* via Lense—Thirring occurs *rapidly*, hole spins up to keep a ~ M, accretion efficiency is *high*

• but L—T *also* vanishes for *counteralignment*

• alignment or not? (King, Lubow, Ogilvie & Pringle 05)



Lense—Thirring:
plane of circular geodesic precesses
about black hole spin axis: dissipation causes alignment or
counteralignment

Torque on hole is pure precession, so *orthogonal to spin*.

Thus general equation for spin evolution is

$$\frac{\mathrm{d}\mathbf{J}_h}{\mathrm{d}t} = -K_1[\mathbf{J}_h \wedge \mathbf{J}_d] - K_2[\mathbf{J}_h \wedge (\mathbf{J}_h \wedge \mathbf{J}_d)].$$

Here K_1, $K_2 > 0$ depend on disc properties. First term specifies precession, second alignment.

Clearly magnitude J_h is constant, and vector sum $\mathbf{J}_t$ of $\mathbf{J}_h$, $\mathbf{J}_d$ is constant. Thus $\mathbf{J}_t$ stays fixed, while tip of $\mathbf{J}_h$ moves on a sphere during alignment.

Using these, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{J}_h \cdot \mathbf{J}_t) = \mathbf{J}_t \cdot \frac{\mathrm{d}\mathbf{J}_h}{\mathrm{d}t} = \mathbf{J}_\mathrm{d} \cdot \frac{\mathrm{d}\mathbf{J}_h}{\mathrm{d}t}.$$

thus

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{J}_h \cdot \mathbf{J}_t) = K_2[J_d^2 J_h^2 - (\mathbf{J}_d \cdot \mathbf{J}_h)^2] \equiv A > 0.$$

But J_h, J_t are constant, so angle θ_h between them obeys

$$\frac{\mathrm{d}}{\mathrm{d}t}(\cos\theta_h) > 0$$

— hole spin *always* aligns with *total* angular momentum

Can further show that J_d^2 always *decreases* during this process – dissipation

Thus viewed in frame precessing with $\mathbf{J}_h$, $\mathbf{J}_d$,

$\mathbf{J}_t$ stays fixed: $\mathbf{J}_h$ aligns with it while keeping its length constant

J_d^2 decreases monotonically because of dissipation

Since

$$J_t^2 = J_h^2 + J_d^2 - 2J_h J_d \cos(\pi - \theta)$$

there are two cases, depending on whether

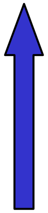
$$\cos\theta < -\frac{J_d}{2J_h}$$

or not. If this condition fails, $J_t > J_h$ and alignment follows in the usual way – older treatments implicitly assume

