CLAUDE OPUS 4.5

EDITED BY ADAM S. JERMYN

LECTURES ON DISKS AND JETS

Contents

Preface

These notes explore one of the most beautiful and universal phenomena in astrophysics: the formation of rotating disks around compact objects and the launching of powerful jets from those disks. From young stars still gathering their mass to the supermassive black holes at the centers of galaxies, nature repeatedly produces the same striking geometry—a thin, rotating disk of gas surrounding a central gravitational well, with narrow beams of material shooting outward along the rotation axis.

Why should this be? The answer lies in the interplay of gravity, angular momentum, and magnetic fields. A gas cloud that begins to collapse under its own gravity cannot simply fall inward if it carries any rotation. The rotation speeds up as the cloud shrinks, just as a figure skater spins faster when pulling in their arms. But somehow the material does reach the central object—billions of solar masses of gas have accumulated in quasars over cosmic time. And somehow, from this swirling infall, nature extracts energy and channels it into jets that can span millions of light-years.

These notes assume familiarity with classical mechanics, thermodynamics, and electromagnetism. We will develop the physics of disks from first principles, starting with the puzzle of why matter spirals rather than falls, and building toward an understanding of how jets might be launched and accelerated. The mathematics will emerge from the physics, not the other way around.

1
Why Matter Spirals

Drop a rock into a well and it falls straight down. Drop a cloud of gas onto a star and—well, it doesn't fall straight down. Instead, it swirls around, spreading into a flattened disk that might persist for millions of years before the gas finally reaches the central object. This simple observation leads to some of the most spectacular phenomena in the universe: the glowing disks around black holes that outshine entire galaxies, the jets of material that stretch across millions of light-years, and the dusty nurseries where planets form around young stars.

Why should gas behave so differently from a rock? The rock falls straight because we dropped it straight—we didn't give it any sideways motion. But a gas cloud in space almost always has some rotation. Maybe it inherited angular momentum from the larger structure it fragmented from. Maybe nearby objects gave it a gravitational nudge. Whatever the cause, even a tiny amount of initial rotation leads to dramatic consequences as the cloud shrinks under gravity.

1.1 The Skater's Spin

Every physics student learns the figure skater demonstration: a skater spinning with arms extended pulls their arms inward and spins faster. The explanation is conservation of angular momentum. If no external torque acts on the skater, the product of their moment of inertia I and angular velocity ω must remain constant:

$$L = I\omega = \text{constant}$$

When the skater pulls in their arms, I decreases, so ω must increase to compensate.

The same physics applies to a collapsing gas cloud, but the numbers are far more extreme. Consider a cloud of gas with mass $M = 10^{30}$ kg (about half a solar mass) and initial radius $R_0 = 10^{15}$ m (about 0.1 light-years). If this cloud is rotating with a period of 10

million years—hardly spinning at all by everyday standards—its angular velocity is

$$\omega_0 = \frac{2\pi}{10^7 \text{ yr}} = \frac{2\pi}{3 \times 10^{14} \text{ s}} \approx 2 \times 10^{-14} \text{ rad/s}$$

The angular momentum of this slowly rotating cloud, treating it crudely as a uniform sphere, is

$$L = \frac{2}{5} M R_0^2 \omega_0 \approx \frac{2}{5}(10^{30})(10^{15})^2(2 \times 10^{-14}) \approx 8 \times 10^{45} \text{ kg m}^2/\text{s}$$

Now suppose this cloud tries to collapse onto a star with radius $R_* = 7 \times 10^8$ m (like the Sun). If all the angular momentum were preserved and the material somehow reached the stellar surface, the rotation period would be

$$P = \frac{2\pi}{\omega} = \frac{2\pi I}{L} = \frac{4\pi M R_*^2}{5L}$$

Plugging in our numbers:

$$P = \frac{4\pi(10^{30})(7 \times 10^8)^2}{5(8 \times 10^{45})} \approx 150 \text{ seconds}$$

The star would need to rotate once every 2.5 minutes. At that rate, you'd need to hold on tight—except of course there wouldn't be much star left to hold onto. The maximum rotation rate for a star before it tears itself apart is roughly when centrifugal acceleration equals gravitational acceleration at the equator:

$$\omega_{\text{max}}^2 R_* = \frac{GM}{R_*^2}$$

For a solar-mass star, this gives $\omega_{\text{max}} \approx 6 \times 10^{-4}$ rad/s, or a minimum rotation period of about 3 hours. The angular momentum from our leisurely rotating cloud would spin the star about 70 times faster than it can survive.

Something has to give. Either the star cannot form from this cloud, or the angular momentum must go somewhere else.

You might say, "Well, perhaps stars just don't form from rotating clouds." But we observe them forming all the time, and the clouds *always* rotate. Every molecular cloud we've ever measured has angular momentum. The universe, it seems, does not provide stationary initial conditions.

1.2 *The Angular Momentum Barrier*

Let us think about this more carefully by following a small parcel of gas as it tries to fall inward. If the parcel is orbiting at radius r with

angular velocity ω, it feels two forces in the rotating reference frame: the gravitational pull inward and the centrifugal force outward.[1]

For a circular orbit, these forces balance:

$$\frac{v_\phi^2}{r} = \frac{GM}{r^2}$$

where $v_\phi = r\omega$ is the orbital velocity. Solving for v_ϕ:

$$v_\phi = \sqrt{\frac{GM}{r}}$$

This is the Keplerian orbital velocity. It increases as r decreases—material closer to the center orbits faster.

Now consider what happens if we try to push a parcel of gas inward while conserving its angular momentum. The specific angular momentum (angular momentum per unit mass) is

$$\ell = r v_\phi = r^2 \omega$$

If ℓ is conserved as the parcel moves to a smaller radius r', its new angular velocity must be

$$v'_\phi = \frac{\ell}{r'}$$

But the Keplerian velocity at r' is

$$v_K(r') = \sqrt{\frac{GM}{r'}}$$

Let's compare these. At the original radius r, the parcel was moving at the Keplerian velocity, so $\ell = r v_K(r) = r\sqrt{GM/r} = \sqrt{GMr}$. At the new radius:

$$v'_\phi = \frac{\sqrt{GMr}}{r'} = \sqrt{\frac{GMr}{r'^2}}$$

while

$$v_K(r') = \sqrt{\frac{GM}{r'}}$$

The ratio is

$$\frac{v'_\phi}{v_K(r')} = \sqrt{\frac{r}{r'}}$$

If $r' < r$, then $v'_\phi > v_K(r')$. The parcel is moving too fast for a circular orbit at its new radius. The centrifugal force exceeds gravity, and the parcel gets pushed back outward.

This is the angular momentum barrier. A parcel of gas that conserves its angular momentum cannot move inward to a circular orbit at a smaller radius. It will oscillate around its original radius, perhaps executing an elliptical orbit, but it cannot spiral inward.

[1] We could equally well work in an inertial frame and talk about the centripetal acceleration required to maintain circular motion. The physics is the same.

1.3 *A Paradox of Cosmic Proportions*

Here we face a genuine puzzle. Angular momentum is conserved. Material orbiting a star or black hole cannot simply fall in while keeping its angular momentum. Yet we observe accretion happening everywhere in the universe:

- Quasars—the luminous nuclei of distant galaxies—shine with the power of a trillion suns. This energy comes from matter falling onto supermassive black holes at rates of several solar masses per year.
- X-ray binaries—neutron stars and stellar-mass black holes orbiting normal stars—show variability on timescales of milliseconds, indicating that matter is crashing onto surfaces or disappearing through event horizons.
- T Tauri stars—young stars still gathering mass—are surrounded by dusty disks from which they continue to grow.

The numbers are striking. The supermassive black hole at the center of the galaxy M87 has a mass of about 6×10^9 solar masses. This mass had to come from somewhere. If it accumulated over the age of the universe (about 13.7 billion years), the average accretion rate was roughly half a solar mass per year. Every year, for billions of years, material equivalent to half our entire Sun had to lose its angular momentum and fall into the black hole.

You might say, "Perhaps the angular momentum was never very large to begin with. Maybe the gas started with very little rotation." But this doesn't work. Any realistic cloud of gas has some rotation—even a tiny initial rotation gets amplified enormously as the cloud contracts. A cloud that starts with a leisurely spin once per million years will, by the time it contracts to disk scales, be rotating furiously. The angular momentum problem cannot be avoided by initial conditions.

How? The answer, as we will see, involves a conspiracy between viscosity, magnetic fields, and geometry that allows angular momentum to be transported outward even as mass flows inward. But before we can understand the solution, we need to appreciate the full depth of the problem.

1.4 *The Minimum Radius*

Given a parcel of gas with specific angular momentum ℓ, there is a minimum radius it can reach while remaining in a bound orbit. At this radius, called the circularization radius, all of the parcel's kinetic

energy is in circular motion. If it tried to go closer, it would need more velocity than angular momentum conservation allows.

The circularization radius is found by setting the parcel's velocity equal to the Keplerian velocity:

$$\frac{\ell}{r_{\text{circ}}} = \sqrt{\frac{GM}{r_{\text{circ}}}}$$

Solving:

$$r_{\text{circ}} = \frac{\ell^2}{GM}$$

Let's work out what this means for our collapsing gas cloud. We found earlier that $\ell = \sqrt{GMr}$ for material starting in a Keplerian orbit at radius r. For material starting far from the central object with some small angular velocity, the specific angular momentum is roughly $\ell \sim \omega_0 r_0^2$ where r_0 is the initial radius and ω_0 is the initial angular velocity. The circularization radius is then

$$r_{\text{circ}} = \frac{\omega_0^2 r_0^4}{GM}$$

For our earlier example of a cloud with $r_0 = 10^{15}$ m, $\omega_0 = 2 \times 10^{-14}$ rad/s, and $M = 10^{30}$ kg collapsing toward a central mass $M_* = 2 \times 10^{30}$ kg:

$$r_{\text{circ}} = \frac{(2 \times 10^{-14})^2 (10^{15})^4}{(6.67 \times 10^{-11})(2 \times 10^{30})} \approx 3 \times 10^{12} \text{ m}$$

This is about 20 astronomical units—roughly the radius of Uranus's orbit around the Sun. Our slowly rotating gas cloud cannot collapse any further than this without somehow getting rid of angular momentum.

Notice that the circularization radius depends on the fourth power of the initial radius and the second power of the initial angular velocity. Even small amounts of initial rotation lead to substantial circularization radii. This is why disks are ubiquitous in astrophysics—any collapsing cloud with the slightest rotation will form a disk rather than falling directly onto the central object.

1.5 Why a Disk?

We've established that infalling material cannot reach the central object while conserving angular momentum. But why should it form a disk—a thin, flattened structure in a single plane—rather than some other shape?

Consider a gas cloud with angular momentum vector $\vec{L}$ pointing in some direction, which we'll call "up." The cloud has material at all

different angles relative to this axis. Some gas is near the equatorial plane (perpendicular to $\vec{L}$), while other gas is near the poles (along $\vec{L}$).

As the cloud collapses, material near the poles has a problem. It's trying to move inward, but "inward" for polar material means toward the rotation axis. This material has angular momentum—it's orbiting around the axis with the same angular velocity as the rest of the cloud. When it tries to move toward the axis, it experiences the same angular momentum barrier we discussed, but now the barrier prevents motion toward the axis itself, not toward the central object.

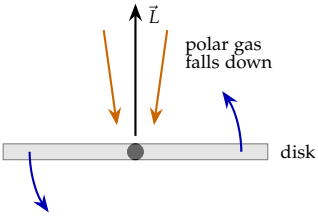

Figure 1.1: Polar material falls toward the equatorial plane without changing its distance from the rotation axis, forming a flattened disk.

But here's the key insight: polar material can fall toward the equatorial plane without changing its distance from the rotation axis. As long as it stays at the same cylindrical radius, it keeps the same angular momentum about the axis. So gas near the poles collapses down toward the equator, while gas near the equator stays more or less in place.

The result is a disk. Material from all directions settles into the equatorial plane, where it orbits the central object. The disk can be quite thin—in some systems, the ratio of thickness to radius is only a few percent.

This explains the ubiquity of disks in astrophysics. They're not a special configuration that requires fine-tuning. They're the natural outcome of gravitational collapse when angular momentum is present. Any collapsing cloud with net rotation will form a disk.

1.6 The Central Problem

Now we can state the central problem of accretion physics clearly. Gas settling into a disk reaches the circularization radius and stops. To get any closer to the central object, it must lose angular momentum. But angular momentum is conserved.

Where does the angular momentum go?

There are only a few possibilities:

1. **It leaves the system entirely.** Angular momentum could be carried away by outflows, jets, or radiation. Some of it certainly is—we'll discuss jets in later chapters. But this can't be the whole story, because we observe steady accretion in systems with no obvious outflows.

2. **It goes to the central object.** The accreted material could spin up the star or black hole it falls onto. This does happen, but it's limited. A black hole, for instance, has a maximum spin rate. Once it reaches this rate, it can't accept any more angular momentum per unit mass than the material already carries. And we've already

seen that material carries far too much angular momentum to fall in directly.

3. **It's transported outward within the disk.** If some of the disk material moves outward, carrying its angular momentum with it, then other material can move inward. This is like two ice skaters spinning together—if one pushes the other outward while staying connected, the inner skater speeds up while the outer one slows down. The total angular momentum is conserved, but it's been rearranged.

The third option turns out to be the key. In an accretion disk, angular momentum flows outward while mass flows inward. The material that carries angular momentum outward eventually reaches the edge of the disk and escapes (or gets expelled), taking the excess angular momentum with it. The material that flows inward loses angular momentum to its outer neighbors and eventually falls onto the central object.

But this raises a new question: what mechanism transports angular momentum outward? In ordinary fluids, this would be viscosity—the friction between fluid layers moving at different speeds. But as we'll see in Chapter 4, ordinary viscosity is far too weak to explain observed accretion rates. Something else must be responsible.

The recognition of this problem has a long history. As early as the 1940s, Carl Friedrich von WeizsÃd'cker and others realized that the solar nebula—the disk of gas and dust from which the planets formed—faced exactly this angular momentum crisis. The Sun contains 99.8% of the solar system's mass but only about 2% of its angular momentum. Somehow, angular momentum was transported outward to the planets while mass fell inward to form the Sun. By the 1970s, with the discovery of X-ray binaries and quasars, the problem had become urgent. These objects required accretion rates that no known mechanism could explain. The solution, when it came in 1991, would revolutionize the field.

1.7 The Plan Ahead

In the chapters that follow, we'll develop the physics of accretion disks and jets systematically:

In Chapter 2, we'll look more carefully at angular momentum—how it's defined, how it's conserved, and what happens when material with different specific angular momenta mixes together.

In Chapter 3, we'll develop the theory of thin disks—the most common configuration in astrophysics. We'll derive the temperature and density structure of a thin disk and see how bright it should be.

In Chapter 4, we'll confront the mystery of viscosity. What could possibly transport angular momentum efficiently enough to explain observed accretion rates? The answer remained elusive for decades.

In Chapter 5, we'll learn what accretion disks look like—their spectra, their variability, their appearance in images. This is how we connect theory to observation.

In Chapter 6, we'll see what happens when the assumptions of the thin disk model break down. Very high and very low accretion rates lead to qualitatively different structures.

In Chapters 7 and 8, we'll explore the role of magnetic fields. It turns out that magnetic fields are not just a small correction to disk physics—they're essential for angular momentum transport and are probably responsible for launching jets.

In Chapter 9, we'll follow jets as they travel outward from their sources, examining how they're accelerated and collimated into the narrow beams we observe.

Finally, in Chapter 10, we'll survey the remarkable universality of disk-jet systems across the mass scale, from forming planets to supermassive black holes.

The story of accretion disks and jets is one of the great achievements of theoretical astrophysics. It connects the rotation of a gas cloud to the brightest objects in the universe. And it all begins with a simple observation: matter doesn't fall straight.

2
Angular Momentum: The Cosmic Traffic Jam

What happens when a river of gas, all trying to flow in the same direction, has to squeeze through a bottleneck? In ordinary life, we call this a traffic jam. Cars pile up, tempers flare, and everyone inches forward at a crawl. In the cosmos, something similar happens when gas tries to fall onto a compact object—except the bottleneck isn't a lane closure but a fundamental conservation law.

In the previous chapter, we saw that angular momentum prevents gas from falling directly onto a central object. The gas gets stuck at the circularization radius, where it has just enough angular velocity to maintain a circular orbit. To accrete, the gas must somehow transfer its angular momentum elsewhere.

But we glossed over an important question: what exactly is angular momentum in a continuous fluid, and how does it get redistributed? A gas cloud isn't a figure skater with a well-defined moment of inertia. It's a collection of particles, each following its own trajectory, each carrying its own angular momentum. When we talk about angular momentum being "transported," what do we mean?

This chapter develops the machinery we'll need to understand angular momentum in accretion flows. The key quantity turns out to be the specific angular momentum—angular momentum per unit mass—which is conserved along the trajectory of each fluid element in the absence of torques. Understanding how specific angular momentum varies with position in a disk, and how it flows from one place to another, is essential for understanding accretion.

2.1 Specific Angular Momentum

For a single particle of mass m at position $\vec{r}$ moving with velocity $\vec{v}$, the angular momentum about the origin is

$$\vec{L} = m\vec{r} \times \vec{v}$$

The specific angular momentum is simply

$$\vec{\ell} = \vec{r} \times \vec{v}$$

For a particle in the equatorial plane of a disk, moving in a circular orbit at radius r with velocity v_ϕ in the azimuthal direction, the specific angular momentum has magnitude

$$\ell = r v_\phi$$

and points perpendicular to the plane (in the z direction).

Why focus on specific angular momentum rather than total angular momentum? Because in a fluid, mass is constantly being redistributed. A parcel of gas might gain or lose mass through compression, mixing, or flow across boundaries. But in the absence of torques, the specific angular momentum of each fluid element is conserved. It's the natural quantity to track.

For Keplerian orbits, where $v_\phi = \sqrt{GM/r}$, the specific angular momentum is

$$\ell_K = r\sqrt{\frac{GM}{r}} = \sqrt{GMr}$$

This increases with radius. Outer parts of a Keplerian disk have more specific angular momentum than inner parts. This simple fact—that $\ell_K \propto r^{1/2}$—has profound consequences.

2.2 *The Traffic Jam Analogy*

Imagine a highway with multiple lanes merging into a single lane. Cars in the outer lanes are traveling at highway speeds. To merge into the inner lane, they must slow down and wait for a gap. The result is a traffic jam—cars pile up in the merge zone, unable to move forward at the rate they'd like.

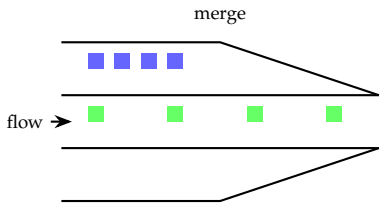

Figure 2.1: Cars merging from outer lanes pile up at the merge point, just as gas with high angular momentum piles up as it tries to move inward in a disk.

Now think of the outer lanes as representing large radii in a disk, with high specific angular momentum. The inner lane represents small radii, with low specific angular momentum. Gas trying to move inward faces a merge—it has too much angular momentum for the inner lanes. If it can't shed that angular momentum somehow, it backs up, forming a disk.

But here's where the analogy gets interesting. On a highway, cars can merge when there's a gap. The congestion propagates backward from the merge point. In a disk, there's no shortage of space at small radii—the central object would happily accept more mass. The bottleneck is purely due to angular momentum. Gas at large radii has to give its angular momentum to... what?

The answer is: to other gas at even larger radii. If gas at radius r can transfer angular momentum to gas at radius $r + \Delta r$, then the

inner gas can move inward while the outer gas moves outward. Mass flows inward; angular momentum flows outward. The disk acts as a conveyor belt, with angular momentum being passed from hand to hand outward until it reaches the edge, where it's expelled from the system.

2.3 *The Flux of Angular Momentum*

Let's make this idea quantitative. Consider a ring of gas in the disk, extending from radius r to $r + dr$. The ring has surface density Σ (mass per unit area), so its mass is $dm = 2\pi r \Sigma dr$. Its angular momentum is

$$dL = dm \cdot \ell = 2\pi r \Sigma \ell \, dr$$

where $\ell = \ell(r)$ is the specific angular momentum at radius r.

Now consider what happens to this ring over time. Mass can flow into the ring from the outside (at $r + dr$) and out of the ring toward the inside (at r). Let $\dot{M}(r)$ be the mass flow rate through radius r—positive means inward flow. The mass flux into the ring is

$$\frac{d(dm)}{dt} = \dot{M}(r + dr) - \dot{M}(r) = \frac{\partial \dot{M}}{\partial r} dr$$

But angular momentum can also flow through the ring without being carried by mass. Neighboring rings exert torques on each other—through viscosity, magnetic stresses, or gravitational forces. Let $\dot{L}(r)$ be the torque exerted by the material inside radius r on the material outside radius r. Positive torque means angular momentum is being transferred outward.[1]

[1] This sign convention is standard but can be confusing. Think of it this way: if the inner material speeds up the outer material, angular momentum has flowed outward, so $\dot{L} > 0$.

The angular momentum of the ring changes due to both the mass flux (carrying its own angular momentum) and the torque:

$$\frac{d(dL)}{dt} = \ell(r + dr)\dot{M}(r + dr) - \ell(r)\dot{M}(r) + \dot{L}(r) - \dot{L}(r + dr)$$

The first two terms account for angular momentum carried in and out by the mass flow. The second two terms account for angular momentum transferred by torques.

Expanding to first order in dr:

$$\frac{\partial}{\partial t}(2\pi r \Sigma \ell) = \frac{\partial(\ell \dot{M})}{\partial r} - \frac{\partial \dot{L}}{\partial r}$$

In a steady state, $\partial/\partial t = 0$, so

$$\frac{\partial(\ell \dot{M})}{\partial r} = \frac{\partial \dot{L}}{\partial r}$$

Integrating from the inner edge $r_{\rm in}$ to some radius r:

$$\ell(r)\dot{M}(r) - \ell_{\rm in}\dot{M}_{\rm in} = \dot{L}(r) - \dot{L}_{\rm in}$$

If we assume that the torque vanishes at the inner edge (a reasonable approximation for many systems), then

$$\ell(r)\dot{M}(r) = \dot{L}(r) + \ell_{\rm in}\dot{M}_{\rm in}$$

This is a fundamental relation for accretion disks. The angular momentum flux $\dot{L}$ at each radius is determined by the mass flux and the specific angular momenta at that radius and at the inner edge. For accretion to occur ($\dot{M} > 0$ inward), angular momentum must flow outward ($\dot{L} > 0$).

2.4 *A Keplerian Disk in Steady State*

Let's work out what this means for a Keplerian disk. In a Keplerian disk, $\ell = \sqrt{GMr}$, so

$$\frac{d\ell}{dr} = \frac{1}{2}\sqrt{\frac{GM}{r}}$$

For a steady accretion rate $\dot{M}$ (constant with radius in steady state), the angular momentum equation becomes

$$\dot{M}\sqrt{GMr} = \dot{L}(r) + \dot{M}\sqrt{GMr_{\rm in}}$$

Solving for the torque:

$$\dot{L}(r) = \dot{M}\sqrt{GM}\left(\sqrt{r} - \sqrt{r_{\rm in}}\right)$$

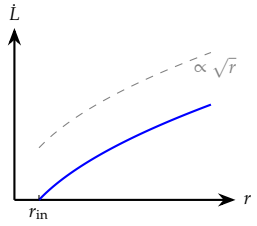

Figure 2.2: Angular momentum flux $\dot{L}$ as a function of radius in a steady Keplerian disk. The flux grows as $\sqrt{r}$ at large radii.

Several features are worth noting. First, the angular momentum flux increases with radius. More angular momentum flows through the outer parts of the disk than through the inner parts. This makes sense: material at small radii has already shed most of its angular momentum on the way in.

Second, the angular momentum flux goes to zero at $r = r_{\rm in}$. This is our boundary condition—we assumed no torque at the inner edge. In reality, the inner boundary condition depends on what's there. For a black hole, there really is no torque once material crosses the innermost stable circular orbit. For a star with a magnetosphere, the star exerts a torque on the inner disk.

Third, and most importantly, for the torque to be positive (angular momentum flowing outward), something must exert it. In a simple gas, that something is viscosity.

2.5 *What Viscosity Does*

Viscosity in a fluid arises because adjacent fluid layers moving at different speeds exert shear stresses on each other. Faster-moving layers drag slower-moving layers along; slower-moving layers hold

back faster-moving layers. The effect is to transfer momentum from fast regions to slow regions.

In an accretion disk, the orbital velocity decreases outward (for Keplerian rotation, $v_\phi \propto r^{-1/2}$). Inner parts of the disk orbit faster than outer parts. Viscosity causes the fast inner material to drag on the slow outer material, speeding it up. But remember, orbital velocity and angular momentum are different things. Speeding up a piece of outer material increases its angular momentum. So viscosity transfers angular momentum outward even though it transfers linear momentum inward.

To see this mathematically, consider two adjacent rings of gas at radii r and $r + dr$. The inner ring moves at velocity $v_\phi(r)$; the outer ring at $v_\phi(r + dr)$. The shear between them is

$$\frac{dv_\phi}{dr} = \frac{d}{dr}(r\Omega) = \Omega + r\frac{d\Omega}{dr}$$

For Keplerian rotation, $\Omega = \sqrt{GM/r^3}$, so

$$\frac{d\Omega}{dr} = -\frac{3}{2}\sqrt{\frac{GM}{r^5}}$$

and

$$\frac{dv_\phi}{dr} = \sqrt{\frac{GM}{r^3}} - \frac{3}{2}\sqrt{\frac{GM}{r^3}} = -\frac{1}{2}\sqrt{\frac{GM}{r^3}}$$

The shear is negative because v_ϕ decreases outward.

