Relativistic Astrophysics

Charles Gammie

July 25, 2016

1 Lecture 1 — Accretion Disks

1.1 Child's Garden of Astrophysical Disks

Here is a (incomplete) list of astrophysical disks

- Galactic disk: spiral/elliptical
- Supermassive BH: quasar/Seyfert/LINER/LLAGN/TDE
- Stellar mass BH: microquasar/GRB
- Neutron star: LMXB/HMXB/GRB
- White dwarf: dwarf nova/nova
- Protostar: protoplanatory/debris
- Planet: protolunar disk/planetary rings

Consider the galaxy NGC 4258. If we zoom in to 0.1 pc, we see maser spots around the central engine. The velocities of these spots fall nicely on a Keplerian orbit. The combination of the acceleration, velocity, and angular position of these maser spots allows us to measure to mass of the central object.

Here let us introduce the first dimensionless parameter of a thin disk

$$\frac{H}{R} \ll 1, \quad H = \frac{c_s}{\Omega} \propto T^{1/2} \tag{1.1}$$

For this system this parameter is $\sim 10^{-3}$. This is an example of a thin disk. The disk of NGC 4258 is not a flat disk, since one can see warps following the maser spots. The disk is actually made of rings at the positions of the masers.

This particular disk is relatively far away from the object, and the disk is heated due to the radiation from the central object.

The next example is the center of our galaxy. We know that the accretion flow near $\sim 10r_g$ emits synchrotron radiation which is the target of the Event Horizon Telescope (EHT). The mean free path of Coulomb scattering in Sgr A* is very large, so the plasma is almost collisionless. This allows us to introduce the second number

$$\kappa_n = \frac{\lambda_{mfp}}{R} \tag{1.2}$$

which is about 10^5 for this system.

The next example is the HL Tau. This is a cold disk of protoplanetary disk. At 50 au the temperature is less than $100\,K$. Here we introduce the third number which is the magnetic Reynolds number

$$\mathrm{Re}_M = \frac{c_s H}{\eta} \tag{1.3}$$

In regions of this system that we can measure with ALMA, this number is less than 1.

The last number we introduce is

$$Q = \frac{c_s \Omega}{\pi G \Sigma} \tag{1.4}$$

where Σ is the surface density of the disk. When this number is less than one, self gravity becomes important.

The last example is the moon. When people look at the moon most people do not see a disk. The theory of the formation of the moon involves an impact on earth, throwing some mass from the earth which forms a disk around it. This is a system where the magnetic Reynolds number might play a role.

One more system we want to mention is the SS Cyg which is a Cataclysmic variable, formed by a massive star overflowing its Roche lobe, feeding a disk around its companion which is a WD. We have observed this system since 1896 and it alternates between a quiescent state and an active state. This system may have a low magnetic Reynolds number at quiescent state, but has a high magnetic turbulence during its outbursts.

1.2 Disk Evolution

Disks are a class of special objects in astrophysics. The angular momentum is conserved, but its kinetic energy is easily converted to thermal energy due to various instabilities in the disk, and eventually radiated away. Therefore understanding the disk evolution is all about understanding the evolution of angular momentum.

In a thin disk, dynamical equilibrium is reached in a dynamic timescale $\Delta t \sim \Omega^{-1}$, where

$$\Omega = \left(\frac{GM}{R^3}\right) 1/2 + O\left(\frac{H}{R}\right)^2 \tag{1.5}$$

However it is possible for there to be long living excitations which live in circular orbits, in addition to tilts and warps in the disk.

Thermal equilibrium is reached when $Q^+ \sim Q^-$, when heating rate balances the cooling rate. The time scale is

$$\Delta t \sim \Sigma c_s^2 / Q^+ \sim (\alpha\Omega)^{-1} \tag{1.6}$$

where the α parameter is very important and it describes the intensity of turbulence in the disk, and the above equation can be taken as the definition. It relates the thermal time scale to the dynamical time scale. Its value, usually at 10^{-2} means that the thermal time scale is much longer than the dynamical time scale.