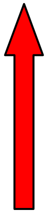
$$J_d >> J_h$$

so predicted alignment

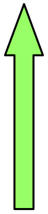
$\mathbf{J}_h =$

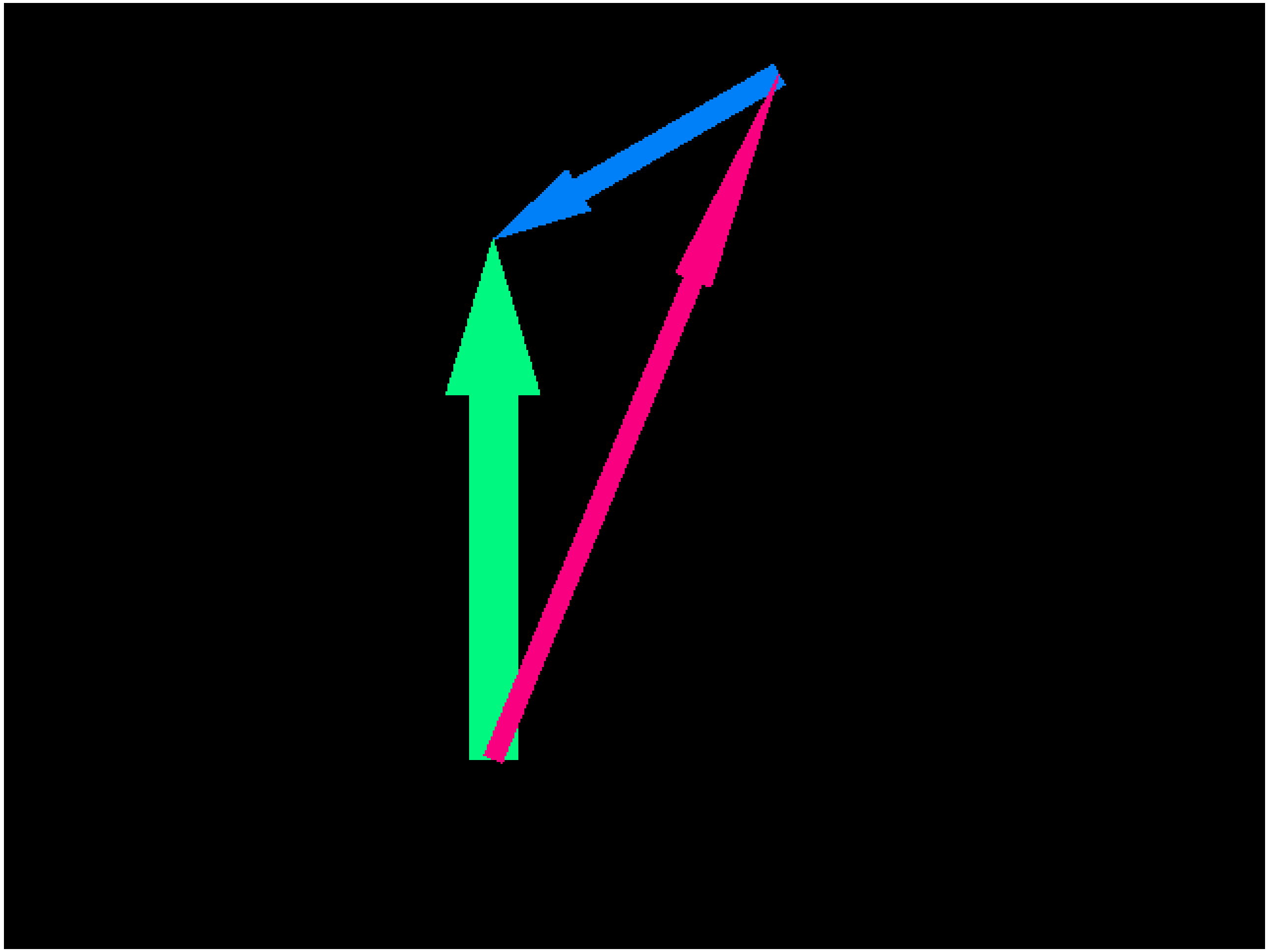