The viscous stress (force per unit area) between the rings is proportional to the shear:

$$\tau = \eta\frac{dv_\phi}{dr}$$

where η is the dynamic viscosity. But what matters for angular momentum transport is the torque per unit area, which is the stress times the lever arm r:

$$\text{torque per unit area} = r\tau = \eta r\frac{dv_\phi}{dr}$$

For a disk of thickness H, the total torque exerted by the inner disk on the outer disk across a cylindrical surface at radius r is

$$\dot{L} = 2\pi r \cdot H \cdot r\tau = 2\pi r^2 H\eta\frac{dv_\phi}{dr}$$

It's conventional to define the kinematic viscosity $\nu = \eta/\rho$, where ρ is the density. Then, since $\Sigma = \rho H$:

$$\dot{L} = 2\pi r^2\Sigma\nu\frac{dv_\phi}{dr}$$

For Keplerian rotation:

$$\dot{L} = 2\pi r^2\Sigma\nu \cdot \left(-\frac{1}{2}\sqrt{\frac{GM}{r^3}}\right) = -\pi\Sigma\nu\sqrt{\frac{GM}{r}}$$

Wait—this is negative! Did we make a mistake? No. The negative sign reflects our sign convention. Positive $\dot{L}$ means angular momentum flowing outward, which means the inner disk is pushing the outer disk to go faster. But viscosity makes the inner disk push the outer disk in the direction of its own motion (trying to speed it up to match), which is the prograde direction. In a Keplerian disk where inner orbits are faster, this means speeding up the outer material, increasing its angular momentum. The physical torque is indeed outward.

The resolution is that we should compare with the angular momentum gradient, not just take the sign literally. The viscous torque acts to flatten the angular momentum profile—moving angular momentum from where there's more to where there's less. In a Keplerian disk, there's more angular momentum at large r, but the material is going slower. Viscosity transfers momentum (linear) inward but angular momentum outward.

2.6 *The Viscous Timescale*

How long does it take for viscosity to transport angular momentum outward and allow accretion to proceed? We can estimate this with a dimensional argument.

The viscous stress transports momentum over a distance r on a timescale

$$t_\nu \sim \frac{r^2}{\nu}$$

This is the viscous timescale. For accretion to proceed on this timescale, we need

$$\dot{M} \sim \frac{M_{\text{disk}}}{t_\nu} \sim \frac{\Sigma r^2}{r^2/\nu} = \Sigma\nu$$

The accretion rate is proportional to the product of surface density and kinematic viscosity. This is dimensionally correct: $[\Sigma] = \text{kg/m}^2$ and $[\nu] = \text{m}^2/\text{s}$, so $[\Sigma\nu] = \text{kg/s/m}$. Multiplying by a characteristic length gives kg/s, a mass rate.

Now let's put in numbers for a real accretion disk. Consider an accretion disk around a stellar-mass black hole, with

- Black hole mass: $M = 10M_\odot = 2 \times 10^{31}$ kg
- Disk radius: $r = 10^{10}$ m (about 0.1 AU)
- Accretion rate: $\dot{M} = 10^{-8}M_\odot/\text{yr} = 6 \times 10^{14}$ kg/s

The Keplerian orbital velocity at this radius is

$$v_K = \sqrt{\frac{GM}{r}} = \sqrt{\frac{(6.67 \times 10^{-11})(2 \times 10^{31})}{10^{10}}} \approx 3.7 \times 10^5 \text{ m/s}$$

The orbital period is

$$P = \frac{2\pi r}{v_K} = \frac{2\pi(10^{10})}{3.7 \times 10^5} \approx 1.7 \times 10^5 \text{ s} \approx 2 \text{ days}$$

If the disk has surface density $\Sigma \sim 10^4$ kg/m^2 (a reasonable estimate), then for accretion at the observed rate:

$$\nu \sim \frac{\dot{M}}{\Sigma} \sim \frac{6 \times 10^{14}}{10^4} = 6 \times 10^{10} \text{ m}^2/\text{s}$$

The viscous timescale is then

$$t_\nu \sim \frac{r^2}{\nu} = \frac{(10^{10})^2}{6 \times 10^{10}} \approx 1.7 \times 10^9 \text{ s} \approx 50 \text{ years}$$

So the disk evolves viscously on a timescale of years—much longer than the orbital period of days, but much shorter than the age of the system. This separation of timescales is typical and important: the disk maintains approximate Keplerian rotation (set by gravity, acting on an orbital timescale) while slowly evolving through viscous processes.

2.7 *The Mystery Deepens*

Now comes the puzzle. What is this viscosity $\nu \sim 10^{10}$ m^2/s? Is it the ordinary molecular viscosity of gas?

Molecular viscosity arises from random thermal motions of gas particles. A particle in a faster-moving layer occasionally wanders into a slower-moving layer and deposits its excess momentum there. The kinematic viscosity is approximately

$$\nu_{\text{mol}} \sim v_{\text{th}} \lambda$$

where v_{th} is the thermal velocity and λ is the mean free path between collisions.

For ionized gas at temperature $T \sim 10^6$ K (typical for the inner parts of an accretion disk around a black hole), the thermal velocity is

$$v_{\text{th}} \sim \sqrt{\frac{k_B T}{m_p}} \sim \sqrt{\frac{(1.4 \times 10^{-23})(10^6)}{1.7 \times 10^{-27}}} \sim 10^5 \text{ m/s}$$

The mean free path in a plasma is limited by Coulomb collisions. For electron density $n_e \sim 10^{20}$ m^{-3} and temperature $T \sim 10^6$ K:

$$\lambda \sim \frac{T^2}{n_e} \sim \frac{(10^6)^2}{10^{20}} \times \text{constants} \sim 10^{-2} \text{ m}$$

(The actual calculation involves the Coulomb logarithm and various factors, but the order of magnitude is correct.)

So the molecular viscosity is

$$\nu_{\text{mol}} \sim (10^5)(10^{-2}) \sim 10^3 \text{ m}^2/\text{s}$$

Compare this to the required viscosity: $\nu \sim 10^{10}$ m^2/s. Molecular viscosity is too small by a factor of 10^7!

You might say, "Perhaps we made an error somewhere. A factor of ten million is a lot." But we didn't. The calculation is solid, and the discrepancy is real. Molecular viscosity gives us an answer that's wrong by seven orders of magnitude. In physics, we call this "not even close."

This was the central puzzle of accretion disk theory for decades. We observe accretion happening. We can calculate the viscosity required to explain it. But the actual microscopic viscosity of the gas is far too small. Something else must be transporting angular momentum.

2.8 *The Alpha Prescription*

Faced with this puzzle, Shakura and Sunyaev in 1973 made a practical choice. They didn't know what was causing the anomalously high viscosity, but they wanted to build models that could be compared with observations. So they parameterized their ignorance.

They proposed that whatever the viscosity mechanism, it probably scales with the local properties of the disk. The largest velocity in the disk is the orbital velocity v_K. The largest length scale is the disk thickness H. Dimensional analysis suggests

$$\nu \sim vH$$

where v is some characteristic velocity and H is some characteristic length. But in a Keplerian disk, the sound speed c_s is related to the orbital velocity by the disk's aspect ratio: $c_s \sim (H/r)v_K$. So

$$\nu \sim \alpha c_s H$$

where α is a dimensionless parameter that encodes our ignorance.

This is the famous alpha prescription. The parameter α could be anywhere from 0 (no viscosity) to 1 (maximum conceivable viscosity—motions approaching the sound speed over a scale comparable to the disk thickness). Fits to observations of various accretion disks suggest $\alpha \sim 0.01$ to 0.1.

The alpha prescription is not a theory. It's a way of parameterizing what we don't know so that we can make progress. It says: whatever the viscosity mechanism, it probably doesn't exceed the "turbulent" limit where motions at the sound speed transport momentum over

the disk thickness. And it probably doesn't depend sensitively on details we haven't thought of.

For decades, theorists built increasingly sophisticated disk models using the alpha prescription, hoping that eventually someone would figure out what α really represented. The answer, when it came, was magnetic.

2.9 *Why This Matters*

Understanding angular momentum transport might seem like an abstract exercise, but it has concrete observational consequences. The rate at which angular momentum flows through the disk determines the accretion rate. The accretion rate determines the luminosity—how brightly the disk shines. Different values of α predict different relationships between disk mass, size, and luminosity.

More subtly, the nature of the angular momentum transport affects the disk's variability. If angular momentum is transported by large-scale magnetic fields, the disk might show quasi-periodic oscillations as magnetic structures form and dissipate. If it's transported by small-scale turbulence, the variability might be more random. Observations of flickering and oscillations in X-ray binaries provide clues about the underlying transport mechanism.

And understanding angular momentum in disks is essential for understanding jets. Jets remove angular momentum from the system extremely efficiently. A jet that carries off material with high specific angular momentum can allow much more rapid accretion than viscosity alone. The interplay between viscous accretion through the disk and magnetically driven outflows in jets is one of the key questions in modern accretion physics.

We'll return to these themes in later chapters. But first, in Chapter 3, we'll develop the standard thin disk model and see what it predicts for the structure and appearance of an accretion disk.

3
The Thin Disk

Here is something that should astonish us, if we stop to think about it. In the chaos of gravitational collapse, in the violent environment near black holes and young stars, nature builds structures of extraordinary geometric precision. Disks thinner than knife blades, stretched across distances that dwarf our solar system. The disk around the supermassive black hole in NGC 4258, mapped by water masers, has a thickness less than 1% of its radius—flatter, proportionally, than a DVD. The protoplanetary disk around the young star HL Tau shows concentric rings so flat they cast sharp shadows. X-ray observations of accreting black holes reveal material heated to millions of degrees, yet still confined to a thin equatorial plane.

Why should this be? What principle of physics enforces such exquisite flatness? You might expect that all the violence near a black hole—the tremendous shearing, the magnetic turbulence, the radiation pressure—would puff the gas up into a chaotic blob. But nature has other ideas. In this chapter, we'll develop the standard model of a thin accretion disk—the Shakura-Sunyaev model that has guided the field for half a century. The thin disk approximation allows us to solve for the disk's temperature, density, and luminosity as functions of radius, all in terms of the mass of the central object, the accretion rate, and the mysterious viscosity parameter.

3.1 Why Disks Are Thin

A disk in the equatorial plane of a rotating system is held against gravitational collapse in the vertical direction by pressure. In the radial direction, gravity is balanced by centrifugal force (or, in an inertial frame, the orbital motion provides the centripetal acceleration needed to curve the trajectory). But in the vertical direction, there's no rotation to provide support. The disk must support itself thermally.

Consider a parcel of gas at height z above the midplane, at cylin-

drical radius r from the central object. The gravitational acceleration toward the midplane is approximately

$$g_z = \frac{GMz}{(r^2 + z^2)^{3/2}}$$

For $z \ll r$, this simplifies to

$$g_z \approx \frac{GMz}{r^3} = \Omega_K^2 z$$

where $\Omega_K = \sqrt{GM/r^3}$ is the Keplerian angular velocity.

The disk is in hydrostatic equilibrium in the vertical direction:

$$\frac{dP}{dz} = -\rho g_z = -\rho \Omega_K^2 z$$

For an isothermal disk with sound speed $c_s = \sqrt{P/\rho}$ (assuming ideal gas), we can write $P = \rho c_s^2$. Then

$$c_s^2 \frac{d\rho}{dz} = -\rho \Omega_K^2 z$$

This has the solution

$$\rho(z) = \rho_0 \exp\left(-\frac{z^2}{2H^2}\right)$$

where

$$H = \frac{c_s}{\Omega_K}$$

is the disk scale height.

The ratio of scale height to radius is

$$\frac{H}{r} = \frac{c_s}{\Omega_K r} = \frac{c_s}{v_K}$$

where $v_K = r\Omega_K$ is the Keplerian orbital velocity.

This is a key result: the geometric thickness of the disk is determined by the ratio of sound speed to orbital velocity. In most astrophysical situations, the sound speed is much smaller than the orbital velocity. Gas temperatures are typically 10^4 to 10^7 K, giving $c_s \sim 10$ to 300 km/s. But orbital velocities near compact objects are a substantial fraction of the speed of light: $v_K \sim 0.1c$ to $0.5c$ near a black hole. Even around a normal star, $v_K \sim 30$ km/s at 1 AU.

So $H/r \sim c_s/v_K$ is typically 0.01 to 0.1. Disks are geometrically thin because they're cold—cold compared to what it would take for pressure to provide support comparable to the rotational support in the radial direction.

3.2 *The Thin Disk Equations*

In a thin disk, we can simplify the three-dimensional fluid equations by averaging over the vertical direction. This is analogous to how we describe the ocean with surface waves rather than tracking the full three-dimensional flow.

Define the surface density as the vertical integral of the volume density:

$$\Sigma = \int_{-\infty}^{\infty} \rho \, dz$$

For the Gaussian density profile $\rho = \rho_0 e^{-z^2/2H^2}$, this gives

$$\Sigma = \sqrt{2\pi}\rho_0 H$$

The equations governing a thin disk are:

Mass conservation:

$$\frac{\partial \Sigma}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}(r\Sigma v_r) = 0$$

where v_r is the (vertically averaged) radial velocity.

Angular momentum conservation:

$$\frac{\partial}{\partial t}(\Sigma r^2 \Omega) + \frac{1}{r}\frac{\partial}{\partial r}(r\Sigma v_r \cdot r^2\Omega) = \frac{1}{r}\frac{\partial}{\partial r}(r \cdot \text{torque})$$

For a Keplerian disk with the alpha viscosity prescription, the torque per unit circumference is $\nu\Sigma r^2 d\Omega/dr$. After some algebra (which we'll skip—it's straightforward but tedious), the mass conservation and angular momentum equations combine into a single diffusion equation for the surface density:

$$\frac{\partial \Sigma}{\partial t} = \frac{3}{r}\frac{\partial}{\partial r}\left[r^{1/2}\frac{\partial}{\partial r}(\nu\Sigma r^{1/2})\right]$$

This is remarkably elegant. The surface density evolves by diffusion, but in a peculiar geometry dictated by the $r^{1/2}$ factors from Keplerian rotation. Matter diffuses inward (toward the central object) while angular momentum diffuses outward (toward the disk edge).

In steady state, $\partial\Sigma/\partial t = 0$, and the equation can be integrated to give

$$\nu\Sigma = \frac{\dot{M}}{3\pi}\left(1 - \sqrt{\frac{r_{\text{in}}}{r}}\right)$$

where $\dot{M}$ is the mass accretion rate and r_{in} is the inner edge of the disk.

3.3 *Energy Generation*

As gas spirals inward through the disk, it loses gravitational potential energy. Where does this energy go?

Consider a ring of gas at radius r with mass dm. Its gravitational potential energy is

$$U = -\frac{GM\,dm}{r}$$

As the ring moves inward by dr, the potential energy changes by

$$dU = \frac{GM\,dm}{r^2}dr$$

But the gas also has kinetic energy from its orbital motion. For a Keplerian orbit, the kinetic energy is exactly half the magnitude of the potential energy (by the virial theorem):

$$K = \frac{1}{2}dm \cdot v_K^2 = \frac{1}{2}dm \cdot \frac{GM}{r} = -\frac{1}{2}U$$

The total mechanical energy is

$$E = K + U = -\frac{1}{2}\frac{GM\,dm}{r}$$

As the gas moves inward from r to $r - dr$, its mechanical energy changes by

$$dE = \frac{1}{2}\frac{GM\,dm}{r^2}dr$$

This is positive—the gas gains mechanical energy as it falls inward. But wait, shouldn't it lose energy to spiral in?

You might object: "If the gas is gaining energy, how can we call this accretion? Shouldn't it fly apart?" This is exactly the right question. The resolution is that half the liberated gravitational potential energy goes into kinetic energy (the gas speeds up as it moves to smaller radii), while the other half is dissipated by viscosity and radiated away. The gas does gain kinetic energy—it orbits faster at smaller radii—but it loses even more energy to radiation.

The energy dissipation rate per unit area of the disk (summing both faces) is

$$D(r) = \frac{3GM\dot{M}}{4\pi r^3}\left(1 - \sqrt{\frac{r_{\rm in}}{r}}\right)$$

Notice that $D(r) \to 0$ as $r \to r_{\rm in}$. This is because of our boundary condition: zero torque at the inner edge. With no torque, there's no shear stress, and no viscous dissipation. In reality, the boundary condition is more complicated—material plunging into a black hole, or magnetic coupling to a stellar surface—but the standard model takes the simplest assumption.

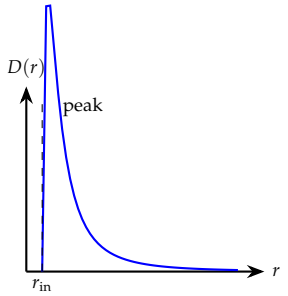

Figure 3.1: Energy dissipation rate per unit area as a function of radius. The dissipation peaks just outside the inner edge, not at $r_{\rm in}$ itself due to the boundary condition.

Integrating the dissipation rate over the whole disk gives the total luminosity:

$$L = \int_{r_{\rm in}}^{\infty} 2\pi r\, D(r)\, dr = \frac{GM\dot{M}}{2r_{\rm in}}$$

This beautiful result says that the total luminosity is just half the gravitational potential energy released as material falls from infinity to $r_{\rm in}$. The other half remains as kinetic energy of the gas at the inner edge, and is either carried into the black hole or deposited on the stellar surface.

3.4 *Temperature Structure*

If the disk radiates as a blackbody (a reasonable approximation when the disk is optically thick), then the local effective temperature is determined by the local dissipation rate:

$$\sigma T_{\rm eff}^4 = D(r)$$

where σ is the Stefan-Boltzmann constant. Solving:

$$T_{\rm eff}(r) = \left[\frac{3GM\dot{M}}{8\pi\sigma r^3}\left(1-\sqrt{\frac{r_{\rm in}}{r}}\right)\right]^{1/4}$$

Let's put in numbers for a disk around a stellar-mass black hole. Take $M = 10M_\odot$, $\dot{M} = 10^{-8}M_\odot$/yr, and $r_{\rm in} = 6GM/c^2$ (the innermost stable circular orbit for a non-spinning black hole). Then $r_{\rm in} \approx 9 \times 10^4$ m.

At $r = 10r_{\rm in}$:

$$T_{\rm eff} = \left[\frac{3(6.67\times10^{-11})(2\times10^{31})(6\times10^{15})}{8\pi(5.67\times10^{-8})(9\times10^5)^3}(1-1/\sqrt{10})\right]^{1/4}$$

Working through the arithmetic: $T_{\rm eff} \approx 10^7$ K.

This is remarkably hot—ten million degrees! At those temperatures, the atoms themselves are having a bad day: most electrons have long since abandoned their nuclei. The disk emits primarily in X-rays. This is why stellar-mass black holes in accreting binary systems are called X-ray binaries. The disk is so hot that it glows in X-rays.

System	$T_{\rm eff}$ at $r = 10r_{\rm in}$	Emission
Stellar-mass BH ($10M_\odot$)	10^7 K	X-rays
Supermassive BH ($10^8M_\odot$)	10^5 K	UV
White dwarf ($1M_\odot$)	10^4 K	Optical/UV
T Tauri star ($1M_\odot$, large $r_{\rm in}$)	10^3 K	Infrared

Table 3.1: Characteristic disk temperatures for different accreting systems. The temperature depends on the compactness—how much gravitational energy is released per unit area.

The temperature varies with central mass because the inner radius scales with mass (for a black hole, $r_{\rm in} \propto M$), and the temperature depends on the energy released per unit area. For a given accretion rate in Eddington units, the disk temperature scales as $T \propto M^{-1/4}$. Bigger black holes have cooler disks.

3.5 The Temperature Profile

Far from the inner edge, where $r \gg r_{\text{in}}$, the temperature profile simplifies to

$$T_{\text{eff}} \propto r^{-3/4}$$

This is the characteristic temperature profile of a thin accretion disk. The disk is hottest near the center and cools off toward the edge. The $r^{-3/4}$ scaling comes from the r^{-3} dependence of the dissipation rate combined with the fourth-root in the Stefan-Boltzmann law.

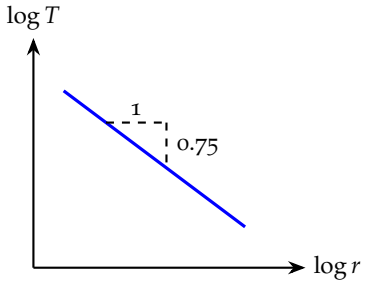

Figure 3.2: Temperature profile of a thin disk. On a log-log plot, $T \propto r^{-3/4}$.

This temperature profile has important observational consequences. Because different radii have different temperatures, they emit at different wavelengths. The inner disk (hot) emits at short wavelengths; the outer disk (cool) emits at long wavelengths. The spectrum of a disk is therefore a superposition of blackbodies at different temperatures—a "multi-color blackbody."

We can calculate this spectrum by integrating over the disk. The flux at frequency ν from a disk at distance d is

$$F_\nu = \frac{\cos i}{d^2} \int_{r_{\text{in}}}^{r_{\text{out}}} B_\nu(T(r)) \cdot 2\pi r \, dr$$

where i is the inclination angle (face-on is $i = 0$) and B_ν is the Planck function.

At high frequencies (short wavelengths), only the hottest inner regions contribute, and the spectrum follows the Rayleigh-Jeans tail of the hottest blackbody: $F_\nu \propto \nu^2$.

At intermediate frequencies, many annuli contribute, and the spectrum takes a characteristic form. The number of annuli contributing at frequency ν scales as the range of radii with $T \sim h\nu/k_B$. Since $T \propto r^{-3/4}$, this range scales as $r \propto T^{-4/3}$. The contribution from each annulus scales as $r \cdot B_\nu(T) \propto r \cdot T^3 \propto r \cdot r^{-9/4} \propto r^{-5/4}$. Integrating over the contributing annuli gives $F_\nu \propto \nu^{1/3}$.

At low frequencies, the entire disk contributes, including the cool outer regions that dominate the area. The spectrum approaches the Rayleigh-Jeans form again.

The predicted spectrum $F_\nu \propto \nu^{1/3}$ at intermediate frequencies is a distinctive signature of a thin accretion disk. Observations of many X-ray binaries and active galactic nuclei show this spectral shape, providing strong evidence for the thin disk model.

3.6 The Disk Structure Equations

So far we've focused on observable quantities: temperature and luminosity. But the thin disk model also predicts the internal structure

of the disk—its density, thickness, and opacity as functions of radius. These "structure equations" were first worked out by Shakura and Sunyaev in 1973.

The equations are:

Hydrostatic equilibrium:

$$H = \frac{c_s}{\Omega_K}$$

Energy balance: The energy generated by viscous dissipation equals the energy radiated from the surface:

$$\frac{3}{2}\nu\Sigma\Omega_K^2 = \sigma T_{\rm eff}^4$$

(This is a more direct form of the energy equation than our earlier expression for $D(r)$.)

Alpha viscosity:

$$\nu = \alpha c_s H$$

Equation of state: For an ideal gas,

$$c_s^2 = \frac{k_B T_c}{\mu m_p}$$

where T_c is the central (midplane) temperature and μ is the mean molecular weight.

Vertical energy transport: The relationship between midplane temperature T_c and surface temperature $T_{\rm eff}$ depends on the opacity. For an optically thick disk:

$$T_c^4 \approx \frac{3}{8}\tau T_{\rm eff}^4$$

where $\tau = \kappa\Sigma/2$ is the optical depth from midplane to surface, and κ is the opacity.

Opacity: The opacity depends on temperature and density. In different regions of the disk, different opacity sources dominate: electron scattering at high temperatures, free-free absorption at intermediate temperatures, and molecular/dust opacity at low temperatures.

These six equations involve six unknowns: H, Σ, ν, c_s, T_c, and τ (or equivalently κ). Given the central mass M, accretion rate $\dot{M}$, and viscosity parameter α, the equations can be solved at each radius to give the disk structure.

The solutions are often written as power laws in radius, accretion rate, and central mass. For example, in the inner, radiation-pressure-dominated region of a disk around a black hole (where electron scattering dominates the opacity):

$$H \propto \dot{M} \quad \text{(independent of } r\text{!)}$$
$$\Sigma \propto \dot{M}^{-1} r^{3/2}$$
$$T_c \propto r^{-3/4} M^{-1/4}$$

The fact that H is independent of radius in this regime is striking. It means the disk has constant thickness, flaring neither inward nor outward. This occurs because the radiation pressure (which supports the disk vertically) scales the same way with radius as the gravitational compression.

3.7 Stability

Thin disks can be unstable in several ways. The most important is thermal instability, which occurs when the heating rate increases faster than the cooling rate as temperature rises.

Consider a small perturbation that raises the temperature slightly. This increases the pressure, which puffs up the disk. If the viscous heating rate increases more than the radiative cooling rate, the disk heats up further, puffs up more, and runs away. Conversely, a small decrease in temperature leads to collapse.

The condition for thermal stability turns out to depend on which pressure source dominates. When gas pressure dominates ($P_{\rm gas} \gg P_{\rm rad}$), thin disks are thermally stable. When radiation pressure dominates ($P_{\rm rad} \gg P_{\rm gas}$), they're thermally unstable.

For disks around stellar-mass black holes accreting near the Eddington limit, the inner regions are radiation-pressure-dominated and therefore thermally unstable. This may explain some of the dramatic variability seen in X-ray binaries—the disk might oscillate between hot, thick states and cool, thin states on thermal timescales.

Another important instability is the viscous instability, which affects the surface density rather than the temperature. If the viscosity-surface density relationship has the wrong slope, perturbations in Σ can grow rather than decay. This leads to propagating waves of enhanced and depleted surface density—the disk breaks up into rings.

The detailed stability properties of thin disks are complex and depend on the opacity regime, the equation of state, and the nature of the viscosity. But the general point is that thin disks are not always the steady, smooth structures they appear in textbooks. Real disks can be dynamic, variable, and structured.

3.8 The Limits of the Thin Disk Model

The thin disk model makes several assumptions that can break down:

The disk must be geometrically thin: We assumed $H \ll r$. This requires $c_s \ll v_K$, which means the disk can't be too hot. At very high accretion rates, the disk becomes so luminous that radiation pressure puffs it up, and it's no longer thin.

The disk must be optically thick: We assumed the disk radiates

as a blackbody, which requires $\tau \gg 1$. At very low accretion rates, the density drops and the disk becomes optically thin. It can no longer cool efficiently, the temperature rises, and the gas puffs up into a thick, hot structure.

The disk must be steady: We derived steady-state solutions, but real disks can vary on many timescales. Outbursts in dwarf novae and X-ray transients show that disks can undergo dramatic changes on timescales of days to weeks.

The disk must be axisymmetric: We averaged over azimuthal angle. But disks can have spiral waves, warps, and other non-axisymmetric structures.

Relativistic effects are ignored: Near a black hole, general relativity modifies the physics significantly. The innermost stable circular orbit, the redshift of emitted radiation, and the dragging of inertial frames all affect the observable properties of the disk.

We'll explore some of these complications in Chapter 6. But first, we need to confront the deepest mystery of all: what is the viscosity, really?