The last time scale is the inflow equilibrium $\dot{M} \sim$ const. The time scale is

$$\Delta t \sim \frac{M_{\mathrm{disk}}}{\dot{M}} \sim (\alpha\Omega)^{-1} \left(\frac{R}{H}\right)^2 \tag{1.7}$$

which is again longer by a factor of the scale height squared.

Lets write down the disk evolution equation

$$\partial_t \Sigma = \frac{2}{r}\partial_r \left(\frac{\Omega}{r\kappa^2}\partial_r (r^2 W_{r\phi}) - \frac{\Omega}{\kappa^2}\tau \right) + \dot{\Sigma}_{\rm ext} \tag{1.8}$$

where Σ is the surface density, Ω the orbital frequency, κ the epicyclic frequency which is similar to Ω, $W_{r\phi}$ is the shear stress, τ is connected to external torques per area, and finally $\dot{\Sigma}_{\rm ext}$ is the mass gain/loss from infall or from wind loss. The full derivation of this equation is left as an exercise. The hint is to start from the angular momentum conservation equation together with conservation of mass. Another hint is that

$$\frac{d}{dr} j = \frac{r\kappa^2}{2\Omega} \tag{1.9}$$

where j is the specific angular momentum.

The standard model for disks is the α disk model. It was introduced by (Shakura & Sunyaev 1973) and (Linden-Bell & Pringle 1974). It adopts simple scaling argument for diffusion of angular momentum by turbulence, by just modeling the turbulent viscosity

$$\nu \sim \alpha c_s H \tag{1.10}$$

In classical α disk models one ignores external torques, infall/winds, etc. Thus one throws away the last two terms in the disk equation, with only the term with $W_{r\phi}$ surviving.

In the α disk model we assume the disk is thin and Keplerian $\Omega \approx (GM/r^3)^{1/2}$. We also assume vertical equilibrium $H \sim c_s/\Omega$. The opacity is assumed to be related to temperature and density $\kappa \sim \kappa_0 \rho^a T^b$. We have a prescription of vertical integration of the density $\Sigma \sim 2\rho H$, which is a crude integral. We also estimate the optical depth to be $\tau \sim \Sigma\kappa/2$. The surface temperature is determined by the fact that heating by viscous processes $F = \sigma T_{\rm eff}^4 \sim (9/8)\Sigma\nu\Omega^2$, where $T_{\rm eff}^4 \sim \sigma T^4/\tau$. Finally in the steady state $\dot{M} = 3\pi\Sigma\nu$.

As an example, consider a steady state disk around a stellar mass BH. In the inner zone, radiation pressure is much larger than the gas pressure, and electron scattering dominates the opacity. We can find that

$$T \sim 4.3 \times 10^7 \alpha^{-1/4} m^{-1/4} x^{-3/8}\, K \tag{1.11}$$

$$\Sigma \sim 0.4 x^{3/2} \alpha^{-1} \dot{m}^{-1}\, \mathrm{g/cm^2} \tag{1.12}$$

$$H/r \sim 10 \dot{m} x^{-1} \tag{1.13}$$

Now lets come back to the disk evolution equation and ask what are $W_{r\phi}$, τ, and $\Sigma_{\rm ext}$. Part of the answer only lies in the global analysis of turbulence in the disks.

1.3 Turbulence in Disks

In α disk model the turbulent diffusion of angular momentum is simply postulated. Possible sources of turbulence include magnetorotational instability (MRI), gravitational instability, zombie vortex instability, subcritical baroclinic instability, and vertical shear instability. The last three are related to instabilities in fluids. The zombie vortex instability is generated due to vortices generated in the disk sharing angular momentum through sound waves. The subcritical baroclinic instability is loosely related to the baroclinic instability that drives weather. The vertical shear instability is when shearing motion is faster than orbital motion and driving the flow of angular momentum at sound speed.

We will focus on the MRI instability (Balbus & Hawley 1991). This is due to local, linear instability of weakly magnetized disks driven by exchange of angular momentum. The condition of the instability does not involve boundary conditions and occur in a localized region of the disk, and can be treated using the WKB approximation.