$\mathbf{J}_d =$



$\mathbf{J}_t = \mathbf{J}_h + \mathbf{J}_d =$



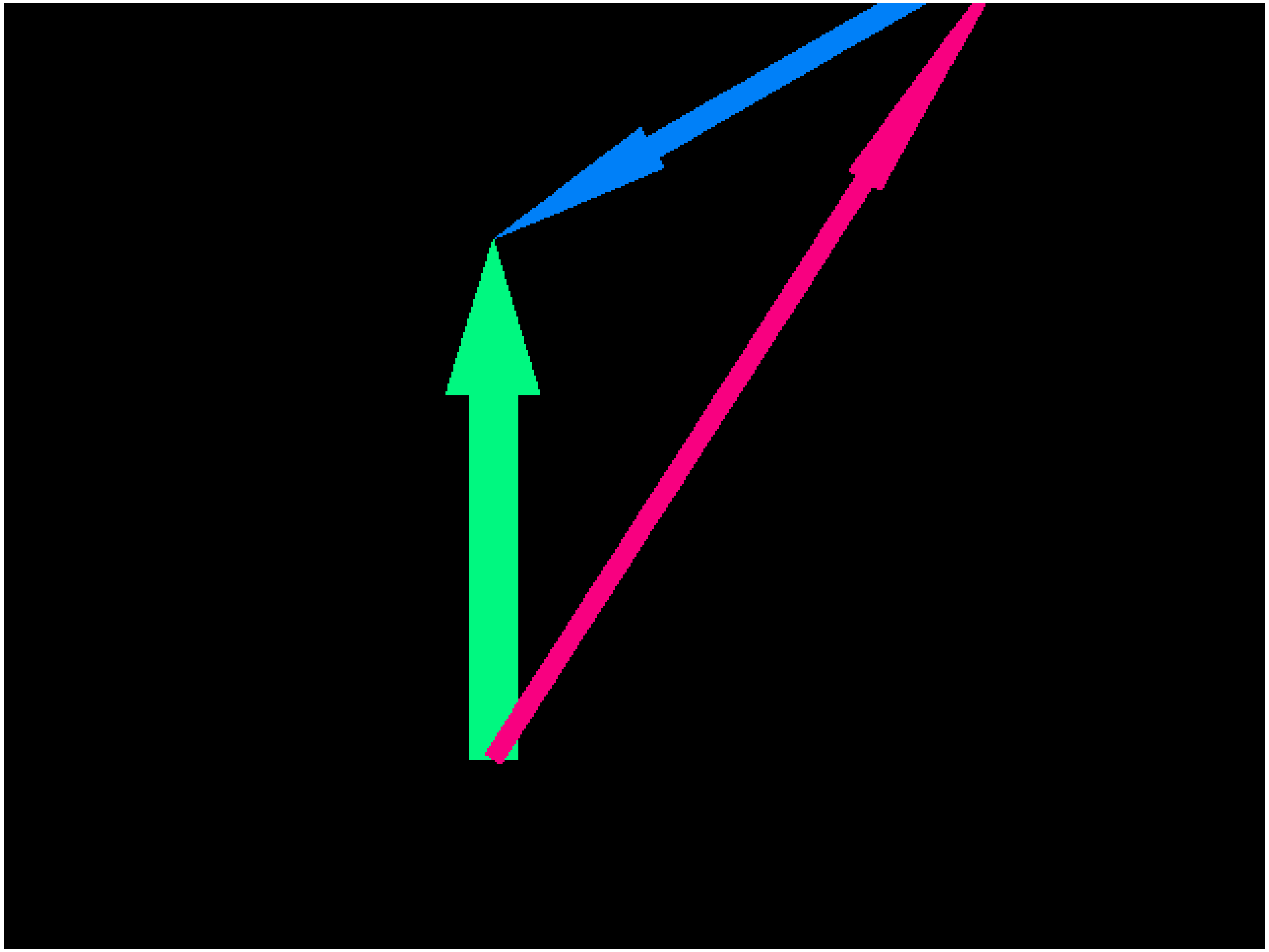


but if $\cos\theta < -\frac{J_d}{2J_h}$ *does* hold,