4
Viscosity: The Mystery Force

For decades, accretion disk theory had a gaping hole at its center. The models worked. They predicted temperatures, spectra, and luminosities that matched observations. But they required a viscosity parameter α that was simply asserted, not derived. What physical mechanism could transport angular momentum efficiently enough to explain observed accretion rates? Molecular viscosity was far too weak. Some kind of turbulence seemed necessary, but what could drive turbulence in a Keplerian disk?

The answer, when it finally came in 1991, was magnetic. A weak magnetic field threading a disk is unstable to a runaway growth of turbulence—the magnetorotational instability, or MRI. This chapter tells the story of how the MRI was discovered, how it works, and why it revolutionized our understanding of accretion disks.

4.1 The Failure of Molecular Viscosity

Let us be precise about the failure of molecular viscosity. In a gas, viscosity arises because particles carry momentum from one place to another through their random thermal motions. A particle in a faster-moving layer wanders into a slower-moving layer, depositing its excess momentum. The kinematic viscosity is

$$\nu_{\rm mol} \sim v_{\rm th} \lambda$$

where $v_{\rm th}$ is the thermal velocity and λ is the mean free path.

In an ionized plasma, the mean free path is set by Coulomb collisions between ions. The Spitzer formula gives

$$\lambda \sim \frac{(k_B T)^2}{n e^4 \ln \Lambda}$$

where T is temperature, n is number density, and $\ln \Lambda \approx 20$ is the Coulomb logarithm.

For conditions typical of the inner regions of an accretion disk around a stellar-mass black hole ($T \sim 10^7$ K, $n \sim 10^{20}$ m^{-3}), the full calculation gives

$$\lambda \sim 0.5 \text{ m}$$

The thermal velocity is

$$v_{\text{th}} \sim \sqrt{\frac{k_B T}{m_p}} \sim \sqrt{\frac{(1.4 \times 10^{-23})(10^7)}{1.7 \times 10^{-27}}} \sim 3 \times 10^5 \text{ m/s}$$

So

$$\nu_{\text{mol}} \sim (3 \times 10^5)(0.5) \sim 10^5 \text{ m}^2/\text{s}$$

Compare this to the effective viscosity implied by the alpha prescription. The alpha viscosity is

$$\nu_\alpha = \alpha c_s H$$

With $\alpha \sim 0.01$, $c_s \sim v_{\text{th}} \sim 3 \times 10^5$ m/s, and $H \sim 0.1r \sim 10^8$ m (for $r \sim 10^9$ m, typical of the inner disk):

$$\nu_\alpha \sim (0.01)(3 \times 10^5)(10^8) \sim 3 \times 10^{11} \text{ m}^2/\text{s}$$

The required viscosity exceeds molecular viscosity by a factor of 10^6!

This isn't a small discrepancy that might be fixed by more careful calculations. It's six orders of magnitude. Something fundamentally different must be happening.

You might say, "Perhaps the temperature or density estimates are wrong." But even if you adjust them by factors of ten, you can't close a million-fold gap. The molecular viscosity calculation is robust. The problem isn't in the numbers; it's in the physics.

4.2 *Turbulence as Effective Viscosity*

The natural candidate is turbulence. In turbulent flow, large-scale eddies transfer momentum much more efficiently than molecular collisions do. The effective viscosity of turbulent flow is roughly

$$\nu_{\text{turb}} \sim v_{\text{eddy}} \cdot l_{\text{eddy}}$$

where v_{eddy} is the velocity of the turbulent eddies and l_{eddy} is their size.

If the eddies have velocities comparable to the sound speed and sizes comparable to the disk thickness, we get

$$\nu_{\text{turb}} \sim c_s H$$

which is exactly the alpha prescription with $\alpha \sim 1$. So turbulence could in principle provide the needed viscosity—if it exists.

But here's the puzzle: Keplerian disks are stable against the usual source of turbulence in fluids. In a pipe or a channel, turbulence arises when the Reynolds number exceeds a critical value. The flow becomes unstable to small perturbations, which grow into turbulent eddies. But this instability requires the velocity profile to have the right shape—specifically, it requires an inflection point in the velocity profile, by the Rayleigh criterion.

A Keplerian disk has no such inflection point. The angular velocity decreases monotonically outward: $\Omega \propto r^{-3/2}$. Moreover, the specific angular momentum increases outward: $\ell = r^2\Omega \propto r^{1/2}$. By the Rayleigh criterion for rotating flows, this configuration is stable.

More precisely, the Rayleigh criterion says that a rotating flow is stable if

$$\frac{d(r^2\Omega)^2}{dr} > 0$$

For Keplerian rotation:

$$\frac{d(r^2\Omega)^2}{dr} = \frac{d(GMr)}{dr} = GM > 0$$

The condition is satisfied everywhere. Keplerian disks should be hydrodynamically stable.

This was the puzzle that perplexed accretion disk theorists for decades. Turbulence was needed, but the standard routes to turbulence were blocked. What could drive turbulence in a stable shear flow?

You might say, "Maybe disks aren't turbulent. Maybe accretion happens some other way." People tried this. They looked for laminar instabilities, for convection-driven transport, for spiral waves. None of it worked well enough. The observations demanded efficient angular momentum transport, and turbulence remained the only viable candidate. The question was what could trigger it.

4.3 A Forgotten Discovery

Here's a story that should be told more often, because it illustrates something important about how science works.

The answer to the viscosity problem had actually been known since 1959, but it was forgotten. In that year, the Soviet physicist Evgeny Velikhov studied the stability of a conducting fluid rotating between two cylinders with a magnetic field parallel to the rotation axis. This wasn't astrophysics—it was laboratory plasma physics, relevant to fusion reactors. Velikhov found that even a very weak

magnetic field could destabilize the flow, causing growing perturbations.

His paper was published in Russian. Few Western scientists read it. A few years later, Subrahmanyan Chandrasekhar—working on his monumental treatise on hydrodynamic stability—independently derived similar results. But Chandrasekhar was systematically cataloging mathematical solutions, not looking for applications to specific astrophysical problems. The connection to accretion disks wasn't made.

For thirty years, the result lay dormant. Think about that. The solution to one of the central problems in astrophysics existed in the literature, in two independent derivations, and nobody noticed. Accretion disk models used the alpha prescription—a parametrization of ignorance—and didn't worry too much about what α represented. Generations of students learned the thin disk equations with α as a free parameter. It was almost as if the field had decided the problem was too hard and moved on.

Then, in 1991, Steven Balbus and John Hawley revisited the problem. Hawley was running simulations of magnetized disks; Balbus was thinking about stability. They realized that the Velikhov-Chandrasekhar instability applied directly to accretion disks—and not just applied, but was unavoidable. Any weakly magnetized disk in Keplerian rotation is unstable. The turbulence that drives accretion isn't some exotic, poorly understood phenomenon. It's a simple consequence of differential rotation plus magnetism.

They called it the magnetorotational instability, or MRI. The discovery transformed the field. Within a few years, numerical simulations confirmed that the MRI drives turbulence in magnetized disks, with effective viscosity parameters $\alpha \sim 0.01$–0.1—exactly what observations required. The mystery of disk viscosity was solved, or at least understood in principle.

4.4 *How the MRI Works*

The MRI is subtle. It requires a magnetic field, but doesn't require the field to be strong. In fact, it works best when the field is weak. And it produces an instability in a flow that is hydrodynamically stable.

Here's the physical picture. Consider two fluid elements at slightly different radii in a Keplerian disk, connected by a magnetic field line. The inner element orbits faster than the outer element (because Ω decreases outward in a Keplerian disk). As the inner element pulls ahead, the field line stretches, like a rubber band connecting two runners on a circular track.

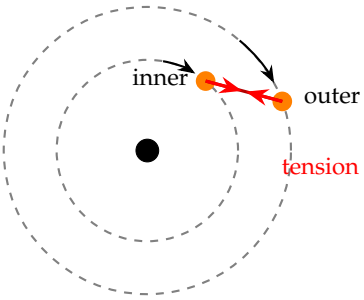

Figure 4.1: The MRI mechanism. Two fluid elements on different orbits are connected by a magnetic field line. The inner element orbits faster, stretching the line. Tension (red) slows the inner element and speeds up the outer one.

The stretched field line exerts tension. The tension pulls backward on the inner element, slowing it down, and forward on the outer element, speeding it up. Now, in a Keplerian disk, slowing down an inner element makes it fall to an even smaller radius (where it orbits even faster!), while speeding up an outer element makes it move to a larger radius (where it orbits even slower).

So the perturbation grows. The inner element, pulled backward, drops inward, while the outer element, pulled forward, moves outward. The field line stretches further, increasing the tension, which drives the elements apart even faster. It's a runaway instability.

This is completely different from ordinary shear instabilities. The Rayleigh criterion says Keplerian flow is stable, and it is—hydrodynamically. But add a magnetic field, even a weak one, and the stability disappears. The field doesn't provide the energy for the instability (gravity does, through the differential rotation). The field just provides the coupling between adjacent fluid elements that allows angular momentum to be transferred.

4.5 *The Dispersion Relation*

Let us make this more precise with a calculation. We'll work in the local approximation, treating a small patch of the disk as if it were a uniformly rotating, uniformly sheared flow. This is valid when the perturbation wavelength is much smaller than the radius.

Consider a disk with angular velocity $\Omega(r)$ threaded by a vertical magnetic field B_z. We perturb the flow with small displacements ξ_r (radial) and ξ_ϕ (azimuthal), and look for modes that grow exponentially as $e^{\gamma t}$.

The linearized equations of motion are:

$$\gamma^2 \xi_r = 2\Omega\gamma\xi_\phi - \frac{d\Omega^2}{d\ln r}\xi_r + \frac{k^2 v_A^2}{1+k^2}\xi_r$$

$$\gamma^2 \xi_\phi = -2\Omega\gamma\xi_r + \frac{k^2 v_A^2}{1+k^2}\xi_\phi$$

Here $v_A = B_z/\sqrt{4\pi\rho}$ is the Alfvén speed (the characteristic speed of magnetic wave propagation), k is the vertical wavenumber of the perturbation, and we've made some simplifications for clarity.

The term $-d\Omega^2/d\ln r$ is the square of the epicyclic frequency κ^2, which measures how strongly a displaced fluid element oscillates about its equilibrium radius. For Keplerian rotation, $\kappa^2 = \Omega^2$.

Looking for growing modes ($\gamma^2 > 0$), we find instability when

$$\frac{d\Omega^2}{d\ln r} < 0$$

For Keplerian rotation, $d\Omega^2/d\ln r = -3\Omega^2 < 0$. The condition is satisfied!

The maximum growth rate occurs when the Alfvén speed times the wavenumber matches the orbital frequency:

$$kv_A \sim \Omega$$

At this wavenumber, the growth rate is

$$\gamma_{\max} = \frac{3}{4}\Omega$$

This is remarkably fast—the instability grows on the orbital timescale. After just a few orbits, small perturbations have grown by orders of magnitude.

4.6 *What the Simulations Show*

The linear analysis tells us that the MRI exists and grows quickly. But it doesn't tell us what happens after the perturbations become large. For that, we need numerical simulations.

The first three-dimensional simulations of the MRI were performed in the mid-1990s by Hawley, Gammie, and Balbus. They started with a disk threaded by a weak magnetic field and watched what happened. The results were striking—and satisfying. After decades of wondering what drove disk turbulence, they could finally see it happen on their screens. The disk didn't just sit there; it churned.

The MRI grows as predicted by linear theory. As the perturbations become nonlinear, they break up into turbulence. The turbulence is sustained by the continuing action of the MRI—as fast as turbulent eddies dissipate, new perturbations are generated by the shear and amplified by the instability.

The turbulent flow transports angular momentum outward. The effective viscosity parameter, measured from the Reynolds and Maxwell stresses in the simulations, is $\alpha \sim 0.01$ to 0.1, depending on the initial field strength and geometry. This is exactly what's needed to explain observed accretion rates.

The simulations also revealed the importance of the magnetic field geometry. A purely vertical field (parallel to the rotation axis) produces sustained turbulence. A purely toroidal field (wrapped around the rotation axis) is less effective—it can still drive the MRI, but the saturation level is lower. Most realistic disks probably have fields with both vertical and toroidal components, tangled by the turbulence itself.

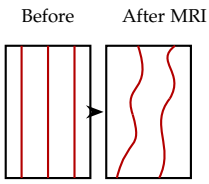

Figure 4.2: MRI-driven turbulence. Left: initial vertical magnetic field. Right: after the MRI develops, field lines become tangled and the flow is turbulent.

4.7 The Role of Net Flux

One subtlety deserves attention. The MRI requires a magnetic field, but where does the field come from? If the disk starts with no net magnetic flux (equal amounts of field pointing up and down), the MRI can still operate, amplifying whatever small-scale fluctuations exist. But the resulting turbulence is weaker than if the disk has a net vertical flux threading it.

This matters because disks probably accumulate magnetic flux over time. Infalling gas drags field lines inward. The field piles up in the inner disk, increasing the net flux. Simulations show that disks with net flux develop stronger MRI turbulence and higher effective viscosities.

In extreme cases, the accumulated flux can become dynamically important, creating a "magnetically arrested disk" (MAD). We'll discuss this more in Chapter 8, because MAD disks are particularly efficient at launching jets.

4.8 Limitations of the MRI Picture

The MRI beautifully explains why accretion disks are turbulent and how they transport angular momentum. But several questions remain open.

Does the MRI operate in all disks? The MRI requires the gas to be sufficiently ionized to couple to the magnetic field. In cool, dense regions—like the midplanes of protoplanetary disks—the ionization fraction may be too low. These "dead zones" might be essentially inviscid, with accretion proceeding only through the ionized surface layers.

What sets the saturation level? The linear theory predicts the growth rate, but not the amplitude at which the turbulence saturates. Simulations suggest $\alpha \sim 0.01$–0.1, but the precise value depends on factors we don't fully understand. Different numerical methods sometimes give different answers, raising concerns about numerical artifacts.

What about angular momentum transport by other mechanisms? The MRI isn't the only way to move angular momentum. Gravitational instabilities in massive disks can drive spiral waves that transport angular momentum. Magnetically driven winds can remove angular momentum entirely from the disk. In some systems, these mechanisms might be more important than MRI turbulence.

How does the MRI interact with other physics? Real disks have radiation, thermal structure, chemical composition, and perhaps even dust. How do these affect the MRI? Radiation pressure can

stabilize or destabilize certain modes. Thermal stratification changes the vertical structure of the turbulence. We're still working out the details.

4.9 The Broader Significance

The discovery of the MRI had impacts beyond just explaining accretion disk viscosity. It showed that magnetic fields can fundamentally change the stability properties of astrophysical flows. A flow that's stable hydrodynamically can be unstable magnetohydrodynamically. This insight has been applied to stellar interiors, planetary cores, and laboratory plasmas.

The MRI also demonstrated the power of local analysis combined with numerical simulation. The linear stability calculation is elegant but limited. The simulations are computationally expensive but reveal the nonlinear behavior. Together, they provide a complete picture of how the instability operates.

Perhaps most importantly, the MRI emphasizes that magnetic fields are not just a perturbation to disk dynamics—they're essential. Without magnetic fields, angular momentum transport in Keplerian disks is a mystery. With magnetic fields, it's natural and inevitable. This sets the stage for our discussion of jets in Chapter 8, where magnetic fields play an even more dramatic role.

5
What Disks Look Like

So far we've been building theoretical models of accretion disks. We've derived temperature profiles, density structures, and angular momentum transport mechanisms. But how do we know any of this is right?

This is a question worth taking seriously. A physicist can derive beautiful equations, but nature is under no obligation to follow them. The proof of a theory lies in its predictions—and more importantly, in predictions that could have come out differently but didn't. It's easy to explain something after the fact; the real test is predicting something you haven't seen yet.

The thin disk model passed this test in a remarkable way. Shakura and Sunyaev published their theory in 1973, before X-ray telescopes had the spectral resolution to test it properly. Their model predicted a specific spectral shape—the $\nu^{1/3}$ power law we derived in Chapter 3. When better X-ray observations came in the late 1970s and 1980s, there it was. The predicted spectrum matched the observed spectrum across orders of magnitude in photon energy. This is the kind of vindication that makes theorists giddy.

But nature is also generous with complications. The thin disk model works beautifully in some systems and fails dramatically in others. Understanding where and why it fails has been just as instructive as understanding where it succeeds. In this chapter, we'll examine what accretion disks actually look like—their spectra, their variability, their images—and see how observations both confirm and challenge the standard model.

5.1 The Multi-Color Blackbody Spectrum

A thin accretion disk is a collection of annuli, each at a different temperature. The inner disk is hot; the outer disk is cool. Each annulus radiates roughly as a blackbody at its local temperature. The observed spectrum is the sum of all these blackbodies, weighted by the

area of each annulus.

We derived the temperature profile in Chapter 3:

$$T(r) \propto r^{-3/4}$$

This means there's a continuous range of temperatures, from a maximum near the inner edge to a minimum at the outer edge. The resulting spectrum has a characteristic shape that differs from a single blackbody.

At high frequencies (short wavelengths), only the hottest regions contribute significantly. The spectrum approaches the Rayleigh-Jeans tail of the hottest blackbody: $F_\nu \propto \nu^2$.

At low frequencies (long wavelengths), the entire disk contributes, including the cool outer regions. If the outer disk dominates the area, the spectrum again approaches Rayleigh-Jeans: $F_\nu \propto \nu^2$.

But at intermediate frequencies, something interesting happens. Many annuli contribute at comparable levels. The superposition produces a spectrum flatter than Rayleigh-Jeans. Detailed calculation shows:

$$F_\nu \propto \nu^{1/3}$$

This is the multi-color blackbody spectrum, or "disk spectrum." The $\nu^{1/3}$ slope is a distinctive signature of a thin accretion disk.

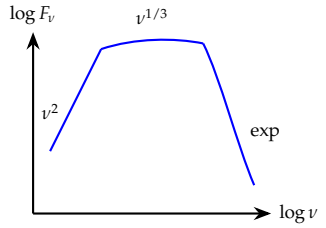

Figure 5.1: Schematic spectrum of a thin accretion disk. The $\nu^{1/3}$ slope at intermediate frequencies is a distinctive signature.

5.2 *X-Ray Binaries: The Classic Test*

The clearest tests of the thin disk model come from X-ray binaries—systems where a compact object (neutron star or black hole) accretes from a companion star. The inner disk is hot enough to emit X-rays, and X-ray satellites can measure the spectrum precisely.

This is where theory meets reality, and the meeting is instructive. Imagine you're a graduate student in the early 1980s, armed with the Shakura-Sunyaev model and fresh X-ray data from the *Einstein* satellite. You fit the model to the data and... it works. Not just vaguely, but precisely. The predicted multi-color blackbody shape matches the observed spectrum. The parameters you extract—temperature, inner radius—make physical sense. You would be forgiven for thinking the problem was solved.

But then you look more closely. The inferred inner radius seems too small. The model works at some times but not others. The same source sometimes looks completely different. Welcome to observational astrophysics, where every answer raises three more questions.

Consider Cygnus X-1, the first widely accepted black hole candidate. Its story is worth telling because it illustrates how observational evidence accumulates. In 1971, the *Uhuru* satellite discovered a bright

X-ray source coincident with a blue supergiant star, HDE 226868. The star showed Doppler shifts with a 5.6-day period—it was orbiting something. From the orbital dynamics, the unseen companion had to be more massive than about 3 solar masses. Too massive for a white dwarf, too massive for a neutron star. The only remaining option was a black hole.

The X-ray emission confirmed this picture. In its "soft state," the X-ray spectrum shows a thermal component that fits a multi-color blackbody remarkably well. The fitted inner disk temperature is $T_{\rm in} \sim 1$ keV (about 10^7 K), consistent with a disk extending to the innermost stable circular orbit around a $\sim 21 M_\odot$ black hole.

The quality of these fits is impressive. Observers fit the data with a model that has just a few parameters: the inner disk temperature, the disk normalization (related to the inner radius and distance), and the inclination angle. The model reproduces the observed spectrum across a factor of 10 in photon energy.

Source	M ($M_\odot$)	$T_{\rm in}$ (keV)	State
Cygnus X-1	21	0.5–1.0	Soft
GRS 1915+105	12	1.5–2.0	Various
LMC X-3	7	0.8–1.0	Soft
GX 339-4	6	0.7–0.9	Soft

Table 5.1: Selected black hole X-ray binaries showing thermal disk emission. The inner disk temperature depends on the accretion rate as well as the mass.

But the fits also reveal something puzzling. The derived inner radius is often smaller than the innermost stable circular orbit (ISCO) for a non-spinning black hole. Does this mean the black holes are spinning? Or is the disk model missing something?

You might say, "The model must be wrong—you can't have a disk inside the ISCO." And you'd be right that material can't orbit stably inside the ISCO. But that doesn't mean the model is wrong. It might mean the black hole is spinning, which shrinks the ISCO. Or it might mean we're measuring something subtly different from what we think.

The answer is probably both. Black holes can indeed spin, and the ISCO shrinks as spin increases (from $6GM/c^2$ for a non-spinning hole to GM/c^2 for maximal spin). But there are also systematic uncertainties in the spectral modeling—color corrections, relativistic effects, and deviations from blackbody emission all affect the derived parameters.

5.3 *Active Galactic Nuclei: The Big Blue Bump*

Supermassive black holes at the centers of galaxies also have accretion disks, but the emission peaks at much lower frequencies. For a

$10^8 M_\odot$ black hole accreting at a few percent of the Eddington rate, the inner disk temperature is only $\sim 10^5$ K. The disk emission peaks in the ultraviolet, not in X-rays.

This UV excess is called the "big blue bump." It's seen in the spectra of many quasars and Seyfert galaxies. The shape of the bump is broadly consistent with multi-color blackbody emission from a thin disk, though the details are complicated by:

- Atomic emission and absorption lines from gas near the disk
- Scattering and reprocessing by material above and below the disk
- Relativistic effects near the black hole
- The unknown spin of the black hole

Despite these complications, the basic picture holds: the UV continuum in AGN comes from thermal emission by accretion disks. The luminosity scales with accretion rate, and the spectral shape is approximately what the thin disk model predicts.

5.4 *Variability: The Flickering Disk*

Accretion disks aren't static. They flicker and flare on timescales from milliseconds to years. This variability provides independent information about disk structure, complementing the spectral analysis.

The fastest variability comes from the innermost regions. Light can't vary faster than the light-crossing time across the emitting region. For a disk around a stellar-mass black hole, the inner few Schwarzschild radii have light-crossing times of milliseconds. Indeed, X-ray binaries show variability on exactly these timescales.

But the variability isn't random noise. It has structure. Power spectra of X-ray light curves often show quasi-periodic oscillations (QPOs)—peaks in the power spectrum at specific frequencies. These QPOs are thought to arise from oscillations or precession of the inner disk, though the exact mechanism is debated.

At longer timescales, variability is driven by changes in the accretion rate. The viscous timescale at radius r is

$$t_\nu \sim \frac{r^2}{\nu} \sim \frac{r^2}{\alpha c_s H} \sim \frac{1}{\alpha \Omega}\left(\frac{r}{H}\right)^2$$

For a thin disk with $H/r \sim 0.1$, the factor $(r/H)^2 \sim 100$. Taking $\alpha \sim 0.1$ and $\Omega \sim 10^{-2}$ s^{-1} (at $r \sim 10^4 GM/c^2$ around a $10 M_\odot$ black hole), $t_\nu \sim 10^5$ s ~ 1 day.

This matches observed timescales for state transitions in X-ray binaries. When the accretion rate changes (perhaps due to instabilities

in the disk or variations in mass transfer from the companion), the disk adjusts on the viscous timescale, producing gradual changes in luminosity and spectral shape.

5.5 *Direct Imaging: Seeing the Disk*

For decades, accretion disks were inferred indirectly—from spectra, variability, and energetics. We were like doctors diagnosing a patient through X-rays and blood tests, never actually seeing the organ itself. The spectra told us about temperatures; the variability told us about sizes; the energetics told us about accretion rates. But we couldn't point at a picture and say "there's the disk."

This changed on April 10, 2019, when the Event Horizon Telescope collaboration released the first direct image of an accretion flow, around the supermassive black hole in M87. The image appeared on front pages worldwide—a fuzzy orange ring with a dark center, looking somewhat like a glazed donut viewed slightly from the side. To the public, it was the first "photograph of a black hole." To astrophysicists, it was something more subtle and in some ways more profound: direct visual confirmation of predictions made decades earlier.

The image shows a bright ring surrounding a dark central region. The bright ring is emission from hot gas in the accretion flow. The dark region is the "shadow" of the black hole—photons that passed too close to the event horizon and were captured. The size of this shadow is directly related to the black hole's mass. From the angular size of the ring (about 42 microarcseconds—the angular size of an orange on the Moon as seen from Earth) and the distance to M87 (about 55 million light-years), the EHT team inferred a mass of 6.5×10^9 solar masses. This agreed beautifully with estimates from stellar dynamics in the same galaxy.

But here's what I find most remarkable about this image: theorists had predicted exactly what it would look like. Not vaguely, but in detail. General relativistic simulations of accretion flows had been producing images like this for years before the EHT had the capability to test them. When the observation finally came, it matched the simulations with almost eerie precision. This is the scientific method at its best—prediction before observation, followed by vindication.

The ring isn't symmetric. One side is brighter than the other. This asymmetry arises from relativistic effects: material on the approaching side of the disk has its emission boosted toward the observer, while material on the receding side has its emission dimmed. The degree of asymmetry tells us about the inclination and spin of the black hole.

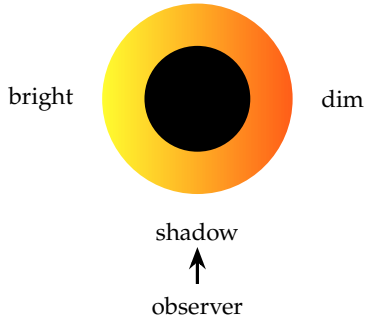

Figure 5.2: Schematic of a black hole shadow with bright emission ring, as seen by a distant observer (below). The left side is brighter because material approaches the observer there.

The EHT results broadly confirm expectations from general relativistic magnetohydrodynamic simulations. The emission comes from a region a few Schwarzschild radii across. The asymmetry is consistent with a disk viewed at moderate inclination. The overall appearance matches simulations of magnetically arrested disks (more on these in Chapter 8).

5.6 Protoplanetary Disks: A Different Regime

Not all accretion disks glow in X-rays. Young stars are surrounded by disks of gas and dust that emit primarily in the infrared. These protoplanetary disks are much cooler and much larger than disks around compact objects, but the same basic physics applies.

