PHYS 231 Lecture Notes – Week 7

Reading from Maoz (2^{nd} edition):

- **Sections 4.1–4.5**

Again, a lot of the material presented in class this week is well covered in Maoz, and we simply reference the book, with additional comments and derivations as needed. References to slides on the web page are in the format "slidesx.y/nn," where x is week, y is lecture, and nn is slide number.

7.1 High-Mass Stars

See Maoz §4.3.1

High-mass stars don't have degeneracy problems that limit their ability to burn heavier elements. As described in the text, they carbon and oxygen into neon, magnesium, silicon, and all the way up to iron. Fusion occurs principally by alpha capture (because of the lower coulomb barrier), bur C-C and other relatively low-mass reactions also occur. The process accelerates greatly as heavier and heavier elements are produced. The core of the star develops a layered structure, with shells of lighter elements surrounding a non-burning and growing nickel-iron core (see slides7.2/15,16).

The star now has a new problem. Up to now, this scenario would have led to a new round of fusion in the core and a temporary restoration of stability, but not in this case. Elements in the so-called iron group (with $A \approx 56$) have the largest binding energy per nucleon of all elements (see slides7.2/19). Simply put, fusing lighter elements (H, He, C, etc.) will increase the total binding energy, creating more tightly bound elements and releasing energy. Similarly, fission of heavy elements (U, Pu, etc.) also leads to more tightly bound nuclei and again releases energy. But the iron group elements are at the peak of the curve. They cannot split into lighter elements or fuse into heavier ones. Iron won't burn, and as it builds up in the core the star is close to the end.

As the degenerate core contracts and heats up, several processes occur:

1. The core is degenerate, despite temperatures approaching 10^{10} K, and its mass approaches the Chandrasekhar mass.

2. *Photodisintegration* splits nuclei apart, ultimately reducing them to protons and neutrons. Recall that, from slides7.2/19) the energy required to spilt a nucleus into lighter ones is a few MeV; for $T = 10^{10}$ K, $kT \approx 1$ MeV. These reactions remove a huge amount of energy (and pressure) from the radiation field.

3. *Neutronization* converts protons to neutrons, removing electrons and their degeneracy pressure, and creating neutrinos, which escape, removing even more energy and pressure from the core. Electron capture is much more common than neutron decay, and the result is that most nucleons are in the form of neutrons, with

$$n_p \;=\; n_e \;\approx\; \tfrac{1}{200} n_n .$$

7.2 Neutron Stars

See Maoz §4.3.2

With its pressure support greatly diminished, the star collapses. The collapse is similar to the idealized pressureless collapse we did back in Homework 1. The core collapses in a free-fall time $t_{ff} \sim (G\rho)^{-1/2} \sim 0.1$ s. The only process that can halt this near-relativistic collapse of a $\sim 1\,M_\odot$

ball of neutrons is *neutron degeneracy pressure*, which becomes effective at roughly nuclear densities. The core by this stage is a *neutron star.*

We can calculate the radius of a neutron star using the same formula as previously for a white dwarf, except that m_e is replaced by m_n. Since the scaling is

$$R \propto \frac{1}{m}\left(\frac{Z}{A}\right)^{5/3} M^{-1/3},$$

where m is the mass of the degenerate particle), we see that replacing m_e with m_n (about a factor of 1800 larger) and $Z/A = 0.5$ with $Z/A = 1$, the radius of a neutron star is about 500 times smaller than that of a white dwarf of the same mass:

$$R_{ns} \approx 14 \text{ km} \left(\frac{M}{M_\odot}\right)^{-1/3}.$$

The mean density of a $1.4\,M_\odot$ neutron star is approximately 2×10^{17} kg/m^3—the density of an atomic nucleus!