We will start with a simple mechanical analogy where 2 masses connected by a spring are in orbit around a central body. The orbital frequency is given by Keplerian motion $\Omega^2 = GM/R^3$. The masses can be thought as fluid elements, and spring can be a model of magnetic field line tension which provides restoring force. We denote the frequency of the spring as γ.

Lets setup x-y coordinates for the orbital motion where x is in radial direction, and y points opposite to the orbital motion. We know that

$$\ddot{x} = -2\Omega\dot{y} + 3\Omega^2 x - \gamma^2 x \tag{1.14}$$

The first term is the Coriolis force, and the second combines the centrifugal force and the gravitational force in an expansion, and the third term is the spring force. The second equation is

$$\ddot{y} = 2\Omega\dot{x} - \gamma^2 y \tag{1.15}$$

As a first go, lets set $\gamma = 0$ and assume x and y scale with $e^{-i\omega t}$, the equations become

$$-\omega^2 x = -2\Omega(-i\omega)y + 3\Omega^2 x \tag{1.16}$$

$$-\omega^2 y = 2\Omega(-i\omega)x \tag{1.17}$$

This simply gives $\omega^2 = \Omega^2$ which means that the natural modes oscillate at the orbital frequency.

Now lets add the γ term, and we get

$$-\omega^2 x = -2\Omega(-i\omega)y + 3\Omega^2 x - \gamma^2 x \tag{1.18}$$

$$-\omega^2 y = 2\Omega(-i\omega)x - \gamma^2 y \tag{1.19}$$

And solving the dispersion relation gives

$$\omega^4 - \omega^2(\Omega^2 + 2\gamma^2) + \gamma^2(\gamma^2 - 3\Omega^2) = 0 \tag{1.20}$$

Notice that there is a zero solution condition when $\gamma^2 = 3\Omega^2$, and it is a transition from positive to negative frequency square. If we plot ω^2/Ω^2 vs γ^2/Ω^2, there are two sets of roots. One is simply spring motion. The other starts with $\omega^2 = 0$, and between $\gamma^2 = 0$ and $3\Omega^2$ we have $\omega^2 < 0$ where there is an instability. The fastest growing mode of the instability is at $\gamma^2/\Omega^2 = 15/16$ and $\omega^2/\Omega^2 = -9/16$.

This is directly analogous to an MHD problem. Lets zoom in on the disk and assume it is penetrated by a weak vertical magnetic field. This instability is related to the magnetic instability where the field is bent, and the fluid element above and below are coupled by Alfven waves. In this case we have

$$\gamma^2 \to (\mathbf{k} \cdot \mathbf{v}_A)^2 \tag{1.21}$$

Lets list some facts about MRI linear theory. The ideal fluid instability only requires $d\Omega^2/dr < 0$. The maximum growth rate is $(3/4)\Omega$ for Keplerian motion. The fastest growing mode is given above. We know that there is a local instability when there is vertical field, but a local instability also occurs even when there is only toroidal field.

However linear theory doesn't tell us what happens in the regime when linear growth saturates. To study this regime we need simulations. In simulations people have used local or global setups. There are models with explicit small dissipation terms, or ILES which self-consistently evolves turbulence cascade from the scale where energy is injected down to the scale of dissipation. There are also isothermal models with equation of state $P = c_s^2\rho$, or one can use energetically self-consistent equation of state.

Some simulations were shown...

Simulations of MRI told us a few things. In 2D MRI leads to turbulence and α. We know that in 2D MRI does not converge, so α depends on resolution. In 3D MRI also leads to turbulence and α, but sometimes it does converge. We know that it converges when explicit dissipation is used. There are issues with ILES models with unstratified shearing boxes. We also learned that α depends on many parameters. It depends on the local height z, and it increases away from the central plane. It also depends on the magnitude of the vertical magnetic field that threads the disk $\langle B_z\rangle$. When the vertical magnetic field increases the viscosity also increases. It depends on the magnetic Reynolds number and $\mathrm{Pr}_M = \nu/\eta$.