The ALMA telescope has revolutionized our view of protoplanetary disks by imaging them at millimeter wavelengths with exquisite angular resolution. The images reveal intricate structure: concentric rings, gaps, spiral arms, and asymmetric features.

The rings and gaps are thought to be carved by forming planets. A planet massive enough to gravitationally perturb the disk clears material from its orbit, creating a gap. The pile-up of material at the gap edges forms bright rings.

The disk around HL Tau, one of the first ALMA targets, shows at least seven concentric rings. If each gap corresponds to a planet, this 1-million-year-old system is already forming a multi-planet system. The disk is thin—the rings cast sharp shadows—just as thin disk theory predicts.

5.7 Spectral Lines: Probing Disk Dynamics

Continuum emission tells us about the temperature structure of disks. Spectral lines tell us about their dynamics.

The most powerful probe is the iron Kα emission line at 6.4 keV. Iron atoms in the inner disk are excited by hard X-rays from the corona (a hot, optically thin region above the disk) and emit a characteristic line. This line is broadened and skewed by:

- **Doppler broadening**: Gas on the approaching side is blueshifted; gas on the receding side is redshifted. The resulting profile has characteristic "horns."
- **Gravitational redshift**: Photons lose energy climbing out of the gravitational potential well near the black hole.
- **Beaming**: Relativistic motion enhances emission from the approaching side and suppresses emission from the receding side.

The combination produces an asymmetric, broad line profile. By fitting this profile, observers can constrain the inner radius of the disk (and hence the black hole spin), the inclination angle, and the emissivity profile.

Iron line spectroscopy has been used to argue that several black holes in X-ray binaries are spinning near their maximum rates. The innermost emission, inferred from the extreme red wing of the line, comes from radii smaller than the ISCO of a non-spinning hole. This requires spin.

5.8 Comparing Theory and Observation

How well does the thin disk model match observations? The short answer is: quite well, with some important caveats.

What works:

- The overall spectral shape (multi-color blackbody) matches X-ray binaries in soft states.
- The temperature scales with mass and accretion rate as predicted.
- Direct images of M87* show the expected size and asymmetry.
- Protoplanetary disk geometries (thin and flat) match predictions.

What doesn't quite work:

- The $\nu^{1/3}$ slope is often too steep; real spectra are flatter.
- The derived inner radii sometimes require implausibly high black hole spins.
- Relativistic effects (ignored in the simple model) are important near the ISCO.
- Real disks show "spectral states" that the steady thin disk model doesn't predict.

The discrepancies point toward physics we've neglected: relativistic effects, magnetic stresses, disk coronae, winds, and time-dependent behavior. The thin disk model is a starting point, not the final answer. But it's a remarkably successful starting point, capturing the essential physics of how matter falls onto compact objects.

5.9 The Importance of Not Seeing

Sometimes what we don't see is as informative as what we do. The thin disk model predicts that disks should be geometrically thin

($H/r \ll 1$) and optically thick ($\tau \gg 1$). But in some systems, we see evidence for thick, optically thin flows.

X-ray binaries in their "hard state" show spectra dominated by hard power-law emission rather than thermal disk emission. The disk either retreats to large radii or transforms into a different structure entirely. We'll explore these alternatives in the next chapter.

AGN with low accretion rates (Low-Luminosity AGN, or LLAGN) also lack the thermal disk signature. The supermassive black hole at the center of our own galaxy, Sgr A*, accretes at a rate far below Eddington, and shows no evidence for a standard thin disk. Something else is happening in these systems.

The absence of thin disk signatures in these systems tells us that the thin disk model has a limited domain of validity. Understanding where and why it breaks down is the subject of our next chapter.

5.10 What Do We Actually Observe?

We've talked about "observing" accretion disks, but what do we really observe? Photons. Photons that left a distant source, traveled across space, passed through our atmosphere (or were collected above it by satellites), interacted with our detectors, and produced electrical signals. Between the disk and our knowledge lies a long chain of inferences.

The photons tell us about conditions at the photosphere—the surface from which radiation escapes. What happens below the photosphere, in the disk interior, we can only infer from models. We assume hydrostatic equilibrium. We assume the gas is in local thermodynamic equilibrium. We assume the opacity follows certain laws. Each assumption could be wrong.

And yet the models work. The same theoretical framework developed in the 1970s continues to match observations from instruments that didn't exist then. The spectrum of Cygnus X-1 looks like a multicolor blackbody. The image of M87* looks like a relativistic accretion flow. The rings in HL Tau's disk are sharp, as expected for a thin disk. The temperatures we calculate correspond to real temperatures. The disk really is thin.

6
When Thin Disks Fail

The thin disk model has been spectacularly successful. But like all models, it has limits. And exploring those limits teaches us more than the model itself ever could.

This is a pattern in physics. We build a theory that works beautifully in some regime. We push it to extremes. It breaks. And in the breaking, we learn something new. Newton's mechanics works beautifully at everyday speeds, but at speeds approaching light, it fails—and that failure led to special relativity. Classical electromagnetism works beautifully for macroscopic charges, but for atomic-scale phenomena, it fails—and that failure led to quantum mechanics.

The thin disk model fails in two opposite regimes: when the accretion rate is too high, and when it's too low. In both cases, the assumptions that made the model tractable—geometric thinness, efficient cooling, optical thickness—break down. The disk transforms into something else entirely: a puffed-up, radiation-dominated structure at high rates, or a hot, tenuous, quasi-spherical flow at low rates.

These "failed" states aren't failures of nature. They're failures of our simplified model to capture nature's full behavior. And they turn out to be remarkably interesting in their own right. Some of the most dramatic phenomena in astrophysics—powerful jets, X-ray state transitions, the underluminous black hole at our galaxy's center—occur precisely where thin disks don't work.

6.1 The Two Failure Modes

Recall the key assumptions of the thin disk model:

1. The disk is geometrically thin: $H \ll r$
2. The disk is optically thick: $\tau \gg 1$
3. The disk radiates efficiently: cooling time $\ll$ infall time
4. The disk is in thermal equilibrium: heating = cooling

These assumptions can fail in two opposite ways:

High accretion rates: When $\dot{M}$ approaches the Eddington limit, radiation pressure becomes important. The disk puffs up, violating assumption 1. Photons get trapped in the flow and are advected inward rather than escaping, violating assumption 3.

Low accretion rates: When $\dot{M}$ is very low, the density drops and the disk becomes optically thin, violating assumption 2. The gas can't cool efficiently; it heats up and puffs up, violating assumptions 1 and 3.

In both cases, the "thin disk" becomes something else. Let's examine each regime.

6.2 The Eddington Limit

Before discussing high accretion rates, we need to understand what limits them. The Eddington luminosity is the luminosity at which radiation pressure on electrons balances gravity on protons:

$$L_{\rm Edd} = \frac{4\pi G M m_p c}{\sigma_T} \approx 1.3 \times 10^{38} \left(\frac{M}{M_\odot}\right) \text{ erg/s}$$

where σ_T is the Thomson scattering cross-section.

If accretion liberates gravitational energy with efficiency η (typically $\eta \sim 0.1$ for thin disks around black holes), then the Eddington accretion rate is

$$\dot{M}_{\rm Edd} = \frac{L_{\rm Edd}}{\eta c^2} \approx 2 \times 10^{-8} \left(\frac{M}{M_\odot}\right) M_\odot/\text{yr}$$

For accretion rates $\dot{M} \gtrsim \dot{M}_{\rm Edd}$, the outward radiation force exceeds gravity. Naively, accretion should stop. But nature finds ways around this limit.

6.3 Slim Disks

When $\dot{M}$ approaches $\dot{M}_{\rm Edd}$, two things happen. First, radiation pressure becomes comparable to gas pressure, puffing up the disk. Second, the photon diffusion time becomes comparable to the infall time—photons get trapped in the flow.

A photon at the disk midplane must diffuse through an optical depth τ to escape. The diffusion time is

$$t_{\rm diff} \sim \frac{\tau H}{c}$$

The infall time at radius r is roughly

$$t_{\rm infall} \sim \frac{r}{v_r} \sim \frac{r}{\alpha c_s} \cdot \frac{r}{H}$$

When $t_{\rm diff} \gtrsim t_{\rm infall}$, photons are advected inward faster than they can escape. The energy they would have radiated is carried into the black hole instead. This is called photon trapping or advection-dominated cooling.

The resulting structure is called a "slim disk." It's thicker than a standard thin disk ($H/r \sim 0.1$–0.5) but not truly thick. The luminosity no longer increases linearly with $\dot{M}$; instead, it saturates near $L_{\rm Edd}$ because excess energy is swallowed by the black hole.

Slim disks can accrete at rates far exceeding $\dot{M}_{\rm Edd}$. The excess energy just disappears into the black hole. Observations of some ultraluminous X-ray sources (ULXs) suggest accretion rates of 10–100 times Eddington, implying slim disk geometries.

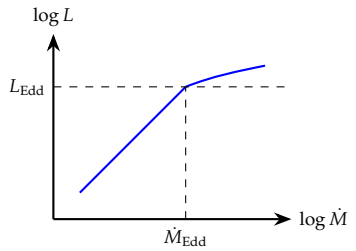

Figure 6.1: Luminosity vs. accretion rate. Below Eddington, $L \propto \dot{M}$. Above Eddington, the luminosity saturates as photons are trapped.

6.4 *The Low-Luminosity Puzzle*

Now consider the opposite extreme. What happens when $\dot{M} \ll \dot{M}_{\rm Edd}$?

The thin disk model predicts that the temperature and density scale with accretion rate. At low $\dot{M}$, the disk should be cooler and more tenuous. But it should still be geometrically thin and optically thick, just dimmer.

Observations tell a different story, and it's a fascinating one.

The black hole at the center of our galaxy, Sagittarius A*, offers the clearest case. We know a lot about this object. It has a mass of about 4 million solar masses, precisely measured from the orbits of stars that swing around it at thousands of kilometers per second. It sits just 26,000 light-years away—close enough that we can resolve individual stars near it. And it's surrounded by hot gas that should, in principle, be accreting.

Here's the puzzle: given the amount of gas we see near Sgr A*, and assuming it falls in with a thin disk's efficiency, the black hole should be about 10^5 times brighter than it actually is. Where is the missing luminosity?

For years, this was a genuine mystery. Maybe the gas isn't actually accreting? Maybe it's being expelled by some unseen wind? Maybe the black hole is spinning in a way that extracts energy rather than releasing it?

The answer turned out to be simpler and stranger: the gas *is* accreting, but it's not forming a thin disk. Instead, it forms a hot, puffy flow that swallows most of its energy before that energy can be radiated. The radiation we don't see isn't being blocked or redirected—it's never being produced in the first place. The gravitational energy released by accretion goes into heating the gas, and that heat is carried into the black hole rather than being radiated away.

This is fundamentally different from a thin disk. In a thin disk, all the gravitational energy is radiated. In this alternative flow, almost none of it is. The accretion is happening; the luminosity is just being suppressed.

6.5 *Advection-Dominated Accretion Flows*

The answer, developed in the 1990s by Narayan, Yi, and collaborators, is the advection-dominated accretion flow, or ADAF.[1]

[1] Also called radiatively inefficient accretion flows (RIAFs) by some authors.

The idea had precursors going back decades—Ichimaru in 1977, Rees and collaborators in the 1980s—but it was Narayan and Yi who put the pieces together into a coherent model and showed how it applied to real systems. Their work in 1994-1995 triggered an avalanche of applications. Suddenly, low-luminosity sources that had been puzzles—Sgr A*, the nuclei of nearby galaxies, quiescent X-ray binaries—all made sense.

The key insight is beautifully simple, once you see it. At low accretion rates, the gas density is so low that collisions between particles are rare. Specifically, the collision timescale between ions (protons) and electrons exceeds the infall timescale. This means ions and electrons can have different temperatures.

Why does this matter? Think about it for a moment. In an ordinary gas, collisions keep all particles at the same temperature. Energy flows from hot to cold until equilibrium is reached. But if collisions are too slow, equilibrium can't be established. Different species can maintain different temperatures.

Now here's the crucial point. In a plasma, viscous heating primarily heats the ions (because they carry most of the momentum). Radiation is emitted primarily by electrons (because they're lighter and radiate more efficiently). In a thin disk, collisions keep ions and electrons at the same temperature, so the energy deposited in ions is quickly shared with electrons and radiated away.

Think of it like a party where the big guests do all the dancing (heating) but the small guests do all the talking (radiating). If people mingle freely, the energy gets shared. But if the room is too sparse for mingling, the dancers keep dancing and the talkers have nothing to say.

But if collisions are too slow, the electrons stay cool while the ions get hot. The ions can reach temperatures approaching the virial temperature:

$$T_i \sim \frac{GMm_p}{kr} \sim 10^{13}\ \mathrm{K} \times \left(\frac{r}{r_g}\right)^{-1}$$

where $r_g = GM/c^2$ is the gravitational radius.

At these temperatures, the ion thermal velocity approaches the speed of light. The gas pressure is enormous, and the disk puffs up to become quasi-spherical ($H \sim r$).

The electrons remain much cooler, perhaps 10^9–10^{10} K. They radiate through synchrotron emission (in the magnetic field), bremsstrahlung (from collisions), and inverse Compton scattering (of synchrotron photons). But the total radiation is far less than the gravitational energy released. Most of the energy is advected into the black hole with the hot ions, hence "advection-dominated."

The efficiency of an ADAF is very low—perhaps $\eta \sim 10^{-3}$ instead of the $\eta \sim 0.1$ of a thin disk. For the same accretion rate, an ADAF is 100 times dimmer than a thin disk. This explains why Sgr A* is so faint despite having a healthy supply of gas to accrete.

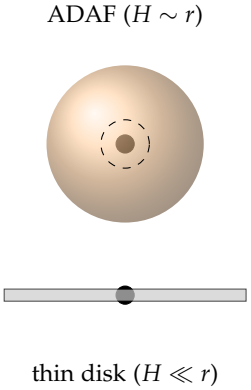

Figure 6.2: Comparison of accretion flow geometries. Top: an ADAF is quasi-spherical ($H \sim r$). Bottom: a thin disk is highly flattened ($H \ll r$).

6.6 *The Critical Accretion Rate*

ADAFs exist only below a critical accretion rate $\dot{M}_{\rm crit}$. Above this rate, the density is high enough that ion-electron collisions are frequent, forcing them to the same temperature. The gas cools efficiently, collapses to the midplane, and forms a thin disk.

The critical rate is roughly

$$\dot{M}_{\rm crit} \sim \alpha^2 \dot{M}_{\rm Edd}$$

For $\alpha \sim 0.1$, this gives $\dot{M}_{\rm crit} \sim 0.01 \dot{M}_{\rm Edd}$. Systems accreting below about 1% of Eddington are likely to have ADAF-like inner regions.

This explains a curious observational pattern. X-ray binaries in their "soft state" (high luminosity) show thermal disk spectra. The same systems in their "hard state" (lower luminosity) show power-law spectra with no thermal disk component. The transition occurs at luminosities of a few percent of Eddington—exactly where the ADAF critical rate predicts.

The picture is that at high accretion rates, a thin disk extends all the way to the ISCO. At low accretion rates, the inner thin disk evaporates into an ADAF, leaving only an outer thin disk at large radii.

6.7 *Jets from Hot Flows*

There's another crucial difference between thin disks and ADAFs. ADAFs seem much better at producing jets.

Observations show that jets are common in the hard state (ADAF-like) and rare or weak in the soft state (thin disk). The radio emission from X-ray binaries, which traces jet activity, is strongly correlated with the hard spectral state.

Why should hot, thick flows produce jets more readily than thin disks? Several factors probably contribute:

- **Magnetic field geometry:** In a thin disk, the field is confined near the midplane and stretched azimuthally. In an ADAF, the field can develop a more vertical component, suitable for launching jets.
- **Rotation support:** In an ADAF, the flow isn't purely Keplerian. Pressure support allows the angular velocity to deviate from Keplerian values, which affects how magnetic field lines are wound up.
- **Energy availability:** ADAFs store most of the gravitational energy as internal energy of the hot ions, rather than radiating it away. This energy reservoir could power jets.

We'll return to the disk-jet connection in Chapters 8 and 10. For now, the key point is that the geometry and thermodynamics of the accretion flow affect more than just the spectrum—they affect the entire outflow structure.

6.8 The Transition Region

Real systems aren't always purely thin disks or purely ADAFs. Often, both structures coexist at different radii.

A common configuration has an outer thin disk that truncates at some radius $r_{\rm tr}$, with an inner ADAF filling the region between $r_{\rm tr}$ and the black hole. The truncation radius depends on the accretion rate: at higher $\dot{M}$, $r_{\rm tr}$ moves inward; at lower $\dot{M}$, it moves outward.

Evidence for this geometry comes from spectral analysis. The thermal disk component has an inner temperature that implies a truncated disk—the disk doesn't extend as close to the black hole as it would if it reached the ISCO. The hard power-law component is interpreted as emission from the hot inner ADAF, possibly Comptonized by hot electrons.

The transition between disk and ADAF isn't sharp. There's probably a gradual transition zone where the disk puffs up and heats up. Modeling this transition zone is challenging and an active area of research.

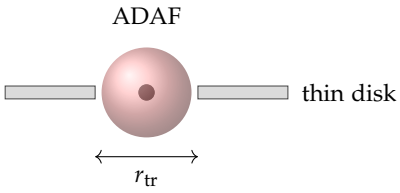

Figure 6.3: A truncated disk geometry: an outer thin disk with an inner ADAF. This configuration is common at intermediate accretion rates.

6.9 Super-Eddington Flows: Outflows and Photon Bubbles

Return to the high-accretion-rate regime. When $\dot{M} \gg \dot{M}_{\rm Edd}$, the disk is highly super-Eddington, and dramatic things happen.

First, radiation pressure launches powerful outflows. The radiation field is so intense that it blows material off the disk surface. These

outflows can carry away most of the accreting mass—only a small fraction actually reaches the black hole.

Second, the disk can develop "photon bubbles"—regions where radiation pressure pushes gas aside, creating low-density channels through which photons can escape. This is a type of convection driven by radiation rather than heat.

Numerical simulations of super-Eddington accretion show highly dynamic flows: thick disks with radiation-driven winds, variable luminosities, and complex geometries. The physics is nonlinear and difficult to model analytically.

Observationally, super-Eddington accretion may occur in:

- Ultraluminous X-ray sources (ULXs): Some ULXs have luminosities implying $\dot{M} \sim 10$–$100\ \dot{M}_{\rm Edd}$ if the accretor is a stellar-mass black hole.

- Tidal disruption events (TDEs): When a star is torn apart by a supermassive black hole, the initial fallback rate can far exceed Eddington.

- Narrow-line Seyfert 1 galaxies: These AGN show evidence for high accretion rates and may have slim disk geometries.

6.10 *The Parameter Space of Accretion Flows*

We can summarize the different accretion flow regimes in terms of two parameters: the accretion rate $\dot{m} = \dot{M}/\dot{M}_{\rm Edd}$ and the radius r (in units of r_g).

$\dot{m}$	Geometry	Properties
$\lesssim 0.01$	ADAF	Hot, thick, radiatively inefficient
0.01–0.3	Thin disk	Cool, thin, efficient
0.3–1	Slim disk	Warm, moderately thick, trapped photons
$\gtrsim 1$	Super-Eddington	Thick, outflowing, photon bubbles

Table 6.1: Rough boundaries between accretion flow regimes. The exact boundaries depend on the viscosity parameter α and other details.

The boundaries aren't sharp, and individual systems can move between regimes as their accretion rates change. X-ray binaries famously exhibit state transitions, shifting between hard states (ADAF-like) and soft states (thin disk) on timescales of weeks to months. These transitions are probably driven by changes in $\dot{M}$ that push the system across the critical rate.

6.11 What We Still Don't Understand

Despite decades of work, several aspects of non-standard accretion flows remain unclear:

The ion-electron coupling: ADAFs require inefficient coupling between ions and electrons. But we don't know the coupling rate from first principles. It depends on plasma physics at scales too small to simulate with current computers.

The transition physics: How exactly does a thin disk evaporate into an ADAF? What sets the truncation radius? Models exist, but they involve assumptions that haven't been tested.

The role of magnetic fields: We've described ADAFs as if magnetic fields were a minor perturbation. But simulations suggest that magnetic pressure can be comparable to gas pressure. This changes the dynamics significantly.

The outflows: ADAFs and super-Eddington flows both drive powerful outflows. But predicting the mass and energy in these outflows from first principles is difficult.

These uncertainties matter because they affect the observables. The spectrum, variability, and jet power all depend on details of the accretion flow structure that we don't fully understand.

Despite these uncertainties, the broad picture is clear: thin disks are just one member of a diverse family of accretion solutions. Nature explores this parameter space, producing flows with different geometries, efficiencies, and observational signatures depending on the accretion rate and the properties of the central object. Jets are more powerful from ADAFs than from thin disks. State transitions in X-ray binaries involve transitions between these regimes. The underluminous black holes in the centers of nearby galaxies are in ADAF states. The thin disk model works beautifully in its domain of validity—and that domain includes most of the accretion we can observe—but the full zoo of accretion solutions is richer than the standard model alone would suggest.

7
Magnetic Fields in Disks

Imagine you could see magnetic field lines. In an accretion disk, you would see them twisted, stretched, wound into spirals, and tangled into impossible knots—yet remarkably, they never break, never intersect. They thread through the plasma like infinitely strong threads sewn into stretchy fabric. Pull the fabric, and the threads stretch with it. Try to pull a thread out, and you tear the fabric itself. This is not just a poetic image; it's the key to understanding why magnetic fields transform accretion disk physics from an academic puzzle into a cosmic engine.

We've mentioned magnetic fields before—the MRI requires them, jet launching depends on them—but we haven't systematically explored their role. This chapter rectifies that omission. Magnetic fields aren't a small correction to the hydrodynamic picture; they're fundamental to how disks work.

In a way, the history of accretion disk theory is a story of gradually recognizing the importance of magnetic fields. First they were ignored entirely—pure hydrodynamics seemed simpler, and anyway, magnetic fields in space were thought to be weak. Then they were invoked to explain viscosity through the MRI. Now we understand that magnetic fields can dominate the dynamics of the inner disk, control the launching of jets, and determine whether a given system is radio-loud or radio-quiet. The story is not finished. We are still discovering ways that magnetic fields surprise us.

7.1 The Full MHD Equations

Before diving into specific phenomena, let's establish the complete set of magnetohydrodynamic equations that govern magnetized accretion flows. I'll write them all down at once—not because you need to memorize them, but because having them in one place helps you see the structure. The equations combine fluid dynamics with electromagnetism, and the magnetic force appears in an interesting

way: as both pressure and tension.

You might ask, "Why start with equations? Shouldn't we build physical intuition first?" Fair point. But these equations are actually simpler than they look, and the physical content emerges from studying them. Think of this section as laying out the playing board before we start the game.

Mass conservation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0$$

Momentum conservation:

$$\rho \frac{\partial \vec{v}}{\partial t} + \rho(\vec{v} \cdot \nabla)\vec{v} = -\nabla P + \frac{1}{4\pi}(\nabla \times \vec{B}) \times \vec{B} + \rho \vec{g}$$

The magnetic force $(\nabla \times \vec{B}) \times \vec{B}/4\pi$ can be rewritten as

$$\frac{1}{4\pi}(\nabla \times \vec{B}) \times \vec{B} = -\nabla \left(\frac{B^2}{8\pi} \right) + \frac{1}{4\pi}(\vec{B} \cdot \nabla)\vec{B}$$

The first term is magnetic pressure; the second is magnetic tension.

Energy conservation:

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\rho v^2 + \frac{P}{\gamma - 1} + \frac{B^2}{8\pi} \right) + \nabla \cdot \left[\left(\frac{1}{2}\rho v^2 + \frac{\gamma P}{\gamma - 1} \right) \vec{v} + \frac{c}{4\pi} \vec{E} \times \vec{B} \right] = \rho \vec{v} \cdot \vec{g} - \mathcal{L}$$

where $\mathcal{L}$ represents radiative losses and $\vec{E} \times \vec{B}$ is the Poynting flux.

Induction equation:

$$\frac{\partial \vec{B}}{\partial t} = \nabla \times (\vec{v} \times \vec{B}) + \eta \nabla^2 \vec{B}$$

Solenoidal constraint:

$$\nabla \cdot \vec{B} = 0$$

These equations are coupled and nonlinear. They support a rich variety of wave modes: Alfvén waves (transverse oscillations of field lines), slow and fast magnetosonic waves (compressional waves modified by magnetic pressure), and various instabilities. The interplay of these modes determines the turbulent state of the disk.

7.2 The Induction Equation in Detail

The induction equation deserves special attention because it governs how magnetic fields evolve. Let's derive it carefully.

Start with Faraday's law:

$$\frac{\partial \vec{B}}{\partial t} = -c \nabla \times \vec{E}$$

We need to relate $\vec{E}$ to the fluid motion and current. In a conducting fluid, Ohm's law in the fluid rest frame is $\vec{J}' = \sigma\vec{E}'$, where primes denote the rest frame. Transforming to the lab frame (to first order in v/c):

$$\vec{E}' = \vec{E} + \frac{\vec{v}}{c} \times \vec{B}, \qquad \vec{J}' = \vec{J}$$

So

$$\vec{J} = \sigma\left(\vec{E} + \frac{\vec{v}}{c} \times \vec{B}\right)$$

Using Ampère's law $\vec{J} = (c/4\pi)\nabla \times \vec{B}$ (neglecting displacement current, valid for non-relativistic flows):

$$\vec{E} = -\frac{\vec{v}}{c} \times \vec{B} + \frac{c}{4\pi\sigma}\nabla \times \vec{B}$$

Substituting into Faraday's law:

$$\frac{\partial \vec{B}}{\partial t} = \nabla \times (\vec{v} \times \vec{B}) - \nabla \times \left(\frac{c^2}{4\pi\sigma}\nabla \times \vec{B}\right)$$

For uniform resistivity $\eta = c^2/(4\pi\sigma)$:

$$\frac{\partial \vec{B}}{\partial t} = \nabla \times (\vec{v} \times \vec{B}) + \eta\nabla^2\vec{B}$$

The magnetic Reynolds number compares advection to diffusion:

$$R_m = \frac{vL}{\eta} = \frac{4\pi\sigma v L}{c^2}$$

For a fully ionized hydrogen plasma, the Spitzer conductivity is

$$\sigma \approx \frac{n_e e^2 \tau_e}{m_e} \approx 10^7 T^{3/2} \text{ s}^{-1}$$

where T is in Kelvin. For $T = 10^7$ K, $\sigma \sim 10^{17}$ s^{-1}, giving

$$\eta \sim \frac{(3 \times 10^{10})^2}{4\pi \times 10^{17}} \sim 10^3 \text{ cm}^2/\text{s}$$

With $v \sim 10^8$ cm/s and $L \sim 10^7$ cm (typical inner disk values):

$$R_m \sim \frac{10^8 \times 10^7}{10^3} \sim 10^{12}$$

This enormous Reynolds number means ideal MHD ($\eta \to 0$) is an excellent approximation for the bulk of the disk. Resistive effects matter only in thin current sheets where gradients are large.