which requires $\theta > \pi/2$ and $J_d < 2J_h$,

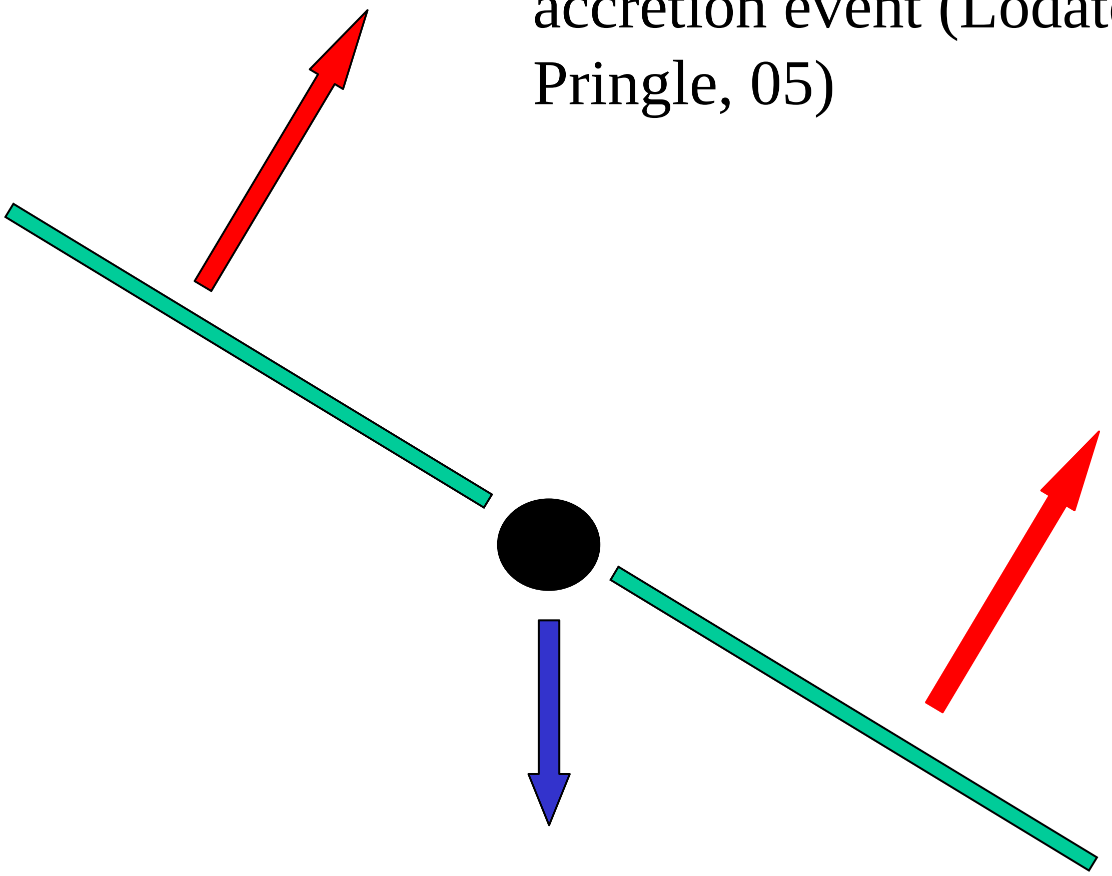
then $J_t < J_h$, and

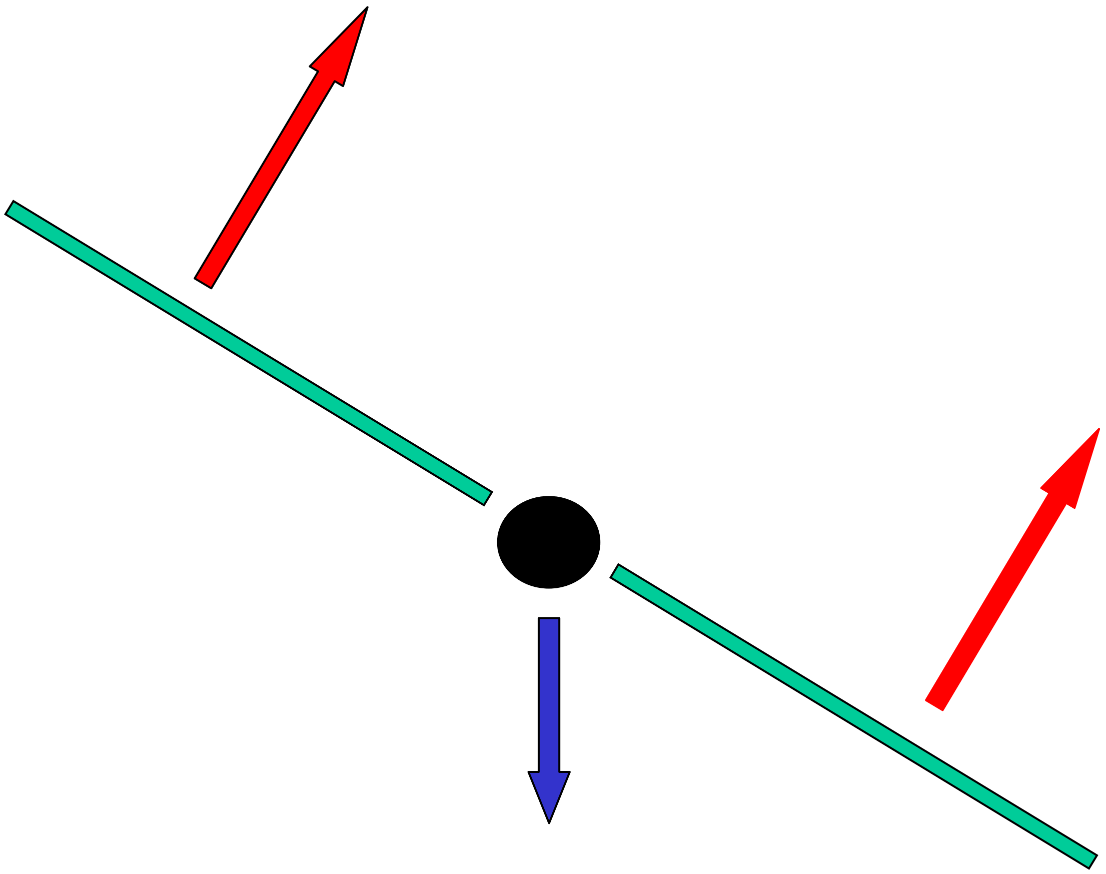