1.4 Current Problems in Disk Theory

Ran out of time...

2 Lecture 2 — Relativistic MHD

2.1 When is Relativistic MHD required?

The first situation when we need relativistic MHD is when one of the wave speeds is of order c. The characteristic wave speeds are usual one of $\mathbf{v}$, $\mathbf{v}_A$, and $\mathbf{v}_s$. The nonrelativistic sound speed can be defined as

$$c_{s,NP}^2 = \gamma\frac{P}{\rho} \tag{2.1}$$

which blows up at high γ and low ρ. If we do it correctly with relativistic theory we can find that

$$c_{s,R}^2 = \frac{\gamma P}{\rho + \frac{\gamma}{\gamma-1}\frac{P}{c^2}} = \frac{c_{s,NR}^2}{1 + c_{s,NR}^2/(\gamma-1)/c^2} \tag{2.2}$$

In relativistic limit we have $\gamma = 4/3$, $\rho = 3P$, and we have sound speed as $c_s^2 = c^2/3$. If we still use ideal gas law $P = nkT$, and $n = \rho/m$ then since we have in the relativistic limit $P \sim \rho c^2$

$$\theta = \frac{kT}{mc^2} \sim 1 \tag{2.3}$$

This dimensionless number for electron $\theta_e = 1$ means $T \sim 5.9 \times 10^7\,K$, and $\theta_p = 1$ means $T \sim 1.1 \times 10^{13}\,K$.

Lets look at Alfven speed. For nonrelativistic case it is defined as $v_A = B/\sqrt{4\pi\rho}$, which blows up when $\rho \to 0$. We are saved in relativistic case by

$$v_{A,R} = \frac{|\mathbf{B}|}{\sqrt{4\pi\rho + B^2/c^2}} = \frac{v_{A,NR}}{\sqrt{1 + v_{A,NR}^2/c^2}} \tag{2.4}$$

Therefore in relativistic case B field inertial comes into play. The dimensionless ratio here is

$$\sigma = \frac{B^2}{4\pi\rho c^2} \tag{2.5}$$

Without looking at speeds, we also need relativistic hydro when

$$\phi \sim \frac{GM}{rc^2} \sim 1 \tag{2.6}$$

which accounts for gravity effects.

2.2 Basic equations

Lets write down the equations for relativistic MHD in conservative form in terms of fluxes and conserved quantities

$$\partial_t \mathsf{U} = -\nabla \cdot \mathsf{F} + \mathsf{S} \tag{2.7}$$

Now we need to be careful about what are the conserved quantities and what is the divergence operator. What we are going to do is write down the nonrelativistic versions and use them to motivate the transition to relativistic version.

The mass conservation equation looks like

$$\partial_t \rho = -\nabla \cdot (\rho \mathbf{v}) = -\partial_i (\rho v_i) \tag{2.8}$$

The relativistic generalization is straightforward:

$$\nabla_\mu (\rho u^\mu) = 0 \tag{2.9}$$

where $u^\mu = dx^\mu / d\tau$. We need to define density a bit more carefully. ρ is defined using the particle density in a frame that is comoving with the plasma. The ∇_μ operator is the covariant derivative operator which was covered in lecture by Lehner. We define the gamma factor of the particles as $u^0 = \Gamma$, and $u^i = \Gamma v^i$. The 4 velocity is normalized with $u^\mu u_\mu = -1$.

The mass conservation equation can be expanded out in a coordinate in the following form:

$$\partial_t (\sqrt{-g} \rho u^t) = -\partial_i (\sqrt{-g} \rho u^i) \tag{2.10}$$

where g is the determinant of the metric tensor. This is because the 4-divergence can be written as $1/\sqrt{-g}\partial_\mu(\sqrt{-g}\rho u^\mu)$. This equation now looks like the form $\partial_t \mathsf{U} = -\nabla \cdot \mathsf{F}$.