In line with our earlier study of white dwarfs, we can ask what the maximum mass of a neutron star. The earlier calculation of the Chandrasekhar mass is still valid (now with $Z/A = 1$), giving $M_{ch} = 5.6\ M_\odot$. However, as discussed in the text, a couple of important factors both tend to reduce this estimate. First, the neutron star is in a regime where general relativity is important in determining its structure—the star's total gravitational energy is about 20 percent of its total rest-mass energy Mc^2. One important fact about general relativity is that pressure contributes to the gravitational field, reducing its effectiveness in counteracting gravity. As a result, taking this into effect, the maximum mass decreases, to about $5\,M_\odot$. Second, the equation of state of matter at nuclear densities is not well known, but again it is thought that the changes relative to the idealized equation of state used earlier also tend to reduce the maximum mass, perhaps to $\sim 2 - 3\,M_\odot$.

7.3 Type II Supernovae

See Maoz §4.3.3

The collapsing stellar core has a lot of energy when it "bounces" at neutron star (nuclear) densities, and much of that energy is transferred into a violent shock wave running outward from the surface of the newborn neutron star through the rest of the star. Theoretical models still have significant areas of uncertainty in explaining exactly how this happens, but most researchers agree that, ultimately, the energy from the core propagates through the star and blows it apart in a *core-collapse supernova.*

Some energetics. The total gravitational energy of the collapsing core by the time it has reached nuclear density is

$$E_{gr} \sim \frac{GM^2}{R} \sim 3 \times 10^{46} \text{ J}.$$

The total kinetic energy in the form of ejected material is observed to be 3×10^{44} J. The total luminous energy is about 1 percent of this, $\sim 3 \times 10^{42}$ J. Together, these are a small fraction of the total available energy, most of which is actually released in the form of *neutrinos.* Nevertheless, even the electromagnetic energy is impressive—emitted over roughly a month (the typical duration of a visible supernova), it amounts to a mean luminosity of about a billion Suns, enough to make a single supernova temporarily burn as bright as its parent galaxy (slides8.1/05). Supernova 1987a (slides8.1/06), which occurred in the Large Magellanic Cloud, a satellite galaxy of the Milky Way, is the only supernova from which neutrinos have definitively been observed. The 20 neutrinos observed from that event were consistent with the energies just described and provided detailed and otherwise unavailable insight into conditions in the core bounce responsible for the explosion.

The details of how the energy emerges from the core are still the subject of intense research. At one time, researchers thought that a powerful spherical shock wave would simply sweep through the star, expelling all overlying material and leaving a neutron star behind. However, they quickly discovered that the combination of neutrino losses from the dense material just behind the outgoing shock and the weight of the overlying layers which, having lost pressure support from below, also begin to fall inward, meant that the shock stalled at a radius of a few hundred km, and never made it to the surface.

Conditions in the core itself are so extreme at these moments that the core is opaque not just to photons but also to neutrinos. The temperature exceeds 10^{10} K and the density is significantly above nuclear. Neutrinos are produced in abundance by reactions such as

$$\gamma + \gamma \ \rightarrow \ e^- + e^+ \ \rightarrow \ \nu + \bar{\nu},$$

where ν can be any of the three neutrino species, but most commonly is ν_e. Using the above mass density and relative numbers of particles, a photon's mean free path for Thomson scattering (cross section 6.7×10^{-29} m^2) is $\sim 2 \times 10^{14}$ m, while neutrinos, whose cross section for an interaction with a neutron at these energies (several MeV) is roughly 10^{-46} m^2, have mean free paths of just tens of meters. Thus to diffuse out of the roughly 10-km core, neutrinos must scatter approximately $10^5 - 10^6$ times, taking tens of milliseconds—a long time in these circumstances—to escape. It is thought that these neutrinos eventually reenergize the stalled shock, boosting its energy to the point where can accelerate outward through the star.