counteralignment occurs



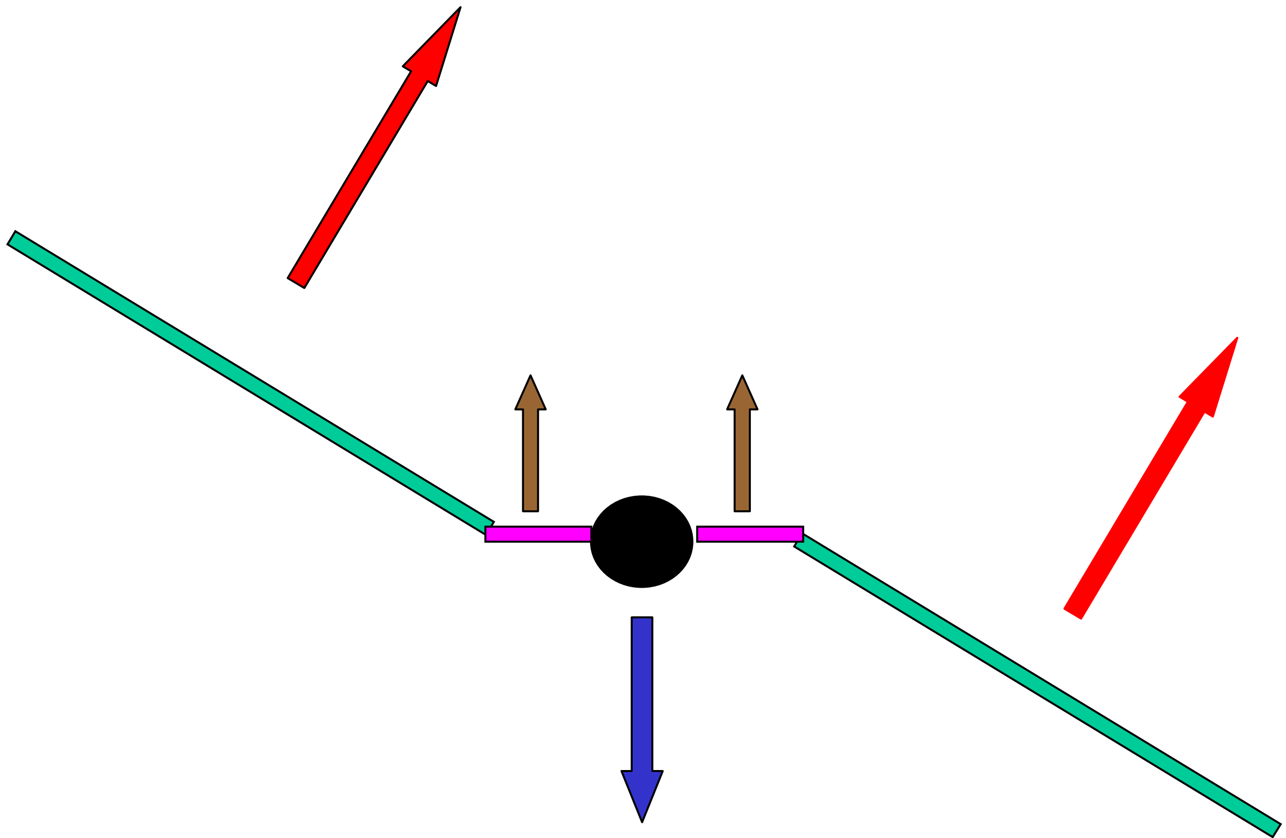
- small counterrotating discs anti—align
- large ones align
- what happens in general?