Now we look at energy momentum conservation equations. The nonrelativistic form looks like

$$\partial_t \left(\frac{1}{2}\rho v^2 + U + \frac{B^2}{8\pi} \right) = -\partial_i \left(\frac{1}{2}\rho v^2 v_i + (U+P)v_i + \frac{1}{4\pi}\left(B^2 v_i - (\mathbf{B}\cdot\mathbf{v})B_i \right) \right) \tag{2.11}$$

$$\partial_t (\rho v_i) = -\partial_j \Pi_{ij} \tag{2.12}$$

where

$$\Pi_{ij} = \rho v_i v_j + P\delta_{ij} + \frac{B^2}{8\pi}\delta_{ij} - \frac{B_i B_j}{4\pi} \tag{2.13}$$

In relativistic case these two equations combine to have the following simple form

$$\nabla_\mu T^{\mu\nu} = 0 \tag{2.14}$$

All the physics is now in the stress-energy tensor. This covariant derivative can be expanded in the coordinate form

$$\partial_t (\sqrt{-g} T^t{}_\mu) = -\partial_i \left(\sqrt{-g} T^i{}_\mu \right) + \sqrt{-g} T^\kappa{}_\lambda \Gamma^\lambda{}_{\mu\kappa} \tag{2.15}$$

where Γ is the connection related to $\partial g_{\mu\nu}$.

Now lets write down the stress-energy tensor

$$T_{\mu\nu} = (\rho + U + P + b^2)u_\mu u_\nu + \left(P + \frac{b^2}{2}\right) g_{\mu\nu} - b_\mu b_\nu \tag{2.16}$$

where b_μ has spatial components b_i which is just the spatial magnetic field. In the fluid rest frame this tensor is diagonal, and if we choose B to be aligned with z direction then we have

$$T_{\mu\nu} = \text{diag}\left(\rho + U + \frac{b^2}{2}, P + \frac{b^2}{2}, P + \frac{b^2}{2}, P - \frac{b^2}{2}\right) \tag{2.17}$$

Notice that the first term in the stress energy tensor has energy terms, so we already take into account the temperature of the plasma in the energy.

Lets look at the magnetic field equation $\partial_t \mathbf{B} = \nabla \times (\mathbf{v} \times \mathbf{B}) = -\partial_j(v_j B_i - v_i B_j)$. It is easy to generalize this by promoting ∂_i to ∇_i and v, B to u and b, we have $\nabla_\mu(u^\mu b^\nu - u^\nu b^\mu) = 0$, and in coordinate basis it looks like:

$$\partial_t \left(\sqrt{-g}(u^t b^i - u^i b^t)\right) = -\partial_i \left(\sqrt{-g}(u^j b^i - u^i b^j)\right) \tag{2.18}$$

This equation has 3 components, and the 4th component is simply the no-monopole constraint

$$\partial_i(\sqrt{-g}\tilde{B}^i) = 0 \tag{2.19}$$

where we define $\tilde{B}^i = b^i u^t - u^i b^t$.

Now we can combine all these equations and write down what is U, F, and S. The perimeter variables we want to use are ρ, U, $\tilde{B}^i$, and u^i. However u^i are not convenient in e.g. black hole ergospheres, so we define

$$\tilde{u}^i = (g^i{}_\mu + n^i n_\mu)u^\mu \tag{2.20}$$

We can transform these into conserved variables U as

$$\left(\sqrt{-g}\rho u^t, \sqrt{-g}(\rho + U + P + b^2)u^t u_t + (P + b^2/2)g^t_t - b^t b_t, \sqrt{-g}\tilde{B}^i\right) \tag{2.21}$$

Lets look at the Kerr metric. The Boyl-Lindquist coordinates is t, r, θ, ϕ, and we have $\partial_\phi = 0$ and $\partial_t = 0$. These two represent two Killing vectors ξ which satisfy

$$\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0 \tag{2.22}$$

For example $\xi^\mu_t = (1, 0, 0, 0)$, and similar for ξ^μ_ϕ. If we define conserved current as $J^\mu = T^\mu_\nu \xi^\nu$, then we automatically have $\nabla_\mu J^\mu = 0$ due to the definition of Killing vectors.