One process now recognized as important is that the spherical stalled shock is unstable, and the core becomes convective. It is likely that this convective motion, driven by neutrinos, can transport energy out to the stalled shock and restart the explosion (see the 3-D simulations on the course web page). The asymmetric motion associated with convection also means that the energy probably also escapes asymmetrically, meaning that there is a net momentum flux of neutrinos in some direction, causing the neutron star core to recoil in the other direction. These recoils have been observed in many neutron stars, which can have space velocities of hundreds of km/s, and were a puzzle until the inherently asymmetric nature of the explosion was understood.

7.4 Observations of Supernovae

See Maoz §4.3.3
Astronomers have long been aware of two types of supernova, which are distinguished principally by the presence or absence of hydrogen lines in their spectra. Supernovae without hydrogen lines are called Type I, those with hydrogen lines are Type II (see slides8.1/10). They may also be distinguished by their *light curves* (slides8.1/09): Type I supernovae shoe a characteristic exponential decline after the initial peak, while Type II supernovae have a characteristic plateau or even a secondary peak a few months after the initial burst.

Without going into any further detail, Type II supernovae are the core collapse supernovae we have just been discussing. They occur in stars with lots of hydrogen in their envelopes, so they have hydrogen in their spectra, and the secondary "bump" in the light curve is consistent with expectations as the hot gas from the explosion expands and cools.

What about Type I supernovae? Those are thought to occur when a white dwarf is pushed over the Chandrasekhar mass, starts to collapse, and undergoes degenerate carbon fusion (see the second simulation on the course web page). The star may be pushed over the limit by accretion from a companion in a binary, or two white dwarfs may collide (for example, as their orbit decays by gravitational radiation emission, see HW5). As we saw earlier nuclear burning an a degenerate gas is unstable, and in this case the entire star explodes once the burning starts. There is some debate about whether the process leaves a remnant behind. Most sources seem to think not, but a vocal minority claim that a neutron star may result. The light curves of Type I supernovae are consistent with the radioactive decay of ^{56}Ni and ^{56}Co produced during the detonation.

Curiously, despite the fact that they stem from very different stellar evolutionary processes and physical environments—Type I supernovae come from the products of low-mass stars in binary systems, while Type II supernovae are produced only by the most massive stars—both types of supernovae have similar peak luminosities, and occur with about the same frequency. About one of each is expected every hundred or so years in a galaxy like the Milky Way.

The heaviest elements (more massive than the iron group) are produced by neutron capture *after* the supernova explosion has begun. Both types of supernova produce huge fluxes of neutrons, which, having no coulomb barrier, easily combine with nuclei to form heavier elements. Depending on the intensity of the neutron flux, neutrons are jammed into nuclei, forcing them to higher and higher A at constant Z, and farther and farther away from stability (which is roughly characterized by $Z \approx A/2$), until the decay timescale of an unstable nucleus is shorter than the time for the next neutron to arrive. The nucleus then undergoes one or more beta decays, turning neutrons into protons and moving back toward stability, and the process continues for as long as the intense neutron flux persists. Calculations of this process (called the *r-process*) in supernovae produce elemental abundances that are in striking agreement with the observed cosmic abundances (slides8.1/12).

7.5 Hypernovae and Gamma-Ray Bursts

See Maoz §4.3.3
Gamma-ray bursts (GRBs) were first observed in the 1960s by U.S. military satellites tasked with verifying the nuclear test ban treaty. Systematic study began in the 1990s with the launch of the Compton Gamma Ray Observatory (CGRO), which observed bursts approximately once per day, distributed uniformly across the sky (slides8.1/18). Initially, distances were not known, so the luminosities of the bursts could not be determined. However, as techniques improved, it became possible for CGRO and other instruments to notify other satellites and ground-based instruments, which could observe a burst as it cooled through the X-ray and visible regimes (the so-called after-glow). These observations identified counterparts to many bursts and allowed their spectra to be measured, leading to the first distance measurements. Almost uniformly, GRBs are at *cosmological distances*—hundreds or thousands of megaparsecs away—meaning that they are extremely bright.