You might say, "That's an enormous Reynolds number! Resistivity must be completely irrelevant." And you'd be mostly right—for the bulk of the disk. But watch carefully what happens in thin current sheets where the field reverses direction. There, gradients become large, L becomes small, and R_m can drop to order unity. That's where the interesting physics happens: reconnection, flares, particle acceleration. The boring bulk obeys ideal MHD; the action is in the singular layers.

7.3 *Flux Freezing and Its Proof*

In ideal MHD ($\eta = 0$), magnetic field lines are "frozen" into the plasma. This is Alfvén's theorem. Let's prove it.

Consider a material surface $S(t)$ that moves with the fluid, bounded by a material curve $C(t)$. The magnetic flux through this surface is

$$\Phi(t) = \int_{S(t)} \vec{B} \cdot d\vec{A}$$

To compute $d\Phi/dt$, we use Leibniz's rule for a moving surface:

$$\frac{d\Phi}{dt} = \int_S \frac{\partial \vec{B}}{\partial t} \cdot d\vec{A} + \oint_C (\vec{v} \times \vec{B}) \cdot d\vec{l}$$

The second term accounts for the motion of the boundary.

Using Stokes' theorem on the line integral:

$$\oint_C (\vec{v} \times \vec{B}) \cdot d\vec{l} = \int_S \nabla \times (\vec{v} \times \vec{B}) \cdot d\vec{A}$$

So

$$\frac{d\Phi}{dt} = \int_S \left[\frac{\partial \vec{B}}{\partial t} + \nabla \times (\vec{v} \times \vec{B}) \right] \cdot d\vec{A}$$

But the ideal induction equation says $\partial \vec{B}/\partial t = \nabla \times (\vec{v} \times \vec{B})$, so

$$\frac{d\Phi}{dt} = 0$$

The flux through any material surface is conserved. This means field lines move with the fluid—if you tag fluid elements lying along a field line, they will always lie along a field line.

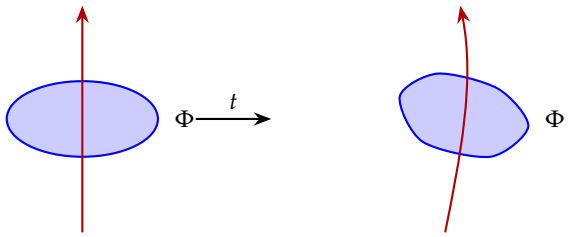

Figure 7.1: Flux freezing. A material loop (blue) deforms with the flow, but the magnetic flux Φ through it remains constant. The field line threads through the loop at all times.

Flux freezing has profound consequences:

- Field lines can be stretched but not broken (in ideal MHD)
- Magnetic topology is preserved
- Field strength increases when fluid is compressed
- Field lines can store energy indefinitely

The last point is particularly important. In ideal MHD, magnetic energy cannot be dissipated. Field lines can be wound up, stretched, and tangled, storing ever more energy. Dissipation requires non-ideal effects—resistivity, which allows field lines to slip through the plasma and reconnect.

This leads to a curious situation. You can pump energy into the magnetic field indefinitely—differential rotation winds up the field, turbulent motions tangle it—but in ideal MHD, you can never get that energy back out. It's like filling a bathtub with no drain. Eventually something has to give. That "something" is the breakdown of ideal MHD in thin current sheets, where reconnection finally provides the drain.

7.4 *Winding Up: Quantitative Treatment*

Let's analyze field line winding more carefully. Consider an initially vertical field $\vec{B} = B_z \hat{z}$ threading a differentially rotating disk.

In cylindrical coordinates (r, ϕ, z), the ideal induction equation component by component:

z-component:

$$\frac{\partial B_z}{\partial t} = -\frac{1}{r}\frac{\partial}{\partial r}(r v_r B_z) - \frac{1}{r}\frac{\partial}{\partial \phi}(v_\phi B_z)$$

For axisymmetry and no radial flow, $\partial B_z/\partial t = 0$. The vertical field is unchanged.

ϕ-component:

$$\frac{\partial B_\phi}{\partial t} = B_z \frac{\partial v_\phi}{\partial z} + B_r \frac{\partial v_\phi}{\partial r} - B_\phi \frac{\partial v_r}{\partial r} + \ldots$$

For Keplerian rotation $v_\phi = r\Omega(r)$ in a thin disk with $\partial v_\phi/\partial z = 0$ and $B_r = 0$ initially:

$$\frac{\partial B_\phi}{\partial t} = B_z \frac{\partial v_\phi}{\partial z} = 0$$

Wait—this gives no winding! The problem is that we assumed a purely vertical field. In reality, the field has some radial structure. Consider a field line that passes through the disk at two different radii, r_1 and r_2. The footpoints rotate at different rates, creating azimuthal shear.

A more careful analysis: suppose the field line has a small radial component due to its curvature. If the field makes an angle θ with the vertical, then $B_r \sim B_z \tan\theta$. The winding term becomes:

$$\frac{\partial B_\phi}{\partial t} = B_r r \frac{d\Omega}{dr} = B_z \tan\theta \cdot r \frac{d\Omega}{dr}$$

For Keplerian rotation, $rd\Omega/dr = -3\Omega/2$, so

$$B_\phi(t) = -\frac{3}{2} B_z \tan\theta \cdot \Omega t$$

After time t, the ratio of toroidal to poloidal field is

$$\frac{B_\phi}{B_z} \sim \frac{3}{2} \tan\theta \cdot \Omega t$$

After one orbital period ($\Omega t = 2\pi$):

$$\frac{B_\phi}{B_z} \sim 10 \tan\theta$$

Even if the field is nearly vertical ($\theta = 5$, $\tan\theta \approx 0.1$), the toroidal field equals the poloidal field after just one orbit. After ten orbits, $B_\phi/B_z \sim 10$—the field is predominantly toroidal.

This is the "Omega effect" in dynamo theory: differential rotation winds poloidal field into toroidal field. It's the most basic mechanism for creating toroidal field in rotating systems.

7.5 The MRI Dispersion Relation

Let's derive the MRI dispersion relation carefully, as it's central to understanding angular momentum transport.

Consider a disk with angular velocity $\Omega(r)$ threaded by a vertical magnetic field B_z. We perturb the system with small displacements of the form $\xi \propto e^{i(k_z z - \omega t)}$, where k_z is the vertical wavenumber.

The linearized equations of motion in the co-rotating frame are:

$$-\omega^2 \xi_r = 2i\omega\Omega\xi_\phi + \left(\frac{d\Omega^2}{d\ln r} + k_z^2 v_A^2\right)\xi_r \tag{7.1}$$

$$-\omega^2 \xi_\phi = -2i\omega\Omega\xi_r + k_z^2 v_A^2 \xi_\phi \tag{7.2}$$

where $v_A = B_z/\sqrt{4\pi\rho}$ is the Alfvén speed.

The first equation includes: the Coriolis force ($2i\omega\Omega\xi_\phi$), the epicyclic restoring force ($d\Omega^2/d\ln r \cdot \xi_r$), and magnetic tension ($k_z^2 v_A^2 \xi_r$).

For non-trivial solutions, the determinant of the coefficient matrix must vanish:

$$\omega^4 - \omega^2\left(\kappa^2 + 2k_z^2 v_A^2\right) + k_z^2 v_A^2\left(k_z^2 v_A^2 + \frac{d\Omega^2}{d\ln r}\right) = 0$$

where $\kappa^2 = 2\Omega(2\Omega + rd\Omega/dr)$ is the square of the epicyclic frequency.

For Keplerian rotation: $\kappa^2 = \Omega^2$ and $d\Omega^2/d\ln r = -3\Omega^2$.

Let $x = k_z v_A/\Omega$. The dispersion relation becomes:

$$\left(\frac{\omega}{\Omega}\right)^4 - \left(\frac{\omega}{\Omega}\right)^2 (1 + 2x^2) + x^2(x^2 - 3) = 0$$

This is a quadratic in $(\omega/\Omega)^2$. The solutions are:

$$\left(\frac{\omega}{\Omega}\right)^2 = \frac{(1+2x^2) \pm \sqrt{(1+2x^2)^2 - 4x^2(x^2-3)}}{2}$$

Instability requires $\omega^2 < 0$, i.e., one of the roots is negative. This occurs when

$$x^2(x^2-3) < 0 \quad \Rightarrow \quad 0 < x^2 < 3$$

The instability exists for $0 < k_z v_A < \sqrt{3}\Omega$. The maximum growth rate occurs at $k_z v_A = (15/16)^{1/2}\Omega \approx 0.97\Omega$, giving

$$\gamma_{\max} = \frac{3}{4}\Omega$$

This is remarkably fast—the instability grows on the orbital timescale. After just two orbital periods, perturbations have grown by $e^{3\pi} \approx 10^4$.

The physical interpretation: magnetic tension couples fluid elements at different radii. The coupling transfers angular momentum from the faster-spinning inner element to the slower-spinning outer element, violating the stability criterion that angular momentum increase outward.

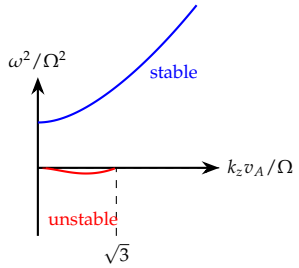

Figure 7.2: MRI dispersion relation. The red curve shows the unstable branch with $\omega^2 < 0$. Instability occurs for $0 < k_z v_A/\Omega < \sqrt{3}$.

7.6 Non-Ideal MHD Effects

While ideal MHD is an excellent approximation for hot, fully ionized disks, several non-ideal effects become important in cooler or partially ionized regions:

Ohmic resistivity: We've already discussed this. Collisions between electrons and ions (or neutrals) impede current flow, allowing field lines to slip through the plasma. The induction equation becomes:

$$\frac{\partial \vec{B}}{\partial t} = \nabla \times (\vec{v} \times \vec{B}) + \eta_O \nabla^2 \vec{B}$$

Hall effect: In a weakly ionized plasma, electrons and ions drift at different velocities in a magnetic field. The Hall term modifies the induction equation:

$$\frac{\partial \vec{B}}{\partial t} = \nabla \times \left[(\vec{v} - \vec{v}_H) \times \vec{B}\right] + \eta_O \nabla^2 \vec{B}$$

where $\vec{v}_H = \vec{J}/(en_e) = c(\nabla \times \vec{B})/(4\pi e n_e)$ is the Hall drift velocity.

The Hall effect doesn't dissipate magnetic energy—it just changes the field evolution. It can alter the MRI, either stabilizing or destabilizing depending on the field orientation.

Ambipolar diffusion: In a plasma with neutrals, the ions are tied to field lines but neutrals aren't. Ion-neutral collisions cause the field

to drift through the neutral gas:

$$\frac{\partial \vec{B}}{\partial t} = \nabla \times \left[(\vec{v} + \vec{v}_{AD}) \times \vec{B} \right]$$

where $\vec{v}_{AD}$ is the ambipolar drift velocity, proportional to $(\nabla \times \vec{B}) \times \vec{B}$.

Ambipolar diffusion does dissipate energy, but only perpendicular to field lines. It's important in molecular clouds and the outer regions of protoplanetary disks.

The relative importance of these effects is characterized by dimensionless numbers. Define the magnetic diffusivities:

$$\eta_O = \frac{c^2}{4\pi\sigma} \quad \text{(Ohmic)} \tag{7.3}$$

$$\eta_H = \frac{cB}{4\pi e n_e} \quad \text{(Hall)} \tag{7.4}$$

$$\eta_A = \frac{B^2}{4\pi \gamma_i \rho_i \rho_n} \quad \text{(Ambipolar)} \tag{7.5}$$

where γ_i is the ion-neutral coupling coefficient.

The ratios of these to the turbulent diffusivity $\eta_T \sim \alpha c_s H$ determine which effects matter:

- $\eta_O / \eta_T > 1$: Ohmic diffusion suppresses MRI (dead zone)
- $\eta_H / \eta_T > 1$: Hall effect modifies MRI
- $\eta_A / \eta_T > 1$: Ambipolar diffusion damps MRI

In protoplanetary disks, all three effects can be important. The midplane at 1–10 AU is often a "dead zone" where the MRI is suppressed. Accretion may proceed only in the ionized surface layers, or through other mechanisms (gravitational instability, vertical shear instability).

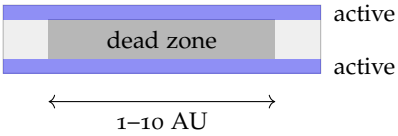

Figure 7.3: Dead zone in a protoplanetary disk. The midplane is too poorly ionized for MRI; only the surface layers are magnetically active.

7.7 *Magnetic Helicity*

Magnetic helicity is a topological quantity that measures the linkage and twist of magnetic field lines:

$$H = \int \vec{A} \cdot \vec{B} \, dV$$

where $\vec{A}$ is the vector potential ($\vec{B} = \nabla \times \vec{A}$).

In ideal MHD, magnetic helicity is exactly conserved:

$$\frac{dH}{dt} = 0$$

This conservation has profound implications. Even as the field is stretched and tangled by turbulence, the total helicity remains fixed. This constrains how much energy can be dissipated.

Consider a field configuration with energy E and helicity H. The energy has a lower bound set by the helicity:

$$E \geq \frac{|H|}{L}$$

where L is the system size. Energy above this bound can be dissipated, but the bound itself cannot be violated without changing the helicity.

In resistive MHD, helicity can be destroyed through reconnection, but the dissipation rate is slow:

$$\frac{dH}{dt} = -2\eta \int \vec{B} \cdot (\nabla \times \vec{B})\, dV$$

For large R_m, helicity is approximately conserved even in turbulent flows. This "selective decay" leads to the field relaxing toward a force-free configuration with $\nabla \times \vec{B} = \alpha \vec{B}$ for some constant α.

Helicity conservation may be important for understanding why disks sustain magnetic fields over long times. The MRI generates helical field structures; their helicity prevents complete dissipation.

7.8 The Parker Instability

Magnetic flux tubes in a gravitational field are buoyant. This is the Parker instability, first analyzed by Eugene Parker in 1966.

Consider a horizontal flux tube in a stratified atmosphere. The gas pressure inside the tube must balance the external pressure:

$$P_{\rm int} + \frac{B^2}{8\pi} = P_{\rm ext}$$

So $P_{\rm int} < P_{\rm ext}$—the interior is underpressured. If the tube is in thermal equilibrium with its surroundings (same temperature), then the interior density is lower:

$$\rho_{\rm int} = \frac{P_{\rm int}}{c_s^2} < \frac{P_{\rm ext}}{c_s^2} = \rho_{\rm ext}$$

The tube experiences a buoyancy force:

$$F_{\rm buoy} = (\rho_{\rm ext} - \rho_{\rm int})g \approx \frac{B^2}{8\pi c_s^2} g$$

But magnetic tension resists bending the flux tube. The stability criterion depends on the wavelength of the perturbation. For perturbations along the field with wavenumber k, the growth rate is:

$$\gamma^2 = \frac{g}{H}\left(\frac{v_A^2}{c_s^2}\right) - k^2 v_A^2$$

where H is the pressure scale height.

Instability ($\gamma^2 > 0$) occurs when

$$k < k_{\rm crit} = \frac{1}{H}\sqrt{\frac{gH}{c_s^2}} \cdot \frac{v_A}{c_s}$$

For $v_A \sim c_s$ (equipartition field) and $gH \sim c_s^2$ (hydrostatic):

$$k_{\rm crit} \sim \frac{1}{H}$$

So perturbations with wavelengths longer than the scale height are unstable. They grow on the dynamical timescale $\sim H/c_s$.

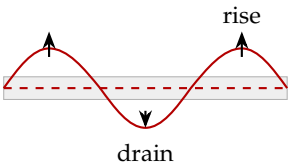

Figure 7.4: The Parker instability. A horizontal flux tube buckles; gas drains into the valleys, and buoyant crests rise.

In accretion disks, the Parker instability can transport magnetic flux from the dense midplane to the low-density corona. The rising flux loops may reconnect above the disk, releasing energy that heats the corona. This is one mechanism for creating the hot, X-ray-emitting coronae observed in accreting systems.

7.9 *Magnetic Reconnection: Detailed Physics*

We introduced reconnection earlier as the mechanism that dissipates magnetic energy. Let's examine it more carefully.

Consider two regions of oppositely directed magnetic field brought together. The field reverses across a thin current sheet. In ideal MHD, the field lines can't break—they just pile up, creating ever-stronger gradients. But in resistive MHD, the field can diffuse across the sheet, allowing lines to reconnect.

Sweet-Parker Reconnection:

The classic model has a steady-state current sheet of length L and thickness δ. Field lines enter from above and below, reconnect in the sheet, and exit sideways.

Mass conservation: plasma enters at velocity $v_{\rm in}$ over area $\sim L$; it exits at velocity $v_{\rm out}$ over area $\sim \delta$:

$$v_{\rm in} L \sim v_{\rm out} \delta$$

The exit velocity is the Alfvén speed (magnetic tension accelerates the plasma):

$$v_{\rm out} \sim v_A$$

The sheet thickness is set by resistive diffusion balancing advection:

$$\delta \sim \frac{\eta}{v_{\rm in}}$$

Combining these:

$$v_{\rm in} \sim v_A \sqrt{\frac{\eta}{v_A L}} = \frac{v_A}{\sqrt{R_m}}$$

For $R_m \sim 10^{12}$, this gives $v_{\rm in} \sim 10^{-6} v_A$—impossibly slow.

Fast Reconnection:

Several mechanisms can speed up reconnection:

Petschek reconnection: The current sheet is much shorter than L, surrounded by slow-mode shocks that process most of the flux. This gives $v_{\rm in} \sim v_A / \ln R_m$—much faster, but controversial whether it's stable.

Plasmoid instability: For $R_m > 10^4$, the Sweet-Parker sheet is unstable to tearing modes. It fragments into a chain of plasmoids (magnetic islands), dramatically increasing the reconnection rate to $v_{\rm in} \sim 0.01 v_A$, independent of R_m.

Collisionless reconnection: When the sheet thickness approaches the ion skin depth $d_i = c/\omega_{pi}$, collisionless effects (Hall term, electron inertia) decouple ions and electrons, enabling fast reconnection.

Observations and simulations suggest that astrophysical reconnection is fast, with $v_{\rm in} \sim 0.01$–$0.1 v_A$. This is fast enough to dissipate field energy on dynamical timescales.

7.10 Magnetic Stress and Angular Momentum Transport

The magnetic field transports angular momentum through Maxwell stress. Let's derive this carefully.

The magnetic force per unit volume is

$$\vec{f}_B = \frac{1}{4\pi}(\nabla \times \vec{B}) \times \vec{B} = \frac{1}{4\pi}\left[(\vec{B}\cdot\nabla)\vec{B} - \nabla\left(\frac{B^2}{2}\right)\right]$$

The first term is magnetic tension; the second is magnetic pressure gradient.

The angular momentum equation in cylindrical coordinates:

$$\rho \frac{\partial}{\partial t}(r v_\phi) + \rho \vec{v}\cdot\nabla(r v_\phi) = r f_{B,\phi} + \ldots$$

The ϕ-component of the magnetic force is

$$f_{B,\phi} = \frac{1}{4\pi}\left[B_r \frac{\partial B_\phi}{\partial r} + B_z \frac{\partial B_\phi}{\partial z} + \frac{B_r B_\phi}{r}\right]$$

The angular momentum flux has a magnetic part:

$$\dot{L}_{\rm mag} = \int r T^{\rm mag}_{r\phi} \cdot 2\pi r \, dz = \int \frac{r^2 B_r B_\phi}{4\pi} \cdot 2\pi \, dz$$

Averaging over a turbulent disk (denoted by $\langle\cdot\rangle$):

$$\langle T^{\rm mag}_{r\phi} \rangle = \frac{\langle B_r B_\phi \rangle}{4\pi}$$

This is the Maxwell stress. For outward angular momentum transport, we need $\langle B_r B_\phi \rangle < 0$ (since r decreases inward).

In MRI turbulence, this correlation is maintained by the instability mechanism: the radial displacement that couples to the shear creates correlated radial and toroidal field perturbations. Simulations consistently find $\langle B_r B_\phi \rangle < 0$, confirming outward angular momentum transport.

The effective viscosity parameter is

$$\alpha_{\rm mag} = \frac{-\langle B_r B_\phi \rangle}{4\pi P}$$

Typical values from simulations: $\alpha_{\rm mag} \sim 0.01$–$0.1$, depending on the net magnetic flux.

7.11 *The Magnetic Prandtl Number*

The magnetic Prandtl number is the ratio of viscosity to magnetic diffusivity:

$$Pm = \frac{\nu}{\eta}$$

In most fluids, $Pm \ll 1$ (magnetic diffusion is faster than viscous diffusion). In hot plasmas, $Pm \gg 1$ (viscosity is faster). Accretion disks span both regimes.

The Prandtl number matters for MRI turbulence because it affects the small-scale structure. At scales below the viscous cutoff, the velocity field is smooth but the magnetic field can still be turbulent. At scales below the resistive cutoff, the magnetic field is smooth.

For $Pm \gg 1$: the resistive scale is larger than the viscous scale. Small-scale kinetic energy can freely cascade to smaller scales, but magnetic energy is dissipated at larger scales. This can suppress the small-scale dynamo.

For $Pm \ll 1$: the viscous scale is larger than the resistive scale. Kinetic energy dissipates at larger scales, but magnetic energy can cascade to smaller scales. This can enhance the small-scale dynamo.

Simulations show that the saturated MRI state depends on Pm:

- For $Pm \gtrsim 1$: $\alpha \sim 0.01$–0.1, relatively independent of Pm
- For $Pm \ll 1$: α can be significantly lower
- For $Pm \gg 1$: α may be enhanced

This creates a potential problem for simulations: achieving realistic Pm values ($\sim 10^{-5}$ for protoplanetary disks, $\sim 10^{10}$ for hot accretion flows) is computationally impossible. The convergence of MRI simulations with Pm remains an active research topic.

7.12 The Magnetically Arrested Disk: Detailed Treatment

The MAD state deserves a more quantitative treatment. Let's derive the saturation flux and jet power.

Define the dimensionless magnetic flux parameter:

$$\phi = \frac{\Phi}{\sqrt{\dot{M} r_g c}}$$

where Φ is the magnetic flux threading the horizon.

In MAD, the magnetic pressure at the horizon balances the ram pressure of the accreting gas:

$$\frac{B^2}{8\pi} \sim \rho v_r^2$$

Using $\dot{M} = 4\pi r^2 \rho v_r$ with $r \sim r_g$ and $v_r \sim c$:

$$\rho \sim \frac{\dot{M}}{4\pi r_g^2 c}$$

$$B^2 \sim 8\pi\rho c^2 \sim \frac{2\dot{M}c}{r_g^2}$$

$$\Phi = \pi r_g^2 B \sim \pi r_g^2 \sqrt{\frac{2\dot{M}c}{r_g^2}} = \sqrt{2\pi^2 \dot{M} c r_g^2}$$

$$\phi_{\rm MAD} = \sqrt{2\pi^2} \approx 4.4$$

GRMHD simulations find $\phi_{\rm MAD} \approx 15$–$50$, depending on the black hole spin and field geometry. The discrepancy reflects simplifications in our estimate (we assumed spherical symmetry and steady state).

The jet power in MAD is:

$$P_{\rm jet} \approx \eta_{\rm jet} \dot{M} c^2$$

where $\eta_{\rm jet}$ is the jet efficiency. For a maximally spinning black hole in MAD:

$$\eta_{\rm jet} \approx 1.4 \left(\frac{\phi}{50}\right)^2 a^2$$

For $a = 0.99$ and $\phi = 50$:

$$\eta_{\rm jet} \approx 1.4$$

The jet power exceeds the rest-mass energy of the accreted material! This is possible because the energy comes from the black hole spin, not just the accretion flow.

Location	B (Gauss)	Notes
ISM	10^{-6}	Seed field
Outer disk (AGN, 1 pc)	0.1–1	Advected
Inner disk (AGN, $10r_g$)	10^3–10^4	MRI saturation
Near horizon (MAD)	10^4–10^5	$\beta \sim 1$
Outer disk (XRB, 1000 r_g)	10^3–10^4	Companion field
Inner disk (XRB, $10r_g$)	10^7–10^8	High density
Near horizon (MAD)	10^8–10^9	
Protoplanetary (10 AU)	0.01–0.1	Partially ionized
Protoplanetary (0.1 AU)	1–10	

Table 7.1: Characteristic magnetic field strengths in different accretion disk environments.

7.13 *The Magnetic Field Strength Ladder*

Let's compile the magnetic field strengths at different scales and in different systems:

These fields are generated and maintained dynamically by the MRI and related instabilities. Even starting from arbitrarily weak seed fields, the disk generates equipartition-level fields within a few tens of orbital periods.

8
How Jets Get Started

Jets are among the most spectacular phenomena in astrophysics. Narrow beams of plasma shoot outward from accreting black holes at speeds approaching the speed of light, maintaining their coherence across distances millions of times larger than the source that produced them. The jet from the galaxy M87 extends 5000 light-years—further than the distance from the Sun to the nearest stars—yet it originates from a region no larger than our solar system.

How do jets get started? What launches material off the disk and accelerates it to relativistic speeds? These questions have occupied theorists for decades. The answer, we now believe, involves magnetic fields in an essential way. But before we can understand how magnetic fields launch jets, we need to understand where the energy comes from.

8.1 Where Does the Energy Come From?

Jets carry enormous power. The jet from M87 has a mechanical luminosity of roughly 10^{44} erg/s, comparable to the total light output of a small galaxy. This energy must come from somewhere.

There are only two possible reservoirs: the gravitational energy released by accreting matter, and the rotational energy stored in the spinning black hole itself.

Let's estimate both.

Accretion Power

The gravitational energy released when mass $\dot{M}$ falls onto a black hole is

$$L_{\rm acc} = \eta \dot{M} c^2$$

where η is the radiative efficiency. For a thin disk around a non-spinning black hole, $\eta \approx 0.06$; for a maximally spinning hole, η can reach 0.42. A typical value is $\eta \sim 0.1$.