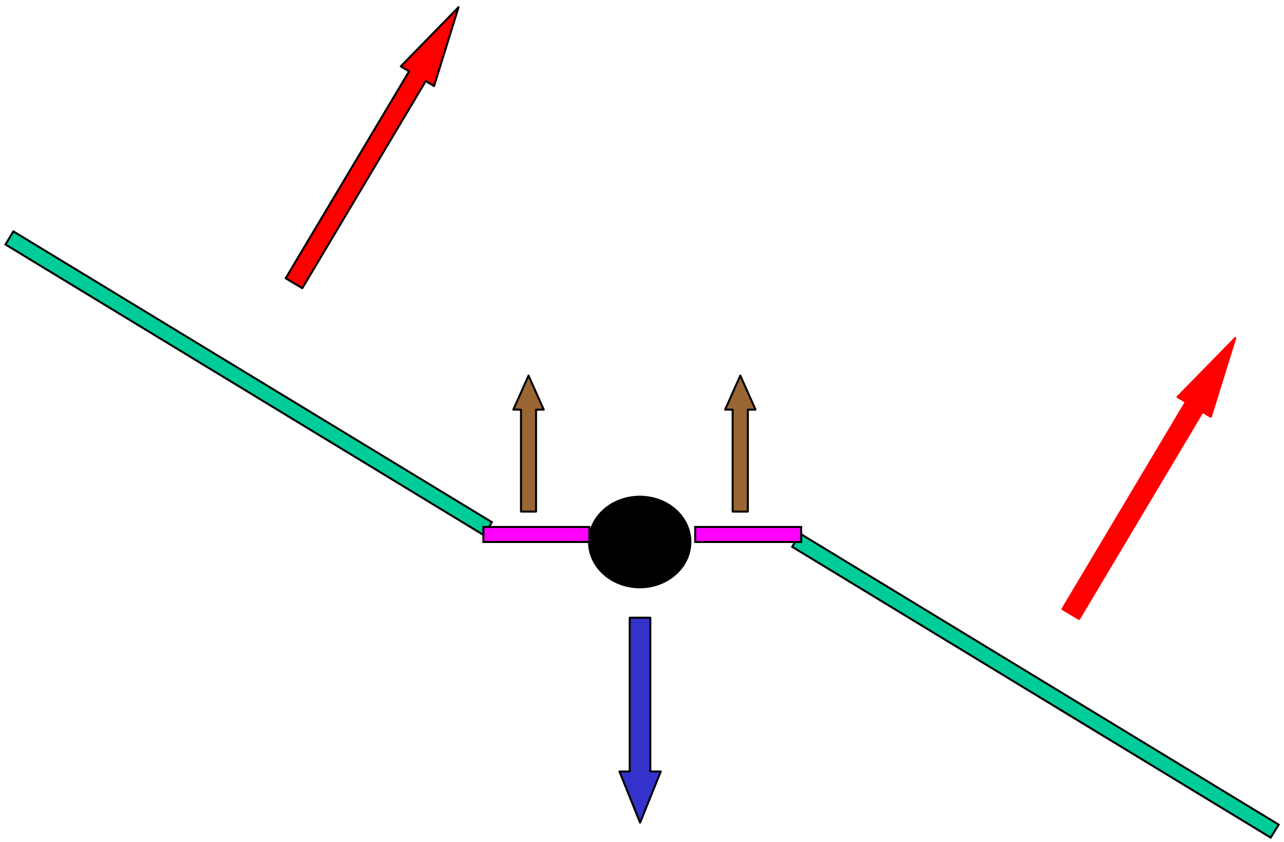
consider an initially counteraligned accretion event (Lodato & Pringle, 05)