2.3 Numerical techniques

There are about a dozen codes out there that can integrate relativistic MHD equations. We start from perimeter variables P^n at timestep n and find the conserved variables U^n, integrate the equations to get U^{n+1} and do an inverse transform to get P^{n+1} which usually involves numerically solving systems of nonlinear equations.

In HARM we don't solve Riemann problems, and we need to estimate $v_\pm$. It is diffusive in comparison to Athena++, but it is much simpler. We still need to keep $\nabla \cdot \mathbf{B} = 0$. One could simply do nothing and hope for the best, but that proved to not work. Another solution is to use the so-called constraint transport, which embed this constraint into the transport equation. Yet another approach is divergence cleaning, which solves periodically the closest B field which has no divergence. A related approach is to evolve an additional field which causes in the induction equation an extra term that relaxes deviation from the constraint to zero over time. The final one is simply to evolve the vector potential, which is guaranteed to have B field divergence free.

The procedures to compute the time evolution of the conserved variables are now pretty well understood. However there is no analytic approach to get the fluxes so that is a challenge. It is also understood now that the inversion process in the end to get the perimeter variables does not dominate the numerical complexity.

One problem that is common for both RMHD and MHD is that sometimes evolution takes conserved quantities to nonphysical values. Some codes simply inspect the quantities and for example put a floor on density values to prevent it from going negative. The reason one could do this is that this kind of scenario shows up in extreme cases and the change of a small negative density to zero does not change the system too much. In HARM it is imposed that

$$\rho > \rho_{\min}, \quad U > U_{\min}, \quad -n^\mu u_\mu < \Gamma_{\max} \tag{2.23}$$

In the procedure of this fixup, one introduces error to the energy and temperature of the system, which is somewhat inevitable.

Feynman said: “The first principle is that you must not fool yourself, and you are the easiest person to fool”. One need to test the code to show correct convergence behavior. Typical code tests include linear wave convergence which only tests the code in the linear regime. Another test is Orszag-Tang vortex (Gammie et. al. 2003). We tested the code against another code VAC. One could also test against changing resolution. We also tested it with Kerr inflow (reversed Parker wind) which has an “exact” solution for us to compare with.

Field loop advection problem and Komissarov’s sadistic explosion problem...

2.4 Beyond ideal MHD

In low $\dot{M}$ BH accretion flows we have $\lambda_{mfp,\parallel}$ for Coulomb scattering much larger than the gravitation scale GM/c^2. However the perpendicular mean free path is still very small. The associated viscosity $\nu \sim v_{\rm th}\partial_{mfp}$ would be highly anisotropic. Particles can carry momentum along the field lines but not across field lines. There is also anisotropic thermal conduction. Since Coulomb collision is not very efficient the electron and ion temperatures will decouple.

One theory include the viscosity and thermal conduction is the covariant extended MHD (Chandra+ 2015). This amounts to adding additional terms to the stress-energy tensor. In this model we add heat flux $q^\mu = qb^\mu$, where b^μ is a unit spacelike four-vector parallel to the b field. There is also possible momentum transport parallel to the field

$$\tau^{\mu\nu} = -\Delta P \left(b^\mu b^\nu - (1/3)h^{\mu\nu}\right) \tag{2.24}$$

where $h^{\mu\nu}$ is the projection tensor perpendicular to u^μ and u^μ is the 4-velocity. This is similar to the Braginskii stress tensor.

A naive theory setting $q \sim \nabla T + Ta$ which is a combination of temperature gradient and 4-acceleration is unstable, mostly due to the 4-acceleration term. One needs to promote q, ΔP to dependent variables and evolve them. One need to have equations for q_0, ΔP_0, and relaxation equations for these. Finally one needs to have closure relation for χ and ν.

An inspirational paper is (Komissarov 1999) called “A Godunov-type scheme for relativistic magnetohydrodynamics”.