The bursts come in two types—short and long (slides8.1/19). Short bursts last only a fraction of a second, while long bursts may last tens or even hundreds of seconds. Long bursts appear to be associated with star-forming regions. Leading models (slides8.1/20) attribute the short bursts to neutron star mergers, driven by gravitational radiation, that cause the combined system to exceed the Chandrasekhar mass and collapse to a black hole. Long bursts are thought to be the result of a massive supernova (a hypernova) in which material not blasted into space falls back on the central neutron star forcing it over the Chandrasekhar limit and again creating a black hole. In each case, the black hole is surrounded by an accretion disk of material orbiting it that ultimately emits the radiation we see.

Models of the energy emission agree that in order to be so bright, both types of bursts must involve beams of energy emerging from the central source. The jet from the accretion disk satisfies this requirement. Thus a burst's energy is only emitted over a small fraction of the sky, allowing the total energy observed to be consistent with the neutron-star and hypernova models of the energy production.

7.6 Pulsars

See Maoz §4.4

7.7 Black Hole Formation

We saw last time that, in a Type II supernova, the core collapses to nuclear densities, and "bounces," creating a shock wave that (eventually) propagates outward through the star, blowing it out into space. The details remain unclear, but the outcome (the supernova) is not

However, not all of the mass of the star is expelled. Some of it remains behind and falls back onto the newborn neutron star. Since the material is rotating, it forms an accretion disk, possibly leading to a gamma-ray burst as the matter works its way inward onto the central object. In many cases, the total accreted mass will push the neutron star over its maximum mass, causing it to collapse into something even more extreme—a *black hole*. It is tempting to conclude that all stellar black hole formation is accompanied by a gamma ray burst, but few researchers would (yet) go quite that far.

However it occurs, it seems clear that the neutron star can be pushed over the edge and must again collapse. Let's explore some consequences of that possibility.

7.8 Black Holes—a Newtonian View

Although black holes are decidedly non-Newtonian objects, let's begin by looking at then with Newtonian eyes.

Escape Speed
Imagine a test particle moving in the gravitation field of a point mass M. As discussed previously, if the particle is at distance r from the mass, its energy (per unit mass) is $E = \frac{1}{2}v^2 - GM/r$, so the escape speed, corresponding to $E = 0$, is given by

$$v_{esc}^2 = \frac{2GM}{r}.$$

We can then ask under what circumstances will the escape speed equal the speed of light—obviously an important value, since, if light can't escape, then nothing else can too. We find that this is the case when

$$r \equiv r_s = \frac{2GM}{c^2} = 3 \text{ km} \left(\frac{M}{M_\odot}\right),$$

a radius called the *Schwarzschild radius*. As we will see, our flawed Newtonian reasoning in fact leads to the correct expression.

We need Einstein's General Relativity to go much further, but first let's just note another important Newtonian result.

Tidal Acceleration
The acceleration at distance r from the mass is

$$a = -\frac{GM}{r^2}.$$

Imagine you are falling into the black hole, feet first. You'll be in free fall, so your actual acceleration isn't something you'll notice. You'll be weightless. However, you may be acutely aware of the *differential acceleration* between your head and your feet. Your feet are closer to the black hole, so the acceleration there is greater than at your head. The difference between the two tends to stretch you out in the radial direction. It is called a *tidal acceleration*, since this differential force due to the Moon and the Sun is responsible for also tides on Earth.

We can easily calculate the tidal force by differentiating the above expression for the acceleration. To first order,

$$\begin{aligned}\Delta a &= \frac{da}{dr}\,\Delta r \\ &= \frac{2GM}{r^3}\,\Delta r,\end{aligned}$$

where Δr here is your height h. Let's take an extreme stance and calculate the tidal acceleration just as you reach the Schwarzschild radius by plugging in $r = r_s$. The result is

$$\Delta a_s = \frac{c^6 h}{4G^2M^2} = 1.1 \times 10^9\, g\ h(\mathrm{m}) \left(\frac{M}{M_\odot}\right)^{-2}.$$

Since the human body can't withstand differential stresses exceeding $10 - 20\,g$, this is obviously bad news for any adventurous astronaut. However, note that the tidal acceleration Δa decreases as the square of the mass. For a sufficiently massive black hole (other than the huge scary looking object looming below your feet), you might not even notice the tidal acceleration as you reached the Schwarzschild radius (see Homework 5).