For a supermassive black hole accreting at $\dot{M} = 1 M_\odot/\text{yr}$ (high but not extreme):

$$L_{\text{acc}} \sim 0.1 \times (2 \times 10^{33}\ \text{g/yr}) \times (3 \times 10^{10})^2\ \text{cm}^2/\text{s}^2$$

Converting years to seconds (1 yr $\approx 3 \times 10^7$ s):

$$L_{\text{acc}} \sim 0.1 \times \frac{2 \times 10^{33}}{3 \times 10^7} \times 9 \times 10^{20} \sim 6 \times 10^{45}\ \text{erg/s}$$

This is more than enough to power the M87 jet. But there's another reservoir.

Black Hole Spin Energy

A spinning black hole stores rotational energy. How much?

The spin of a black hole is characterized by the dimensionless spin parameter a, which ranges from 0 (non-spinning) to 1 (maximally spinning). The rotational energy is

$$E_{\text{rot}} = f(a)\, Mc^2$$

where $f(a)$ ranges from 0 to about 0.29. For a maximally spinning black hole:

$$E_{\text{rot,max}} = 0.29\, Mc^2$$

Let's put in numbers for M87's black hole, which has mass $M \approx 6.5 \times 10^9 M_\odot$:

$$E_{\text{rot,max}} = 0.29 \times (6.5 \times 10^9) \times (2 \times 10^{33}) \times (9 \times 10^{20}) \approx 3 \times 10^{63}\ \text{erg}$$

This is an almost incomprehensibly large number. If we extracted this energy at the rate of 10^{44} erg/s (the M87 jet power), how long would it last?

$$t = \frac{3 \times 10^{63}}{10^{44}} = 3 \times 10^{19}\ \text{s} \approx 10^{12}\ \text{years}$$

A trillion years! The rotational energy of a supermassive black hole could power a jet for far longer than the current age of the universe.

But Can We Extract It?

Here's the puzzle. The rotational energy is locked inside the black hole, behind the event horizon. Nothing can escape from inside the horizon—that's the defining property of a black hole. So how can we possibly extract the spin energy?

You might say, "Well, you can't. It's inside a black hole. That's cheating." And at first glance, you'd be right to object. Physicists

spent years being skeptical of any claim to extract energy from a black hole. But nature, as it turns out, has found a loophole.

The answer lies in a remarkable property of rotating black holes: frame dragging. Near a spinning black hole, spacetime itself is dragged around by the rotation. This creates a region called the *ergosphere* where something extraordinary happens: you can extract energy from the black hole without ever crossing the horizon.

Understanding frame dragging is the key to understanding jet launching. Let's build up the picture carefully.

8.2 Frame Dragging: When Spacetime Flows

Near any massive object, spacetime is curved. Near a *rotating* massive object, spacetime is also dragged around by the rotation. This effect, called frame dragging or the Lense-Thirring effect, is a prediction of general relativity with no Newtonian analog.

An Intuitive Picture

Imagine spacetime as a viscous fluid, like honey. A rotating object immersed in honey drags the honey around with it—faster close to the object, slower farther away. Similarly, a rotating black hole drags spacetime around.

What does it mean for spacetime to be "dragged"? It means that the natural state of motion—what we'd call "at rest"—is actually rotating. An observer who tries to hover at a fixed position near a spinning black hole finds that they must fire rockets to resist the drag. If they turn off their rockets, they start orbiting, carried along by the rotating spacetime.

The dragging angular velocity at distance r from a black hole with spin parameter a is approximately

$$\omega_{\text{drag}} \approx \frac{2aGM}{c^2 r^3}$$

Close to the horizon, this becomes comparable to the black hole's own angular velocity.

The Ergosphere

Close enough to a spinning black hole, frame dragging becomes so strong that *nothing* can remain stationary—not even light. This region is called the ergosphere (from the Greek *ergon*, meaning work, because work can be extracted here).

The outer boundary of the ergosphere, called the static limit, occurs where the dragging velocity equals the speed of light. At the

equator:

$$r_{\text{static}} = 2GM/c^2 = 2r_g$$

independent of spin. The ergosphere extends from this radius down to the event horizon.

Inside the ergosphere, an observer falling freely (with no rocket thrust) is inevitably dragged around the black hole. They cannot hover, cannot orbit retrograde (against the spin), cannot even fall straight in. The rotation of spacetime is inescapable.

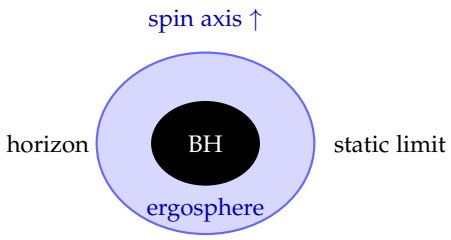

Figure 8.1: The ergosphere of a spinning black hole (view from above the spin axis). Between the static limit and the horizon, nothing can remain at rest.

Negative Energy States

Here's where things get strange. In general relativity, the energy of a particle depends on who measures it. A distant observer assigns energy based on their notion of time. But inside the ergosphere, the distant observer's "time direction" becomes partly spatial (technically: the Killing vector ∂_t becomes spacelike).

The consequence is that a particle inside the ergosphere can have *negative energy* as measured by a distant observer, while still having positive energy as measured locally. This isn't a mathematical trick—it's physically meaningful. If such a negative-energy particle falls into the black hole, the black hole's total mass decreases.

This is the key insight: the ergosphere allows processes where a black hole can lose mass and angular momentum by swallowing particles that carry negative energy.

8.3 The Penrose Process: Proof of Concept

Roger Penrose discovered in 1969 that the ergosphere enables energy extraction from a spinning black hole. His mechanism, now called the Penrose process, works as follows:

1. A particle enters the ergosphere with energy $E_0 > 0$ (as measured from infinity).

2. Inside the ergosphere, it splits into two particles.

3. One particle (call it particle 1) is arranged to have negative energy $E_1 < 0$.

4. The other particle (particle 2) escapes with energy $E_2 = E_0 - E_1 > E_0$.

When particle 1 falls into the black hole, it reduces the hole's mass by $|E_1|/c^2$ and reduces its angular momentum by L_1. The energy gained by particle 2 comes from the black hole's rotation.

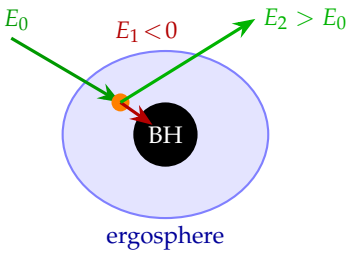

Figure 8.2: The Penrose process. A particle enters the ergosphere, splits, and one fragment with negative energy falls in while the other escapes with more energy.

Efficiency Limits

The Penrose process has a maximum efficiency. For a single particle decay, the energy gain is limited to about 20% of the particle's rest mass:

$$\frac{E_2 - E_0}{m_0 c^2} \lesssim 0.2$$

The "collisional Penrose process," where two particles collide in the ergosphere, can do better—up to $(\sqrt{2} - 1) \approx 41\%$ efficiency in idealized scenarios.

Why Penrose Isn't Practical for Jets

The Penrose process is important conceptually—it proves that black hole spin energy can be extracted. But it's not how jets work. The process requires carefully arranged particle trajectories, and the extracted energy emerges as randomly-directed particles, not a collimated jet.

What we need is a mechanism that extracts energy *continuously* and channels it into a *directed outflow*. This is where magnetic fields enter the story.

8.4 The Blandford-Znajek Mechanism: Overview

In 1977, Roger Blandford and Roman Znajek showed that magnetic fields threading a spinning black hole can extract energy continuously, powering a relativistic outflow. This is now called the Blandford-Znajek (BZ) mechanism, and it's believed to be the primary power source for jets from AGN and microquasars.

The Basic Idea

Imagine magnetic field lines threading the event horizon of a spinning black hole. (We'll see later how this field structure arises.) The rotating horizon drags the field lines, twisting them by frame dragging.

The twist propagates outward along the field lines as an Alfvén wave. This wave carries energy and angular momentum away from the black hole. The energy comes from the hole's rotation; the black hole gradually spins down.

This is an electromagnetic analog of the Penrose process. Instead of arranging particle trajectories, we let the magnetic field do the work. The field extracts energy from the spinning spacetime and transmits it outward as electromagnetic (Poynting) flux.

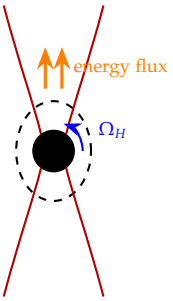

Figure 8.3: The Blandford-Znajek mechanism. Field lines threading the rotating horizon are twisted by frame dragging, generating an outward Poynting flux.

An Analogy: The Faraday Disk

A helpful analogy is the Faraday disk (or homopolar generator). In this device, a conducting disk rotates in a magnetic field. The motion induces an EMF between the axis and rim of the disk, which drives current through an external circuit.

The spinning black hole plays the role of the rotating disk. The magnetic field threading the horizon is twisted by the rotation, inducing an effective EMF. This EMF drives currents through the external "circuit"—the magnetosphere surrounding the black hole.

The analogy isn't perfect (the black hole isn't really a conductor), but it captures the essential physics: rotation plus magnetic field equals power generation.

8.5 The Membrane Paradigm: An Intuitive Picture

The mathematics of the BZ mechanism involves general relativistic electrodynamics in curved spacetime—not for the faint of heart. Fortunately, Kip Thorne and collaborators developed an intuitive picture called the "membrane paradigm" that makes the physics transparent.

The Horizon as a Conductor

In the membrane paradigm, we pretend the event horizon is a physical membrane with electrical properties. This membrane has surface resistivity

$$R_H = \frac{4\pi}{c} \approx 377 \text{ ohms}$$

This is exactly the "impedance of free space"—the same quantity that appears in antenna theory and electromagnetic wave propagation. It's a remarkable coincidence (or perhaps not a coincidence) that makes the physics work out elegantly.

If this seems like a strange way to think about a black hole, you're right. The membrane paradigm doesn't claim the horizon really is a conducting surface. It claims that if you pretend it is, you get the right answers. This is the kind of trick physicists love: a wrong picture that gives correct results is often more useful than a correct picture that's too complicated to calculate with.

The Circuit Analogy

Think of the system as an electrical circuit:

- The spinning, magnetized horizon acts as a **battery** (or generator)

with EMF

$$\mathcal{E} = \frac{\Omega_H \Phi}{2\pi c}$$

where Ω_H is the horizon angular velocity and Φ is the magnetic flux threading one hemisphere.

- The horizon itself acts as an internal **resistor** with resistance R_H.
- The external region (magnetosphere, jet) acts as the **load** with resistance R_L.

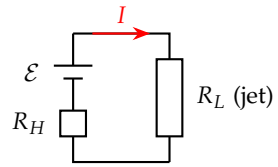

Figure 8.4: Circuit analog of the BZ mechanism. The spinning horizon acts as a battery with internal resistance; the jet is the load.

The current in the circuit is

$$I = \frac{\mathcal{E}}{R_H + R_L}$$

The power delivered to the load (the jet) is

$$P_{\text{jet}} = I^2 R_L = \frac{\mathcal{E}^2 R_L}{(R_H + R_L)^2}$$

This is maximized when $R_L = R_H$ (impedance matching), giving

$$P_{\max} = \frac{\mathcal{E}^2}{4R_H}$$

Physical Interpretation

What do the "current" and "resistance" really mean physically?

The current corresponds to electric currents flowing in the magnetosphere—charges moving along and across field lines. These currents are carried by electron-positron pairs created in the strong fields near the hole.

The horizon "resistance" represents the rate at which electromagnetic energy is absorbed by the horizon. If all the power were absorbed (infinite load resistance, no external circuit), the horizon would still dissipate energy through its internal resistance. The extracted power is the difference between what the "battery" generates and what the horizon absorbs.

The load "resistance" represents the external astrophysical environment. A good conductor (low R_L) shorts out the battery. A poor conductor (high R_L) draws little current. Maximum power extraction requires matching the load to the source.

8.6 The BZ Power Formula

Let's derive the BZ power more carefully. The electromagnetic energy flux (Poynting flux) extracts energy from the black hole and carries

it outward. For a slowly rotating black hole with a split-monopole magnetic field (the simplest analytic model), the power is

$$P_{\rm BZ} = \frac{1}{6\pi c}\Omega_F(\Omega_H - \Omega_F)\Phi^2$$

where Ω_H is the horizon angular velocity, Ω_F is the angular velocity of the magnetic field lines, and Φ is the magnetic flux.

This is maximized when $\Omega_F = \Omega_H/2$, giving

$$P_{\rm BZ,max} = \frac{\Omega_H^2\Phi^2}{24\pi c}$$

For a more general field geometry, the scaling is

$$P_{\rm BZ} \propto a^2\Phi^2/M^2$$

where a is the spin parameter and M is the black hole mass.

Numerical Estimate

Let's put in numbers. For a supermassive black hole:

$$P_{\rm BZ} \approx 10^{45}\left(\frac{a}{M}\right)^2\left(\frac{\Phi}{10^{32}\ {\rm G\ cm^2}}\right)^2\left(\frac{M}{10^9 M_\odot}\right)^{-2}\ {\rm erg/s}$$

This is comparable to the M87 jet power for reasonable parameters. The BZ mechanism can explain the most powerful jets in the universe.

The Horizon Angular Velocity

What is Ω_H? For a Kerr black hole:

$$\Omega_H = \frac{ac}{2Mr_H} = \frac{a}{r_H^2 + a^2}\cdot\frac{c}{r_g}$$

where $r_H = r_g(1 + \sqrt{1-a^2})$ is the horizon radius and $r_g = GM/c^2$.

For a maximally spinning hole ($a \to M$), $\Omega_H \to c/(2r_g)$—half the speed of light divided by the gravitational radius. For M87, this corresponds to a period of about one day.

8.7 Magnetic Flux: The Key Ingredient

The BZ power depends quadratically on magnetic flux. Where does this flux come from, and how much can thread the horizon?

Flux Supply by the Disk

Magnetic flux is supplied by the accretion disk. As magnetized gas spirals inward, it carries flux toward the black hole. The flux accumulates near the horizon.

But flux can't accumulate indefinitely. Eventually, the magnetic pressure at the horizon becomes so strong that it resists further accretion. This sets a maximum flux—the "magnetically arrested disk" or MAD limit.

The MAD Limit

At the MAD limit, magnetic pressure balances the ram pressure of infalling gas:

$$\frac{B^2}{8\pi} \sim \rho v_r^2 \sim \frac{\dot{M}}{4\pi r_H^2} \cdot c$$

where we've used $v_r \sim c$ near the horizon.

Solving for the maximum field strength:

$$B_{\max} \sim \sqrt{\frac{2\dot{M}c}{r_H^2}}$$

The maximum flux is then

$$\Phi_{\max} \sim B_{\max} \cdot \pi r_H^2 \sim \pi r_H \sqrt{2\dot{M}c}$$

The Dimensionless Flux Parameter

It's useful to define a dimensionless flux parameter:

$$\phi = \frac{\Phi}{\sqrt{\dot{M}c \cdot r_g^2}}$$

Numerical simulations show that the MAD state corresponds to $\phi \approx 50$. Below this value, accretion proceeds smoothly. At $\phi \approx 50$, the field becomes strong enough to temporarily halt accretion, leading to episodes of flux eruption and chaotic variability.

MAD Jet Power

Substituting the MAD flux into the BZ formula:

$$P_{\text{BZ,MAD}} \sim \frac{a^2}{M^2} \cdot \dot{M}c \cdot r_g^2 \sim a^2 \dot{M}c^2$$

For $a \sim 1$, the BZ power in the MAD state approaches $\dot{M}c^2$—the maximum possible efficiency! The jet extracts energy from the

black hole spin at a rate comparable to the rest-mass energy of the accreting material.

This remarkable result—confirmed by numerical simulations—means that rapidly spinning black holes in the MAD state can have jet powers exceeding their accretion luminosities. The black hole acts as a flywheel, storing rotational energy that the jet can tap.

8.8 Force-Free Electrodynamics

The BZ mechanism operates in a regime where the electromagnetic field dominates over matter. The magnetic energy density far exceeds the plasma rest-mass energy density:

$$\frac{B^2}{8\pi} \gg \rho c^2$$

In this limit, we can neglect the inertia of the plasma and treat the field as "force-free." The plasma adjusts instantaneously to whatever configuration the field requires.

The Force-Free Condition

The force-free condition states that electromagnetic forces on the plasma vanish:

$$\rho_e \vec{E} + \frac{\vec{J} \times \vec{B}}{c} = 0$$

where ρ_e is the charge density and $\vec{J}$ is the current density.

This doesn't mean there are no fields or currents—it means the electric and magnetic forces exactly balance. The plasma provides whatever charges and currents are needed to maintain this balance.

Maintaining Force-Free: Pair Production

Where do the charges come from? Near a black hole, the required charge density is

$$\rho_e = -\frac{\Omega_F}{2\pi c} B_z$$

This is called the "Goldreich-Julian" density, named after its analog in pulsar magnetospheres.

In number density:

$$n_{GJ} \sim \frac{\Omega_F B}{2\pi e c} \sim 10^{-2} \left(\frac{B}{10^4\ \mathrm{G}}\right) \left(\frac{\Omega_F}{10^{-4}\ \mathrm{s}^{-1}}\right)\ \mathrm{cm}^{-3}$$

This is a very low density—practically a vacuum by laboratory standards. But it must be maintained everywhere in the magnetosphere.

If the density falls below n_{GJ}, electric fields parallel to $\vec{B}$ develop. These "gaps" accelerate particles to extreme energies. The accelerated particles emit gamma rays, which pair-produce on ambient photons or the magnetic field itself:

$$\gamma + \gamma \to e^+ + e^-, \qquad \gamma + B \to e^+ + e^-$$

The newly created pairs fill the gap, restoring the force-free condition. This process is self-regulating: gaps form where needed and are immediately filled by pair cascades.

8.9 *The Blandford-Payne Mechanism*

The BZ mechanism extracts energy from the black hole. A second mechanism—the Blandford-Payne (BP) mechanism—extracts energy from the accretion disk.

Magnetocentrifugal Launching

In the BP mechanism, magnetic field lines anchored in the rotating disk act like rigid wires. Material attached to a field line is flung outward by centrifugal force as the disk rotates.

The physics is familiar: think of a bead on a rotating wire. If the wire is vertical, the bead just sits there. But if the wire is tilted outward, the bead slides along it, driven by the centrifugal force.

For disk material, the "wire" is a magnetic field line. If the field line is inclined outward from the vertical (toward larger radii at greater heights), the centrifugal force has a component along the field that accelerates material away from the disk.

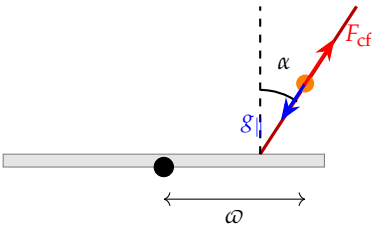

Figure 8.5: Magnetocentrifugal launching. A gas parcel on an inclined field line feels forces projected along the field: gravity pulls it back ($g_\parallel$), while centrifugal force pushes it outward ($F_{\rm cf}$). If the field is tilted enough, $F_{\rm cf} > g_\parallel$ and material is launched.

The Critical Angle

Not any inclination works. The field line must be tilted far enough that centrifugal force wins over gravity. Let's derive the critical angle.

Consider a particle at the disk surface at radius r_0, attached to a field line at angle α from the vertical. In the rotating frame (angular velocity $\Omega = \sqrt{GM/r_0^3}$), the effective potential is

$$\Phi_{\rm eff} = -\frac{GM}{r} - \frac{1}{2}\Omega^2 \varpi^2$$

where ϖ is the cylindrical radius (distance from the rotation axis).

Along a straight field line:

$$r = r_0 + s\cos\alpha, \qquad \varpi = r_0 + s\sin\alpha$$

where s is the distance along the field from the disk surface.

The force along the field is $F = -d\Phi_{\text{eff}}/ds$. At $s = 0$:

$$F(0) = \frac{GM}{r_0^2}(\sin\alpha - \cos\alpha)$$

This is positive (outward) when $\tan\alpha > 1$, i.e., $\alpha > 45$ from vertical.

But mere launching isn't enough. The particle must escape to infinity without encountering a potential barrier. This requires the potential to have no maximum along the field line, which gives a more stringent condition:

$$\tan^2\alpha > 2 \qquad \Rightarrow \qquad \alpha > \arctan(\sqrt{2}) \approx 54.7$$

This corresponds to an angle of about 35 from the disk plane. Field lines must be inclined at least this much for effective wind launching.

The famous "30Âř criterion" from Blandford & Payne's original paper comes from a slightly different analysis with specific assumptions about the field geometry.

Angular Momentum Extraction

A key function of the BP wind is to extract angular momentum from the disk. This allows accretion to proceed without requiring internal viscosity.

The wind carries angular momentum in two forms:

Material angular momentum: The gas itself carries $\dot{M}_w \varpi^2 \Omega$, where $\dot{M}_w$ is the mass loss rate.

Magnetic angular momentum: The twisted field exerts a torque that extracts additional angular momentum from the disk.

The total specific angular momentum carried by the wind is

$$L = \varpi^2\Omega - \frac{\varpi B_\phi}{\eta}$$

where $\eta = 4\pi\rho v_p / B_p$ is the mass-to-flux ratio and B_ϕ is the toroidal field component.

At the Alfvén surface (where the flow speed equals the Alfvén speed), the magnetic and material contributions are equal. This defines the Alfvén radius ϖ_A. The wind extracts angular momentum as if the disk extended to ϖ_A—even though material only leaves the disk at $r_0 \ll \varpi_A$.

This "magnetic lever arm" is why magnetized winds are so effective at driving accretion. A small mass loss rate can extract the angular momentum of a much larger accretion rate.

8.10 BZ vs BP: Which Dominates?

We now have two jet-launching mechanisms: BZ (powered by black hole spin) and BP (powered by disk rotation). When does each dominate?

The Power Comparison

BZ power scales as:

$$P_{\rm BZ} \propto a^2 \Phi^2 / M^2$$

BP power scales as:

$$P_{\rm BP} \propto \Phi^2 \Omega_{\rm disk}^2 \propto \Phi^2 / r^3$$

For field lines anchored at the innermost stable circular orbit (ISCO), $r \sim r_g$, so $P_{\rm BP} \propto \Phi^2 / r_g^3 \propto \Phi^2 / M^3$.

Comparing:

$$\frac{P_{\rm BZ}}{P_{\rm BP}} \propto a^2 M$$

For rapidly spinning ($a \sim 1$) supermassive black holes, BZ dominates. For slowly spinning or stellar-mass black holes, BP may be more important.

The Spine-Sheath Structure

In reality, both mechanisms operate simultaneously, producing a layered jet structure:

Inner spine: Field lines threading the black hole horizon, powered by BZ. Highly relativistic ($\Gamma \sim 10 - 100$), magnetically dominated, carries most of the energy.

Outer sheath: Field lines anchored in the inner disk, powered by BP. Mildly relativistic ($\Gamma \sim 2 - 5$), carries significant mass.

Disk wind: Slower outflow from larger disk radii. Non-relativistic, wide-angle.

This structure has been confirmed by numerical simulations and is consistent with VLBI observations of nearby jets.

8.11 What We Still Don't Understand

The basic physics of jet launching is well established. But important questions remain:

How does flux reach the hole? The disk must transport magnetic flux inward to feed the BZ mechanism. But MRI turbulence tends to diffuse flux outward. How some systems achieve MAD while others don't is unclear.

Why do jets turn on and off? X-ray binaries transition between jet-loud (hard) and jet-quiet (soft) states. The transitions are sharp and reproducible, suggesting a physical threshold—but what sets it?

What determines jet composition? Are jets mostly electron-positron pairs (light and easy to accelerate) or electron-proton plasma (heavy and hard to accelerate)? The loading process is poorly understood.

These questions connect jet launching to the broader physics of accretion flows. The next chapter explores what happens after the jet is launched: how it accelerates, collimates, and maintains its structure across vast distances.

9
Jet Structure and Dynamics

We've launched a jet. The BZ mechanism extracts energy from a spinning black hole; the BP mechanism taps the disk's rotation. Either way, Poynting flux streams outward along open magnetic field lines, carrying plasma with it. But what we've launched isn't really a jet yet. It's slow. It's wide. Most of its energy is in the magnetic field, not in bulk motion. The jets we observe are something else entirely—relativistic needles of plasma, focused to opening angles of a few degrees, plunging through intergalactic space for millions of years without falling apart.

How does a magnetically-dominated wind become a relativistic jet? This is the problem we face in this chapter. The answer involves beautiful physics: the interplay of magnetic tension, relativistic inertia, and plasma instabilities. But to appreciate the answer, we first need to understand why the problem is hard.

9.1 What Would We Naively Expect?

Let's follow our newly-launched outflow and ask: what should happen to it?

The flow starts near the black hole or inner disk, where the magnetic field is strong. As it moves outward, it expands. The field lines spread apart. The magnetic field strength drops, roughly as $B \propto 1/r^2$ for a radially-expanding flow. The magnetic energy density drops as $B^2 \propto 1/r^4$.

You might think this is good news for acceleration. The magnetic field does work on the plasma, pushing it outward. As the field weakens, doesn't its energy go into speeding up the flow?

In a sense, yes. But there's a problem. In a perfectly conducting plasma—the "ideal MHD" limit that usually applies—the magnetic field lines are frozen into the plasma. They move with it. When the plasma expands, the field lines stretch. When they stretch, tension tries to pull them back. The plasma does work against this tension,

which takes away some of the energy it received.

These effects partially cancel. The field pushes the plasma, but the plasma also stretches the field. The net acceleration is much weaker than you'd naively expect.

Let me be more precise. Define the magnetization parameter

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

as the ratio of magnetic energy density to rest-mass energy density. When $\sigma \gg 1$, the energy is mostly in the field—we call this "magnetically dominated." When $\sigma \ll 1$, the energy is mostly in bulk motion—"kinetically dominated."

Near the base of a jet, σ is typically 10 to 100 or higher. The observed jets have Lorentz factors $\Gamma \sim 10$ to 50. If all the magnetic energy converted to kinetic energy, we'd expect[1]

$$\Gamma \sim \sigma_0 \sim 10\text{–}100$$

[1] If the initial energy per unit rest mass is $\sigma_0 c^2$, and this all becomes kinetic energy $(\Gamma - 1)c^2$, then $\Gamma \sim \sigma_0$.