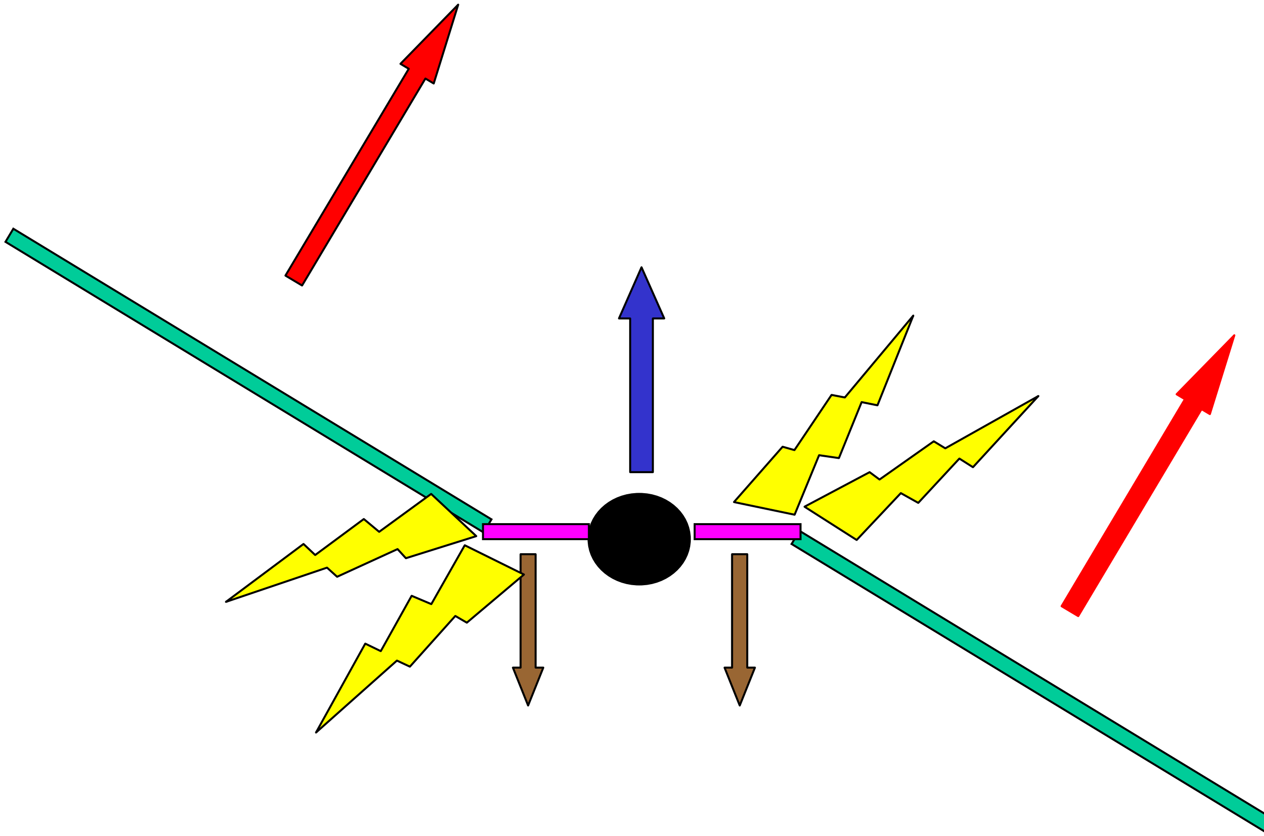


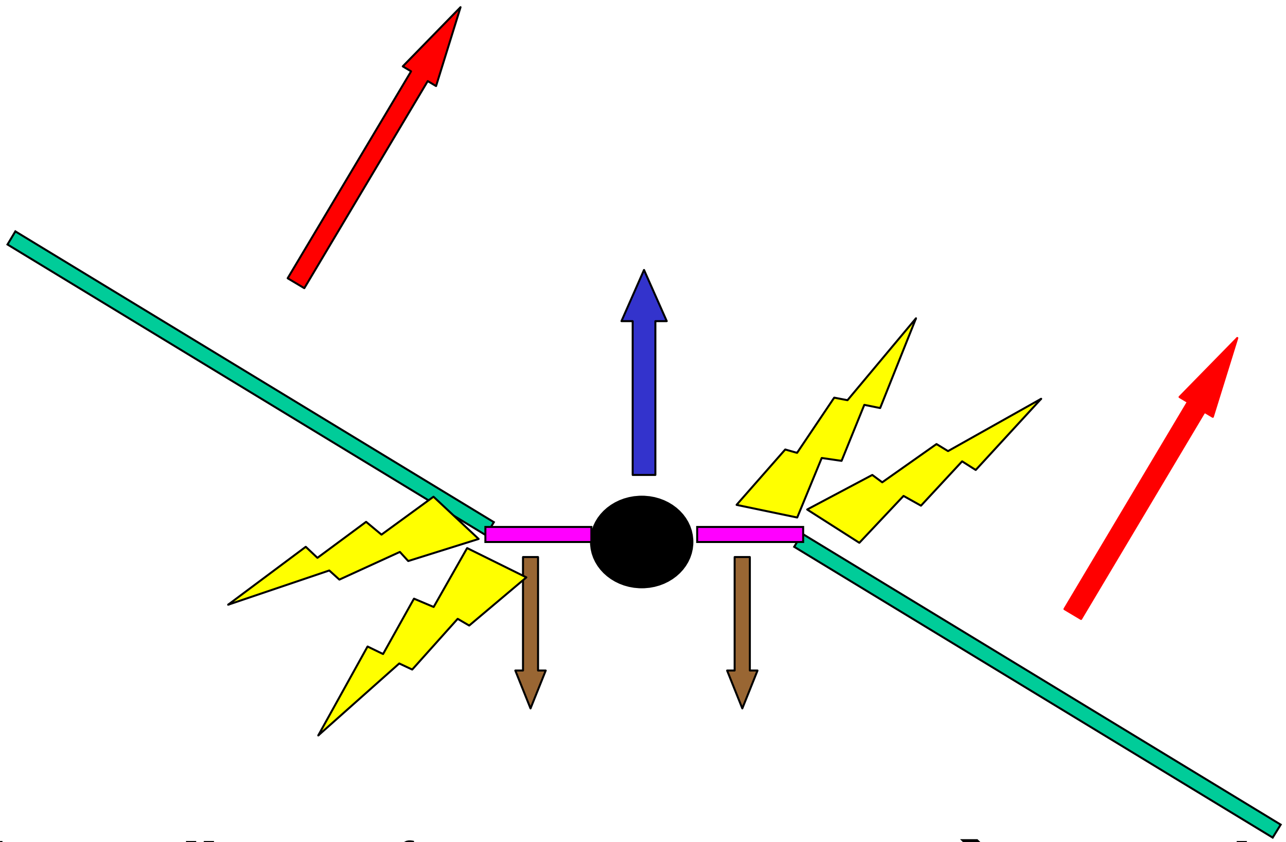
L—T effect first acts on *inner disc*: less a.m. than hole, so central disc *counteraligns*, connected to outer disc by warp: timescale $< 10^8 \mathrm{yr}$





but outer disc has more a.m. than hole, so forces it to *align*, ***taking counteraligned inner disc with it***





resulting collision of counterrotating gas ➔ intense dissipation ➔ high central accretion rate

accretion efficiency initially low (retrograde): a/M may be lower too

- *merger origin of AGN ➔ super—Eddington accretion ➔ outflows*

- these can explain

  1. M—sigma
  2. starbursts simultaneous with BH growth
  3. BH—bulge mass correlation
  4. matter accreting in AGN has high metallicity
  5. superwind connection

- about one—half of merger events lead to

  1. initial retrograde accretion — low efficiency, lower *a/M*
  2. outbursts