That seems fine. But ideal MHD acceleration doesn't work this efficiently.

9.2 *Why Acceleration Is Harder Than It Looks*

In ideal MHD, conservation laws constrain what can happen. The magnetic flux through any surface moving with the plasma is conserved (that's what "frozen-in" means). Angular momentum is conserved. Energy is conserved.

These conservation laws are powerful, but they also limit how energy can be redistributed. In particular, they prevent all the magnetic energy from converting to kinetic energy.

Here's a way to see why. Consider a parcel of plasma moving outward along a magnetic flux tube. The tube expands, so the cross-sectional area A grows. Magnetic flux conservation says $BA = \text{constant}$, so $B \propto 1/A$. If the tube expands as $A \propto r^2$, then $B \propto 1/r^2$, and the magnetic energy density goes as $B^2 \propto 1/r^4$.

But wait—the kinetic energy density also changes as the plasma expands. If the plasma accelerates to Lorentz factor Γ, its kinetic energy per unit volume goes as $\Gamma \rho c^2$. And if mass is conserved, $\rho A v = \text{constant}$, where v is the flow speed. For relativistic flow, $v \approx c$, so $\rho \propto 1/A \propto 1/r^2$.

So both the magnetic and kinetic energy densities drop as the flow expands. The question is which drops faster. If the magnetic energy drops faster, it gets converted to kinetic energy. If they drop at the same rate, there's no net conversion.

In ideal MHD, the conversion is inefficient because the magnetic field tension "holds on" to some of the energy. The detailed calculation shows that the terminal Lorentz factor scales as

$$\Gamma_\infty \sim \sigma_0^{1/3}$$

rather than $\Gamma_\infty \sim \sigma_0$.[2]

For $\sigma_0 = 100$, this gives $\Gamma_\infty \sim 5$. That's much less than the $\Gamma \sim 10$–50 we observe in AGN jets, and far less than the $\Gamma > 100$ required for gamma-ray bursts.

This is the "σ problem." Ideal MHD can't efficiently convert magnetic energy to kinetic energy. Something else must be happening.

The sigma problem emerged gradually through the 1990s and 2000s as theorists developed increasingly sophisticated models of jet acceleration. The pioneers—Michel, Li, Begelman, and others—showed that ideal MHD jets accelerate slowly. By the mid-2000s, it was clear that something beyond ideal MHD was needed. The leading candidate was magnetic reconnection, but showing that reconnection actually happens in jets, and that it happens fast enough to matter, required a new generation of numerical simulations. The problem is still not fully solved.

You might say, "Maybe the observations are wrong. Maybe jets aren't really that fast." But we've measured the speeds carefully—superluminal motion, time delays in gamma-ray flares, all point to the same answer. The jets are fast. The σ problem is real.

[2] This result comes from requiring the flow to pass smoothly through the fast magnetosonic critical surface. The mathematics is involved, but the physical content is that the field manages to hold onto most of its energy.

9.3 *An Analogy: Stretching a Rubber Band*

Before we tackle the σ problem's solution, let's build intuition with an analogy.

Imagine you're holding a rubber band and accelerating a ball attached to it. You pull the ball forward; the rubber band stretches and pulls back. The ball speeds up, but not as fast as if there were no rubber band. Some of the work you do goes into stretching the band, not accelerating the ball.

Now imagine the rubber band is infinitely strong—it can stretch forever without breaking. No matter how hard you pull, the band always exerts tension, always taking some of your work. The ball accelerates, but it never reaches the speed it would without the band.

This is the ideal MHD limit. The magnetic field is like an infinitely strong rubber band. It always exerts tension. It always takes some of the energy. The plasma accelerates, but not as efficiently as it would without the field.

But real rubber bands break. And real magnetic fields can "break" too, through a process called reconnection. When the field topology

changes, energy that was stored in field tension gets released suddenly as heat and particle motion. The rubber band snaps, and the ball shoots forward.

Of course, magnetic field lines don't literally snap—they're not physical strings. What happens is subtler: the topology changes, field lines that were separate become connected, and energy that was locked in the field configuration gets released. But "the rubber band snaps" captures the essential physics, and that's what analogies are for.

9.4 *Magnetic Reconnection: Breaking the Field*

Reconnection happens when magnetic field lines of opposite polarity are pushed together. Imagine two field lines pointing in opposite directions, brought close together. At the interface, the field must change direction over a short distance. In ideal MHD, this is fine—the field just makes a sharp bend. But in any real plasma, the finite conductivity allows the field to "slip" through the plasma at the interface. The field lines break and reconnect in a different topology.

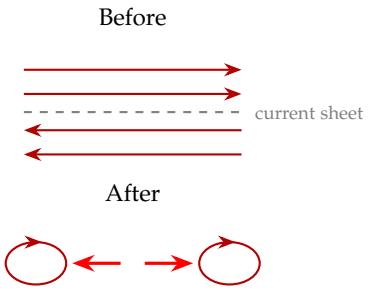

Figure 9.1: Magnetic reconnection. Top: antiparallel field lines meet at a current sheet. Bottom: after reconnection, field lines form loops that are expelled sideways, releasing magnetic energy.

When reconnection happens, the stored magnetic energy is released. Some goes into heating the plasma. Some goes into accelerating particles to high energies. In relativistic reconnection—which occurs when the magnetic energy density exceeds the rest-mass energy density—the outflow from the reconnection region can reach relativistic speeds.

This is how jets may solve the σ problem. Ideal MHD acceleration gets the plasma to modest speeds, $\Gamma \sim$ a few. Then instabilities in the flow bring oppositely-directed field lines together. Reconnection releases the remaining magnetic energy, boosting the plasma to $\Gamma \sim$ 10–100.

What instabilities trigger reconnection? The main candidate is the kink instability, which we'll discuss later. But first, let's understand another crucial aspect of jet structure: collimation.

9.5 *Why Are Jets Narrow?*

Look at an image of the M87 jet. It's remarkably narrow—a thin beam stretching across 5000 light-years, yet only a few hundred light-years wide. The opening angle is about 5.[3]

Why should an outflow stay narrow? Without some focusing mechanism, you'd expect it to expand freely. A wind blown from a star, for instance, spreads into a sphere. Even a supersonic wind spreads conically. What keeps a jet from doing the same?

[3] At smaller scales, near the black hole, the opening angle is larger—the jet collimates as it propagates.

The answer is magnetic—specifically, magnetic tension in the toroidal (wrapped-around) component of the field.

9.6 *Deriving Magnetic Collimation*

Let's work this out from first principles. As the disk rotates beneath the jet, it drags the magnetic field lines around with it. A field line that was initially vertical gets wound into a helix. Near the jet axis, this creates a toroidal field component B_ϕ that wraps around like coils of rope.

Now, what forces does this toroidal field exert? Magnetic fields exert both pressure (pushing perpendicular to the field lines) and tension (pulling along the field lines). For a curved field line, the tension produces a net force toward the center of curvature—just like the tension in a stretched string.

Consider a toroidal field line at cylindrical radius ϖ from the jet axis. The field line is a circle of radius ϖ. Its tension exerts a force per unit length of[4]

$$\frac{B_\phi^2}{4\pi\varpi}$$

directed toward the axis. This is the "hoop stress"—the magnetic analog of the tension that keeps a barrel's hoops from flying apart.

The hoop stress squeezes the jet toward its axis. The stronger the toroidal field, the stronger the squeeze. And as the jet propagates, the disk keeps winding the field, building up B_ϕ. The collimation force grows.

For the jet to be in equilibrium transverse to its axis, the hoop stress must be balanced by outward pressure:

$$\frac{\partial P}{\partial \varpi} = -\frac{B_\phi^2}{4\pi\varpi}$$

This determines the jet's radial structure. The pressure is highest on-axis and drops toward the edge, balanced by the magnetic squeeze.

Let's put in some numbers. For M87, the toroidal field near the collimation region might be $B_\phi \sim 0.1$ G (estimated from polarization observations). At a radius $\varpi \sim 100\, r_g \sim 10^{17}$ cm from the axis:[5]

$$\frac{B_\phi^2}{4\pi\varpi} \sim \frac{(0.1)^2}{4\pi \times 10^{17}} \sim 10^{-21}\ \text{dyn/cm}^3$$

This seems tiny, but it acts over the entire cross-section of the jet, and it acts continuously as the jet propagates. Over thousands of gravitational radii, it's enough to focus a wide outflow into a narrow beam.

[4] The magnetic tension force per unit volume is $(\vec{B}\cdot\nabla)\vec{B}/4\pi$. For a circular field line at radius ϖ, this gives an inward force $B^2/(4\pi\varpi)$.

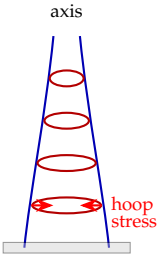

Figure 9.2: Toroidal field lines wrapped around the jet exert inward hoop stress, collimating the flow toward the axis.

[5] For M87's black hole with $M \approx 6.5 \times 10^9 M_\odot$, we have $r_g \approx 10^{15}$ cm.

9.7 Where Does Collimation Happen?

Collimation isn't instantaneous. Near the launching region, the jet is wide and the toroidal field is still building up. As the jet propagates, the field winds tighter and the hoop stress grows. Eventually, the jet reaches its asymptotic, nearly-cylindrical shape.

The characteristic scale for this process is the AlfvÃl'n radius ϖ_A—the cylindrical radius where the poloidal flow speed equals the AlfvÃl'n speed:

$$v_p = v_A = \frac{B_p}{\sqrt{4\pi\rho}}$$

Why is the AlfvÃl'n radius special? Below it, the flow is "sub-AlfvÃl'nic"—slower than magnetic disturbances can propagate. The magnetic field can adjust to the flow, communicating information back to the disk. Above it, the flow is "super-AlfvÃl'nic"—the field is swept along and can't send information backward.

At the AlfvÃl'n surface, the toroidal and poloidal field components become comparable, and the collimation is essentially complete. For typical jets:

$$\varpi_A \sim 100\text{–}1000\, r_g$$

The jet achieves its narrow, collimated structure within about 1000 gravitational radii of the black hole.

9.8 Critical Surfaces: Causal Boundaries

The AlfvÃl'n surface is one of three "critical surfaces" in the jet—places where the flow speed equals a characteristic wave speed. These surfaces aren't just mathematical curiosities. They're causal boundaries that determine what information can propagate where.

Think of it this way. In a flowing medium, waves propagate both upstream and downstream. But if the flow is faster than the wave speed, even upstream-propagating waves get swept downstream. The wave can't outrun the flow. This means no information from downstream can reach upstream—the regions are causally disconnected.

Three Wave Speeds

In magnetized plasma, there are three types of waves with different speeds:

AlfvÃl'n waves oscillate the magnetic field sideways, propagating at $v_A = B/\sqrt{4\pi\rho}$. They're like waves on a guitar string, with magnetic tension providing the restoring force.

Slow magnetosonic waves combine pressure and magnetic forces, propagating slower than both sound and AlfvÃl'n speeds.

Fast magnetosonic waves also combine pressure and magnetic forces, but propagate faster than both.

Each wave type has its critical surface, where the flow speed equals that wave's speed.

The Physical Meaning

Below the slow critical surface, all three wave types can propagate upstream. The flow is fully connected to the disk. Any change in downstream conditions can, in principle, affect the launching region.

Between the slow and AlfvÃl'n surfaces, slow waves are swept downstream, but AlfvÃl'n and fast waves can still propagate up. The flow is partially connected.

Above the fast surface, no MHD wave can propagate upstream. The flow is completely disconnected from its source. Whatever happens downstream—the jet hits something, or dissipates, or gets deflected—that information can never reach the launching region.

This is why the critical surfaces matter physically. They tell us which part of the jet "knows about" the disk and which part is independent. The properties of the jet—its mass loss rate, its energy flux—are determined at the critical surfaces, not downstream.

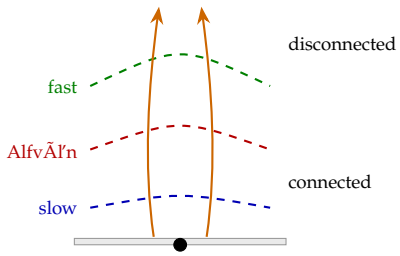

Figure 9.3: The three critical surfaces in a jet. Below the slow surface, all information can propagate back to the disk. Above the fast surface, the jet is causally disconnected from its source.

Regularity Conditions

At each critical surface, the MHD equations develop a mathematical singularity—the coefficients vanish, and the equations become indeterminate. For the solution to pass smoothly through the singularity, specific conditions must be satisfied.

This is exactly analogous to stellar winds. In a stellar wind, the flow must pass through the sonic point where the speed equals the sound speed. This requirement fixes the mass loss rate. In jets, the flow must pass through all three critical surfaces, and the regularity conditions at each one constrain the solution.

9.9 Angular Momentum: The Magnetic Lever Arm

Jets don't just carry energy—they carry angular momentum. This is essential for accretion. If the disk couldn't get rid of angular momentum, material couldn't spiral inward. The jet provides an escape route.

How much angular momentum does a jet carry? More than you might expect, thanks to the magnetic field.

Consider a plasma parcel at radius ϖ rotating at angular velocity Ω. Its angular momentum per unit mass is $\ell = \varpi^2\Omega$. If this parcel flows outward in the jet, it carries this angular momentum with it.

But the magnetic field carries angular momentum too. A twisted field (with $B_\phi \neq 0$) exerts a torque through Maxwell stresses. The magnetic angular momentum flux per unit mass is[6]

$$\ell_{\rm mag} = -\frac{\varpi B_\phi}{4\pi\rho v_p}$$

where v_p is the poloidal (along-the-jet) velocity.

At the AlfvÃl'n surface, these two contributions are equal—the material and magnetic angular momentum fluxes match. This defines the AlfvÃl'n radius ϖ_A. The total specific angular momentum is

$$L = \Omega \varpi_A^2$$

Here's the remarkable thing: the jet extracts angular momentum from the disk as if solid-body rotation extended all the way to ϖ_A, even though the material is launched from much closer in. If ϖ_A is three times the launching radius, the jet extracts *nine times* as much angular momentum as the material alone would carry.

This is the "magnetic lever arm." The field acts like a rigid rod extending from the disk to the AlfvÃl'n surface, giving the jet enormous leverage for extracting angular momentum. This is why magnetized winds are so effective at driving accretion.

[6] The Maxwell stress tensor gives a ϕ-component of momentum flux $\sim B_\phi B_p/4\pi$, which corresponds to angular momentum flux $\varpi \cdot B_\phi B_p/4\pi$.

9.10 *Jet Stability: Why Don't Jets Fall Apart?*

The M87 jet remains coherent over 5000 light-years—about twenty million times the size of the black hole that launched it. Why doesn't it fall apart?

Several instabilities threaten jets. Let's examine the most dangerous ones and see why jets survive.

Kelvin-Helmholtz: The Shear Instability

At the jet boundary, fast-moving jet material slides past slow ambient gas. Velocity shear drives the Kelvin-Helmholtz (KH) instability—the same instability that makes flags flutter and creates ocean waves.

In a simple fluid, the KH growth rate is roughly

$$\gamma_{KH} \sim \frac{k\Delta v}{(1+\rho_j/\rho_{\rm ext})^{1/2}}$$

where k is the wavenumber, Δv is the velocity difference, and ρ_j, $\rho_{\rm ext}$ are the jet and external densities.

For a light jet ($\rho_j \ll \rho_{\rm ext}$) moving at high speed, this grows fast. The jet should be shredded within a few wavelengths.

But jets aren't simple fluids. Several effects stabilize them:

Relativistic motion: In special relativity, inertia increases with Lorentz factor. A relativistic jet acts as if it's much heavier than its rest mass would suggest. This slows the instability.

Magnetic fields: A field component along the flow direction resists bending. The field tension acts like the tension in a garden hose, opposing transverse perturbations.

Gradual shear: If the velocity transition is spread over a "sheath" layer rather than being a sharp discontinuity, the instability is weakened. A boundary layer absorbs the shear.

Kink: The Current-Driven Instability

The toroidal field that collimates the jet carries an electric current (by AmpÃÍre's law: $\nabla \times \vec{B} = (4\pi/c)\vec{j}$). Current-carrying conductors are subject to the kink instability, which displaces the current channel in a helical pattern.

Think of a current-carrying wire. The magnetic field it produces circles around the wire. If the wire is perturbed into a slight helix, the field is stronger on the inside of the helix than the outside. This creates a net force that amplifies the perturbation. The wire kinks.

In jets, the kink instability can grow when the twist exceeds a critical value:

$$\frac{B_\phi}{B_p} > \frac{2\pi\omega}{L}$$

where L is the length of the jet. For a long jet with heavily wound field, this condition is usually satisfied.

But here's the subtlety: the kink instability doesn't necessarily destroy the jet. It may saturate at finite amplitude, producing the helical structures we observe in some jets rather than disrupting them entirely. And when the kink does bring oppositely-directed fields together, it triggers reconnection, which—as we discussed—may be how jets solve the σ problem.

So the kink instability might actually help jets, by enabling efficient acceleration through reconnection while not growing fast enough to destroy the jet.

9.11 The Spine-Sheath Structure

When we look closely at jets—with VLBI radio imaging, for instance—we see that they're not uniform. The edges are often brighter than the center. The velocity varies across the jet. The structure is layered.

The standard picture is a "spine-sheath" structure:

The spine is the inner core, powered by the BZ mechanism. It consists of field lines that thread the black hole horizon. The plasma

here is probably electron-positron pairs, created in the strong electromagnetic fields near the hole. The spine is highly relativistic, with $\Gamma \sim 10$–100, and magnetically dominated ($\sigma \gg 1$) near the base. It carries most of the jet's energy.

The sheath is the outer layer, powered by the BP mechanism. It consists of field lines anchored in the disk. The plasma is probably normal electron-proton matter from the disk. The sheath is mildly relativistic, with $\Gamma \sim 2$–5, and carries less energy but more mass.

The spine-sheath structure isn't imposed by hand—it emerges naturally from the physics. Field lines from the black hole and from the disk are geometrically separate. They carry different amounts of energy per unit mass. They naturally form distinct layers.

Observations support this picture. "Limb brightening"—edges brighter than center—is seen in M87 and Centaurus A. This is consistent with a fast, faint spine (beamed away from us) surrounded by a slower, brighter sheath (more visible because less beamed).

9.12 *What Are Jets Made Of?*

We've discussed jets as "plasma," but what kind of plasma? The answer matters because it affects how jets behave.

Electron-proton plasma is ordinary matter. It's heavy: the proton mass is 1836 times the electron mass. Heavy plasma is hard to accelerate to high speeds.

Electron-positron plasma is made of pairs. It's light: both particles have the electron mass. Light plasma accelerates easily.

How do we tell which kind of plasma is in a jet? Several observational handles exist:

The *spectrum* differs because pair plasmas and proton plasmas radiate differently under the same conditions.

Circular polarization from Faraday rotation depends on the charge-to-mass ratio of the plasma. Pairs have different Faraday properties than electron-proton plasma.

Dynamics differ because the inertia per particle differs by a factor of $\sim$1000.

Current evidence suggests that AGN jets are probably electron-proton dominated in the sheath (which is fed by disk material), but may be pair-dominated in the spine (where pairs can be created by the strong fields near the black hole). Gamma-ray burst jets, which show extreme Lorentz factors $\Gamma > 100$, may require substantial pair content to achieve such high speeds.

The loading mechanism—how plasma gets into the jet in the first place—remains poorly understood.

9.13 What Simulations Tell Us

The physics we've discussed can be simulated numerically. Modern GRMHD (general relativistic magnetohydrodynamics) codes solve Einstein's equations for gravity coupled to Maxwell's equations for electromagnetism and the MHD equations for the plasma. They can resolve the black hole's event horizon, the accretion disk, and the jet launching region all together.

What do these simulations show?

MAD versus SANE: Simulations divide into two classes depending on how much magnetic flux accumulates at the black hole. "SANE" (Standard And Normal Evolution) systems have moderate flux; accretion is relatively steady. "MAD" (Magnetically Arrested Disk) systems have so much flux that it partially blocks accretion; accretion becomes episodic, and jets are powerful.

The jet power differs dramatically. In MAD simulations with a rapidly spinning black hole ($a = 0.99$), the jet power can exceed 100% of the accretion power:

$$\eta_{\rm jet} = \frac{P_{\rm jet}}{\dot{M}c^2} > 1$$

This means the jet extracts energy from the black hole faster than accretion adds it. The black hole is actually spinning down, donating rotational energy to the jet. This is the BZ mechanism in action.

Natural spine-sheath structure: Simulations spontaneously develop the layered structure we discussed. The innermost field lines, threading the horizon, launch the relativistic spine. The outer field lines, anchored in the disk, launch the slower sheath. This structure isn't imposed—it emerges from the physics.

Variability: MAD simulations show chaotic variability on timescales of tens to hundreds of r_g/c—the light-crossing time of the inner accretion flow. For a $10^9 M_\odot$ black hole like M87's, this is weeks to months. Such variability is indeed observed in blazar flares.

9.14 The Disk-Jet Connection

A jet isn't a separate entity from its disk. They're coupled systems, exchanging energy, angular momentum, and magnetic flux.

Energy: The jet draws energy from accretion (BP) or black hole spin (BZ). In MAD states, the jet can dominate the energy budget.

Angular momentum: The jet carries angular momentum away, enabling accretion to proceed. The magnetic lever arm makes this efficient.

Flux: The disk supplies magnetic flux to the system. The jet organizes and uses it.

These couplings create feedback. If the jet becomes more powerful, it may starve the disk of magnetic flux, weakening itself. If the disk accumulates too much flux, accretion may be suppressed, reducing the fuel supply. The system can oscillate, producing variability.

9.15 Open Questions

We've come far in understanding jets, but fundamental puzzles remain.

How does flux accumulate? Reaching the MAD state requires getting magnetic flux to the black hole, but MRI turbulence in the disk tends to diffuse flux outward. Some process must transport flux inward faster than turbulence spreads it. We don't fully understand what.

What triggers state transitions? X-ray binaries switch between jet-loud and jet-quiet states on timescales of days to weeks. Something crosses a threshold, but we don't know what.

How do jets accelerate efficiently? Ideal MHD struggles to achieve $\Gamma > 10$. Reconnection may be the answer, but the details of how reconnection operates in jets are still being worked out.

How do jets stay stable? The observed stability over megaparsec scales isn't fully explained by any single mechanism.

These questions drive current research. New observations—the Event Horizon Telescope's images of M87, multi-messenger detections of jet-associated neutrinos—combined with ever-more-sophisticated simulations are making progress.

The jet problem—how nature converts gravitational energy into relativistic, collimated outflows stable over vast distances—remains one of the great challenges in astrophysics. But the framework we've developed in this chapter, built from magnetic tension, critical surfaces, and energy conversion, gives us the tools to attack it.

10
Jets in Flight

There is something almost absurd about a jet. A stream of matter, thinner than a laser beam relative to its length, traveling at 99.9% of the speed of light, maintaining its coherence across distances that light itself takes millions of years to cross. If the launching region near the black hole were the size of a marble, the jet would stretch from here to the Moon—and arrive still focused, still powerful, still recognizably a jet.

How is this possible? Everything in nature tends toward disorder. Rivers spread into deltas. Smoke disperses. Sound fades. Yet jets—buffeted by instabilities, pushing through ambient gas, radiating away their energy—persist. They punch through entire galaxies. They inflate bubbles in the intergalactic medium the size of small galaxy clusters. They are, in some sense, the longest-lasting structures that active black holes produce.

This chapter follows jets on their journey from engine to lobe. We'll see how relativistic effects make jets appear even stranger than they are, how they interact with their environment, and why some jets remain focused while others disperse.

10.1 The Scale of the Problem

The dynamic range of jets is staggering. Consider the jet from M87:

- Launching region: $\sim 10 r_g \approx 10^{-2}$ pc
- Visible jet length: ~ 2 kpc
- Radio lobe extent: ~ 80 kpc

The ratio of lobe size to launching region is 10^7. If the launching region were the size of a period on this page ($\sim$ 0.3 mm), the lobes would extend 3 km.

No simulation can capture this range. No single observation can resolve all scales. We piece together the story from theory, from ob-

servations at different scales, and from simulations of limited dynamic range.

10.2 *Relativistic Beaming*

Before discussing jet structure, we must understand relativistic beaming—a kinematic effect that profoundly affects how jets appear.

A source moving at relativistic speed toward the observer has its radiation concentrated in a narrow cone in the forward direction. The opening angle of this cone is $\theta \sim 1/\Gamma$, where Γ is the Lorentz factor. For $\Gamma = 10$, the cone is only 6 wide.

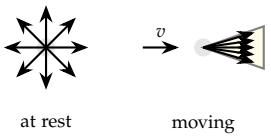

Figure 10.1: Relativistic beaming. A source emitting isotropically at rest (left) has its radiation concentrated in a narrow forward cone when moving relativistically (right).

Beaming has several consequences:

1. **Brightness enhancement:** The intensity is boosted by a factor $\delta^{3+\alpha}$, where $\delta = [\Gamma(1 - \beta\cos\theta)]^{-1}$ is the Doppler factor and α is the spectral index. For a jet pointing at us with $\Gamma = 10$, this can be a factor of 10^4.

2. **Apparent superluminal motion:** A jet moving at angle θ to the line of sight appears to move faster than light if $\beta > \cos\theta$. The apparent velocity is $v_{\rm app} = v\sin\theta/(1 - \beta\cos\theta)$. Maximum superluminal motion of Γv occurs at $\theta = 1/\Gamma$.

3. **Asymmetry:** The approaching jet is brightened while the receding jet is dimmed. In highly relativistic jets, the receding ("counter") jet can be invisible.

Superluminal motion was first observed in the quasar 3C 279 in 1971 by a team using newly-developed Very Long Baseline Interferometry. Components of the jet appeared to move at 10 times the speed of light! The announcement caused a stir. Here was an apparent violation of special relativity, observed in multiple independent measurements.

Now, you might say, "That's impossible. Nothing can travel faster than light." And you'd be right—nothing *is* traveling faster than light. The jet is moving at $\sim 0.998c$ at a small angle to our line of sight. What looks like superluminal motion is a projection effect, like how a lighthouse beam sweeping across distant clouds can appear to move faster than light even though the photons themselves travel at exactly c. The discovery initially perplexed astronomers—some even wondered if relativity needed revision—but the explanation is elegant and purely kinematic.

10.3 Jet Structure on VLBI Scales

Very Long Baseline Interferometry (VLBI) resolves jets on scales of light-weeks to light-years, close to the launching region. At these scales, we see:

The core: An unresolved bright point, usually identified as the base of the jet where it becomes optically thin. The core position depends on frequency—higher frequencies probe closer to the black hole.

Knots: Bright features moving along the jet. Some knots are standing shocks where the jet recollimates. Others are moving components, probably dense regions ejected during outbursts.

Transverse structure: Edge-brightened jets (limb brightening) suggest a spine-sheath structure, with a fast, faint spine surrounded by a slower, brighter sheath.

Proper motion: Knots move, typically at apparent speeds of 5–20c in quasars, corresponding to Lorentz factors $\Gamma \sim 10$–30.

The opening angle of jets on VLBI scales is typically a few degrees. This is wider than the intrinsic opening angle at the launching region (which is less than 1), suggesting that the jet expands as it propagates, probably due to pressure mismatch with the surrounding medium.

10.4 Interaction with the Environment

As the jet propagates, it interacts with ambient gas. The nature of this interaction changes with distance from the source.

Inner region (sub-pc to pc scale): The jet propagates through the broad-line region and dusty torus of the active nucleus. Interactions here can entrain material into the jet, possibly slowing it.

Galactic scale (kpc): The jet pushes through the interstellar medium of the host galaxy. It creates a cavity in the ISM, bordered by shock-heated gas. The jet can be disrupted if the ISM is dense and clumpy.

Intergalactic scale (tens of kpc to Mpc): The jet emerges into the low-density intergalactic medium (IGM). Here it can propagate relatively freely, depositing its energy into large-scale lobes.

At each stage, the jet does work on its environment, decelerating as it does so. A jet that starts with $\Gamma \sim 10$ at the core may slow to $\Gamma \sim 2$ by the kiloparsec scale and become sub-relativistic in the lobes.

10.5 The Jet-Lobe Connection

At the end of its journey, the jet inflates a lobe—a large, amorphous region of relativistic plasma emitting synchrotron radiation. Lobes

are the "exhaust" of the jet engine.

Classic radio galaxies like Cygnus A show bilateral lobes, one on each side of the central galaxy, connected to the nucleus by narrow jets. The lobes can be hundreds of kiloparsecs across and contain 10^{58}–10^{60} erg of energy.

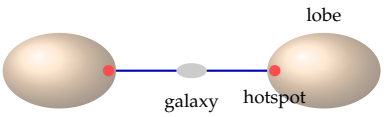

Figure 10.2: Classic radio galaxy morphology: central galaxy, bilateral jets, and extended lobes with hotspots where jets terminate.

Where the jet rams into the lobe material, a strong shock forms—the hotspot. Hotspots are the brightest features in many radio galaxies, outshining the lobes by factors of 10 or more. The shock accelerates particles to ultrarelativistic energies, producing intense synchrotron radiation.

Behind the hotspot, the jet material spreads out to fill the lobe. The lobe expands subsonically into the IGM, driving a weak shock ahead of it. The expansion is approximately self-similar: the lobe grows while maintaining a roughly constant axis ratio.

10.6 FR I and FR II Radio Galaxies

Not all radio galaxies look like Cygnus A. In 1974, Bernard Fanaroff and Julia Riley were cataloguing radio sources when they noticed something curious: the galaxies seemed to fall into two distinct groups, not a continuum. This wasn't what anyone expected. Why should there be two types rather than a smooth range? Their classification scheme, published in a short paper, turned out to be one of the most enduring in extragalactic astronomy. Forty years later, we're still trying to fully explain it.

FR I (edge-darkened): Jets are brightest near the nucleus and fade with distance. No hotspots. Lobes are diffuse and poorly defined. Lower total radio power.

FR II (edge-brightened): Jets are faint near the nucleus and brightest at the edges. Prominent hotspots. Lobes are well-defined with sharp edges. Higher total radio power.

The dividing radio luminosity is about 10^{25} W/Hz at 1.4 GHz, though it depends on the optical luminosity of the host galaxy.

What causes this dichotomy? The leading hypothesis is that FR I jets are decelerated to sub-relativistic speeds within the host galaxy, while FR II jets remain relativistic until they hit the IGM. The deceleration could be caused by entrainment of external material, which slows the jet and makes it turbulent.

Decelerated, turbulent jets deposit their energy gradually, producing diffuse emission (FR I). Fast jets deposit their energy abruptly at the termination shock, producing bright hotspots (FR II).

The transition between FR I and FR II may be controlled by jet power relative to the density of the surrounding medium. More powerful jets can push through the ISM without decelerating. The

fact that the FR I/II boundary depends on host galaxy luminosity supports this: more luminous galaxies have denser environments, requiring more powerful jets to remain FR II.

10.7 Stability and Instabilities

How do jets maintain their coherence over seven orders of magnitude in scale? Why don't they break up into turbulent eddies?

This is exactly the right question to ask. You might think that with all the instabilities we know about—Kelvin-Helmholtz, kink modes, Rayleigh-Taylor—jets shouldn't survive more than a few scale heights. The plasma should shred itself. And yet there they are in the images, narrow and focused across megaparsecs. What's going on?

Jets are subject to several instabilities:

Kelvin-Helmholtz instability: The velocity shear between the jet and the ambient medium generates surface waves that can grow and disrupt the jet. This is the instability that makes flags flutter in the wind.

Current-driven instabilities: The toroidal magnetic field can become unstable to kink modes, bending the jet into a helix that tightens until the jet disrupts.

Rayleigh-Taylor instability: If the jet decelerates, denser external material can penetrate into the lighter jet material.

Yet jets survive. Several factors contribute to their stability:

Relativistic motion: In the jet frame, instabilities grow at a rate proportional to the sound speed. But time dilation stretches the growth time by a factor Γ. A jet moving at $\Gamma = 10$ has 10 times longer to propagate before instabilities grow.

Magnetic tension: The toroidal field provides hoop stress that resists transverse displacements. Helical fields can be particularly stabilizing.

Sheath: The slower, denser sheath surrounding the spine provides a buffer layer that damps instabilities trying to grow at the jet surface.

Expansion: As the jet expands, the growth rate of wavelength-dependent instabilities decreases. Modes that might have disrupted a narrow jet become harmless as the jet widens.

Observed jets do show evidence of instabilities. Wiggles, knots, and oscillations are visible on VLBI scales. Some jets show helical patterns suggestive of kink instabilities. But the instabilities rarely lead to complete disruption—the jets remain coherent.

10.8 Jet Content and Composition

What are jets made of? This seemingly simple question has proven surprisingly difficult to answer. We can see the jets, we can measure their power, we can track their motion—but we still argue about what's actually in them. It's a bit like watching a fire hose from a distance and being unable to tell if it's spraying water or gasoline.

The relativistic particles emitting synchrotron radiation are electrons (and positrons, if present). But electrons alone are not enough to carry the jet's momentum. There must be either:

- Protons (a normal electron-proton plasma)
- Positrons (an electron-positron pair plasma)
- Electromagnetic fields (a Poynting-flux-dominated jet)

Each option has different implications:

Electron-proton: Heavy particles carry most of the momentum and energy. The electrons radiate but are dynamically unimportant. This composition is expected if the jet is made of normal baryonic material from the accretion disk.

Electron-positron: Equal numbers of particles and antiparticles. Pairs can be created by gamma-ray interactions near the black hole. Pair jets are lighter and potentially faster than baryon-loaded jets.

Poynting flux: Energy carried by electromagnetic fields rather than particles. Near the launching region, jets are probably Poynting-flux-dominated. The question is whether they remain so at large distances.

Observations provide mixed evidence. Some jets show circular polarization, which is easier to produce in pair plasmas. But the inferred energy budgets of some jets seem to require baryons.

Current thinking is that jets are probably Poynting-flux-dominated near the source and become particle-dominated at large distances as magnetic energy is converted to kinetic energy. The particle content may be a mixture of pairs and baryons, with the ratio varying between sources.

10.9 Jet Energetics and Feedback

Jets carry enormous amounts of energy. Where does this energy end up?

Some is radiated as synchrotron and inverse-Compton emission. This is what we observe as radio jets and lobes. But the radiated power is typically only a small fraction of the total jet power.

Most of the energy is deposited in the environment through:

- **Shock heating:** The bow shock ahead of the jet heats ambient gas to X-ray temperatures (10^7–10^8 K).
- **Cavity inflation:** The jet inflates a low-density cavity (lobe) that displaces ambient gas. The work done in cavity inflation can exceed the radiated energy.
- **Turbulence and mixing:** Jet material mixing with ambient material creates turbulence that eventually thermalizes.

This energy injection has profound effects on the surrounding gas, a process called "AGN feedback." In galaxy clusters, jets from central AGN heat the intracluster medium, preventing it from cooling and forming stars. Without this heating, cluster cores would have much higher star formation rates than observed.

Source	Jet power (erg/s)	Cavity energy (erg)
Perseus A	10^{45}	10^{59}
M87	10^{44}	10^{57}
Cygnus A	10^{46}	10^{60}

Table 10.1: Jet powers and cavity energies for selected radio galaxies. Cavity energies are inferred from X-ray observations of the displaced hot gas.

The feedback loop is self-regulating. When the central black hole accretes actively, jets heat the surrounding gas, shutting off the supply of cold gas that could fuel further accretion. When accretion drops, the jets weaken, the gas cools, and accretion resumes. This cycle operates over timescales of millions of years.

10.10 The Jet Fate

What ultimately happens to jets? Several endpoints are possible:

Dissipation in lobes: Most of the jet energy ends up in the lobes, where it slowly radiates away over hundreds of millions of years. The relativistic electrons lose energy to synchrotron radiation and inverse-Compton scattering off the cosmic microwave background.

Mixing with the IGM: Eventually, lobe material mixes with the surrounding medium. The magnetic fields become tangled and diluted. The relativistic particles thermalize. The lobe fades into the general IGM.

Fossil cavities: Even after the radio emission fades, the cavities persist as underdense bubbles in the hot cluster gas. These "ghost cavities" can be detected in X-rays and may last for billions of years.

The energy deposited by jets doesn't disappear—it's just distributed over ever-larger volumes. Jets that were narrow beams at the source become diffuse thermal energy spread across the intergalactic medium, contributing to the heating of cosmic gas.

11
The Universal Disk-Jet Connection

We've now developed the physics of accretion disks and jets separately. But perhaps the most remarkable fact about these phenomena is that they appear together, across an astonishing range of scales. From young stars still gathering mass to the billion-solar-mass black holes at the centers of giant galaxies, nature produces disk-jet systems with striking similarities.

This chapter explores the universality of the disk-jet connection. We'll survey the systems where disks and jets appear, look for scaling relations that unify them, and ask whether the same physics operates at all scales.

11.1 The Mass Hierarchy

The central objects in disk-jet systems span more than nine orders of magnitude in mass:

Object	Mass	System
Young stellar objects (YSOs)	0.1–$10 M_\odot$	T Tauri, FU Ori
White dwarfs	0.6–$1.4 M_\odot$	Cataclysmic variables
Neutron stars	1.4–$2.2 M_\odot$	X-ray binaries
Stellar-mass black holes	5–$20 M_\odot$	X-ray binaries
Intermediate-mass black holes	10^2–$10^5 M_\odot$	ULXs, globular clusters?
Supermassive black holes	10^6–$10^{10} M_\odot$	AGN, quasars

Table 11.1: Central objects in disk-jet systems, ordered by mass.

Despite this enormous range, all these systems show qualitatively similar behavior: accretion through a disk, launching of bipolar outflows, and coupling between accretion state and outflow power.

11.2 *Young Stellar Objects*

The closest disk-jet systems are around young stars in our own galaxy. T Tauri stars—young, low-mass stars still contracting toward the main sequence—are surrounded by protoplanetary disks and often drive bipolar outflows.

These outflows are called Herbig-Haro jets, after their discoverers. They appear as chains of glowing knots extending up to parsecs from the central star. The knots are shock-heated gas where fast-moving jet material runs into slower ambient medium.

Herbig-Haro jets are not relativistic. Typical velocities are 100–500 km/s, far below the speed of light. But their collimation and morphology are strikingly similar to extragalactic jets. They're narrow (opening angles of a few degrees), knotty (with emission features moving outward), and often one-sided (the receding jet hidden by the disk).

The jet launching mechanism in YSOs is thought to be magnetocentrifugal—essentially the Blandford-Payne mechanism. The star has a strong magnetic field that threads the inner disk. Disk material tied to field lines is flung outward, then collimated by magnetic hoop stress.

The jet power in YSOs is typically 10^{-2} to 10^{-1} times the accretion luminosity. This ratio is similar to what's seen in X-ray binaries, suggesting a common underlying mechanism.

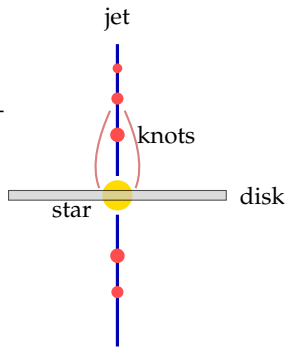

Figure 11.1: Schematic of a T Tauri disk-jet system. The star (center) is surrounded by an accretion disk. Bipolar jets with bright knots extend perpendicular to the disk.

11.3 *Microquasars*

Stellar-mass black holes accreting from companion stars are called X-ray binaries. When they produce relativistic jets, they're sometimes called "microquasars"—scaled-down versions of the quasars powered by supermassive black holes.

The analogy is apt. Microquasars show many of the same phenomena as quasars:

- Relativistic jets with apparent superluminal motion
- Core-jet structure on VLBI scales
- Variability on timescales from milliseconds to months
- Spectral states (hard and soft) with different jet properties

The scaling is dramatic. A microquasar might have $M = 10 M_\odot$, while a quasar has $M = 10^9 M_\odot$. The gravitational radius scales linearly with mass: $r_g \propto M$. So does the orbital timescale at the innermost stable circular orbit: $t_{\rm ISCO} \propto M$. A microquasar with orbital timescales of milliseconds corresponds to a quasar with orbital timescales of hours.

You might say, "So what? Stars and galaxies are different sizes—everything scales." But that misses what's remarkable here. The physics isn't just scaling—it's the *same* physics. The same instabilities, the same magnetic field configurations, the same accretion states. It's as if you found that car engines and aircraft carriers use not just the same principles but the same blueprints, just photocopied at different sizes.

This scaling allows us to study quasar physics in real time. We can watch a microquasar go through an outburst cycle in weeks—a process that takes millions of years in a quasar. The microquasar GRS 1915+105 has been called a "laboratory for black hole accretion physics."

11.4 *The Fundamental Plane of Black Hole Activity*

Is there a quantitative relationship connecting disk-jet systems across the mass scale? In 2003, Merloni, Heinz, and Di Matteo discovered the "fundamental plane of black hole activity"—a correlation relating three quantities:

- Black hole mass M
- X-ray luminosity L_X (a proxy for accretion rate)
- Radio luminosity L_R (a proxy for jet power)

The correlation is:

$$\log L_R = 0.60 \log L_X + 0.78 \log M + \text{const.}$$

This relation holds over 9 orders of magnitude in mass and 12 orders of magnitude in luminosity. It connects stellar-mass black holes in X-ray binaries to supermassive black holes in AGN on a single plane.

What does this relation mean? The $L_R \propto L_X^{0.6}$ scaling suggests that jet power doesn't scale linearly with accretion power. At low accretion rates (where most fundamental plane sources lie), the efficiency of converting accretion power to jet power may increase—consistent with the observation that hard-state, ADAF-like flows are more jet-efficient than soft-state thin disks.

The $L_R \propto M^{0.78}$ scaling indicates that more massive black holes produce more powerful jets at a given Eddington ratio. This could reflect the role of the Blandford-Znajek mechanism: BZ power scales as M^2 (since both the horizon area and the magnetic flux threading it scale with M), modulated by other factors.

The fundamental plane is an empirical relation, not a theoretical prediction. But its existence strongly suggests that the same physics operates in systems differing by 10^9 in mass.

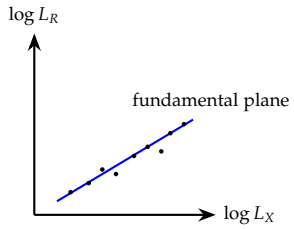

Figure 11.2: Schematic of the fundamental plane. X-ray and radio luminosities are correlated, with the plane tilted by the dependence on black hole mass.

11.5 Jets from Neutron Stars and White Dwarfs

Black holes aren't the only compact objects that produce jets. Neutron stars in X-ray binaries also drive outflows, though they're typically less powerful than black hole jets.

The existence of neutron star jets is significant because neutron stars have solid surfaces, not event horizons. This means the Blandford-Znajek mechanism cannot operate—there's no ergosphere to extract rotational energy from. The neutron star just sits there, stubbornly refusing to have its spin extracted through clever general-relativistic tricks.

Neutron star jets are probably powered by the Blandford-Payne mechanism (launching from the disk) or by magnetic interactions between the neutron star's magnetic field and the accretion flow. The fact that they're weaker than black hole jets at comparable accretion rates suggests that the BZ mechanism contributes significantly to black hole jet power.

White dwarfs in cataclysmic variable systems can also produce jets, though they're even weaker and less collimated than neutron star jets.

A remarkable case is SS 433, an X-ray binary in which a compact object (probably a black hole) launches persistent jets at $0.26c$. The jets precess with a 162-day period, tracing a characteristic corkscrew pattern across the sky. SS 433 demonstrates that steady, well-collimated jets can persist even at super-Eddington accretion rates.

11.6 The Critical Question: What Determines Jet Power?

We can now ask: what makes some disk-jet systems launch powerful jets while others are jet-quiet?

Based on observations and theory, several factors seem important:

Accretion state: Jets are more powerful (relative to accretion luminosity) in hard states than soft states. This suggests that hot, geometrically thick accretion flows are more conducive to jet launching than cold, thin disks.

Magnetic flux: Systems with more magnetic flux threading the black hole (approaching MAD) should produce more powerful jets. This is difficult to measure observationally, but the correlation between jet power and radio loudness supports the idea.

Black hole spin: BZ power scales as a^2. Rapidly spinning black holes should produce more powerful jets. Evidence for this is mixed—some studies find correlations between jet power and spin, others don't.

Environment: The surrounding medium affects jet propagation,

but does it affect launching? Possibly—external confinement could help collimate the jet, while dense environments could impede early-stage jet development.

No single factor explains all the observed variation. Jet power is probably determined by a combination of all these factors, with different factors dominating in different systems.

11.7 Jets and Galaxy Evolution

Jets aren't just interesting in themselves—they affect the galaxies that host them.

The energy deposited by jets into the surrounding gas can exceed the binding energy of that gas. In galaxy clusters, jets from central AGN heat the intracluster medium, preventing it from cooling and forming stars. This "AGN feedback" is now recognized as essential for understanding galaxy evolution.

Without AGN feedback, simulations of galaxy formation produce galaxies that are too massive and too blue (too many stars forming). Including feedback brings the simulations into agreement with observations. Jets are one of the main channels for this feedback.

The disk-jet connection thus links the smallest scales (the accretion disk, at 10^{-5} pc) to the largest (galaxy clusters, at 10^{6} pc). The physics of angular momentum transport near a black hole ultimately determines the appearance of the cosmic web.

11.8 Planets Forming in Disks

At the smallest end of the disk-jet hierarchy are protoplanetary disks around young stars. These disks don't just accrete—they form planets.

The same angular momentum transport mechanisms that drive accretion also redistribute material within the disk. MRI turbulence stirs dust particles, affecting their growth. Disk winds can remove angular momentum without the viscous heating that would evaporate volatiles.

Interestingly, protoplanetary disks may have "dead zones" where the ionization fraction is too low for the MRI to operate. In these regions, angular momentum transport is suppressed, and material piles up. The resulting overdensities might be sites where planets preferentially form.

Jets from young stars can also affect planet formation. The mass carried away by jets reduces the total mass available for planet building. The magnetic fields that drive jets may also influence the migration of planets through the disk.

11.9 A Unified Picture

Let's step back and see the big picture. Wherever matter with angular momentum falls toward a gravitating object, a disk forms. The disk is a natural consequence of angular momentum conservation—the rotating gas cannot simply fall inward.

To accrete, angular momentum must be transported outward. In ionized gas, this transport is provided by the magnetorotational instability, which requires magnetic fields. Even weak fields are amplified by the MRI into sustained turbulence.

Magnetic fields threading the disk can launch outflows. Material along field lines is flung outward by centrifugal force (Blandford-Payne) or energy is extracted from the spinning black hole (Blandford-Znajek). These outflows are collimated by magnetic hoop stress into jets.

The same physics operates at all scales. What changes is the quantitative details: temperatures, densities, velocities, timescales. But the underlying processes—angular momentum transport by magnetic turbulence, jet launching by magnetic stresses—are universal.

11.10 Open Questions

Despite the success of the unified picture, many questions remain:

Why is the disk-jet coupling variable? Some systems are persistently jet-loud; others alternate between jet-loud and jet-quiet states. What triggers the transitions?

What sets the jet Lorentz factor? Jets range from mildly relativistic ($\Gamma \sim 2$) to highly relativistic ($\Gamma > 50$). The launching mechanisms don't obviously predict specific values.

How do jets maintain stability? Theoretical analysis suggests jets should be unstable, yet they remain coherent over enormous distances. What stabilizes them?

Is there a maximum jet power? As jet power approaches the accretion power, the flow must transition to a magnetically dominated state. Are there physical limits on this?

Can we detect jets from planet-mass objects? If the disk-jet connection is universal, do planetary-mass objects with disks (like forming giant planets or brown dwarfs) also launch jets? The signals would be faint but potentially detectable.

11.11 How We Learned It Was Universal

The universality of disk-jet physics wasn't obvious from the start. For decades, astronomers studying T Tauri jets and astronomers study-

ing AGN jets worked in separate communities, attended different conferences, and read different journals. The connection had to be discovered.

The MRI provides a striking example. Velikhov derived the instability criterion in 1959 while studying liquid metal flows in the laboratory. Chandrasekhar rediscovered it in 1960 as a mathematical exercise in stability theory. Neither was thinking about accretion disks. The result sat in the literature for thirty years until Balbus and Hawley, in 1991, recognized it as the solution to the viscosity problem that had plagued disk theory since Shakura and Sunyaev.

You might say, "That's just how science works—ideas get rediscovered." But notice what happened next. Once the MRI was understood in the context of black hole disks, people immediately asked: does it operate in protoplanetary disks? In cataclysmic variables? In neutron star systems? The answer was yes, yes, and yes. A single mechanism explained angular momentum transport across nine orders of magnitude in mass.

Today, observers studying microquasars and observers studying quasars attend the same conferences and speak the same language. A graduate student can learn about jet collimation from M87 and apply it to HH 30. The equations don't know the difference between a ten-solar-mass black hole and a ten-billion-solar-mass one—only the timescales change.

11.12 What We Still Don't Know

We began these lectures with a simple question: why does matter spiral rather than fall? The answer led us through angular momentum conservation, viscous transport, magnetic turbulence, jet launching, and cosmic feedback. We've seen how a figure skater spinning faster when she pulls in her arms connects to the brightest objects in the universe.

But let me end with what we don't understand.

We don't know why some black holes launch powerful jets while others, accreting at the same rate, remain radio-quiet. We don't know what sets the dividing line between FR I and FR II radio galaxies—the answer involves jet power and environment, but the details remain murky. We don't know how magnetic flux accumulates at the black hole to reach the MAD state, or why it happens in some systems and not others.

Most fundamentally, we don't know why magnetic fields. Of all the ways angular momentum could be transported, nature chose magnetized turbulence. Of all the ways energy could be extracted from spinning black holes, nature chose electromagnetic. The MRI

works because magnetic tension couples fluid elements at different radii. The Blandford-Znajek mechanism works because magnetic field lines can thread the ergosphere. Take away the magnetic field and none of this happens—disks don't accrete, jets don't launch, quasars don't shine.

Why should magnetic fields be so central to accretion physics? Is it inevitable, given the basic laws? Or is it a contingent fact about our universe that could have been otherwise? I don't know. But it's the kind of question that makes you want to keep working on the problem.

Bibliography

[1] N. I. Shakura and R. A. Sunyaev. Black holes in binary systems. observational appearance. *Astronomy and Astrophysics*, 24:337–355, 1973.

[2] D. Lynden-Bell. Galactic nuclei as collapsed old quasars. *Nature*, 223:690–694, 1969.

[3] J. E. Pringle. Accretion discs in astrophysics. *Annual Review of Astronomy and Astrophysics*, 19:137–162, 1981.

[4] Steven A. Balbus and John F. Hawley. A powerful local shear instability in weakly magnetized disks. i. linear analysis. *The Astrophysical Journal*, 376:214–233, 1991.

[5] R. D. Blandford and R. L. Znajek. Electromagnetic extraction of energy from kerr black holes. *Monthly Notices of the Royal Astronomical Society*, 179:433–456, 1977.

[6] R. D. Blandford and D. G. Payne. Hydromagnetic flows from accretion discs and the production of radio jets. *Monthly Notices of the Royal Astronomical Society*, 199:883–903, 1982.

[7] Ramesh Narayan and Insu Yi. Advection-dominated accretion: A self-similar solution. *The Astrophysical Journal Letters*, 428: L13–L16, 1994.

[8] M. J. Rees. Black hole models for active galactic nuclei. *Annual Review of Astronomy and Astrophysics*, 22:471–506, 1984.

[9] Juhan Frank, Andrew King, and Derek Raine. *Accretion Power in Astrophysics*. Cambridge University Press, 3rd edition, 2002.

[10] I. D. Novikov and K. S. Thorne. Astrophysics of black holes. *Black Holes (Les Astres Occlus)*, pages 343–450, 1973.

[11] M. A. Abramowicz, B. Czerny, J. P. Lasota, and E. Szuszkiewicz. Slim accretion disks. *The Astrophysical Journal*, 332:646–658, 1988.

[12] M. C. Begelman, R. D. Blandford, and M. J. Rees. Theory of extragalactic radio sources. *Reviews of Modern Physics*, 56:255–351, 1984.

[13] Alexander Tchekhovskoy, Ramesh Narayan, and Jonathan C. McKinney. Efficient generation of jets from magnetically arrested accretion on a rapidly spinning black hole. *Monthly Notices of the Royal Astronomical Society*, 418:L79–L83, 2011.

[14] Roger Penrose. Gravitational collapse: The role of general relativity. *Rivista del Nuovo Cimento*, 1:252–276, 1969.

[15] A. C. Fabian. Observational evidence of active galactic nuclei feedback. *Annual Review of Astronomy and Astrophysics*, 50:455–489, 2012.