In Homework 6 you will see that the tidal field also has a transverse component that squeezes infalling matter laterally almost as strongly as it is stretched radially. Hence the term "spaghettification."

7.9 Gravitational Redshift

A photon loses energy as it moves out of a gravitational potential well according to a very simple rule. If the frequency deep in the well is ν and the frequency higher up, at higher potential $\Delta\phi$ (> 0), is ν', then

$$\frac{\nu'}{\nu} = 1 - \frac{\Delta\phi}{c^2} < 1.$$

This reduction in the frequency of a photon as it climbs out of a gravitation well is called *gravitational redshift*. It doesn't matter whether the gravitational field is uniform ($\Delta\phi = gh$) or due to a point mass ($\Delta\phi = GM(1/r - 1/r')$), or anything else.

The gravitational redshift has something more fundamental to tell us. If we imagine that our rising photon defines a clock (think, number of crests passing us per second), so the "tick" is $\Delta t \propto 1/\nu$, then (if again the prime refers to the measurement higher in the potential well)

$$\frac{\Delta t}{\Delta t'} = 1 - \frac{\Delta\phi}{c^2} < 1.$$

Thus, the upper observer sees the lower observer's clock run slow. This isn't just some sleight of hand due to the peculiar behavior of photons. Clocks really do run slower in a deep gravitational potential well. And if clocks run slow and the speed of light is constant, then our standard rulers also must change with depth in the well. Something strange happens to spacetime in a gravitational field.

7.10 The Equivalence Principle

Einstein's key conceptual breakthrough in his journey to General Relativity had a very Newtonian origin. When we write down Newton's second law, we say

$$F = ma,$$

and when we write down Newton's law of gravitation for the force at distance r due to some object of mass M, we say

$$F_{grav} = -\frac{GMm}{r^2}.$$

But there is no reason at all why the two m's in these equations should be the same. It is not obvious that the *inertial mass*, which determines a body's acceleration in the second law, and the *gravitational mass*, which determines the same body's coupling to the gravitational field of another object, should be the same.

In fact, repeated experiments over several centuries have failed to find any difference between these two masses, so we have to conclude that they are in fact the same. But that has to be another postulate of Newtonian gravity. Within Newton's theory, there's no fundamental reason why it should be so. Quite a coincidence!

If the inertial and gravitational masses are the same, then it follows that the acceleration of an object in a gravitational field is

$$a_{grav} = F_{grav}/m = -\frac{GM}{r^2},$$

which is *independent* of m. More generally, the acceleration of a test particle in a gravitational potential ϕ is $-\nabla\phi$, where ϕ is related to the density ρ of gravitating matter by *Poisson's equation*

$$\nabla^2\phi = 4\pi G\rho.$$

Again, the acceleration is independent of the mass of the particle.

Einstein recognized that the equality of these two masses had fundamental significance. Instead of them being equal to experimental uncertainty, he proposed the *equivalence principle*, which states that they are in fact exactly equal. In that case, gravity is no longer a force. It is an *acceleration* that affects all objects equally—a property of spacetime. The curvature of spacetime follows when we incorporate this global acceleration into the framework of special relativity.

Special relativity deals with the relative motions of frames of reference moving with respect to one another. It tells us that clocks run at different rates in different frames, and measuring rods have different lengths. Applying these ideas to motion under gravity, where nearby observers are accelerating and the properties of clocks and rods are changing continuously with location, it is a short step to concluding that what Newton called gravity is really how we perceive a curved spacetime.

An alternative statement of the equivalence principle is found in Einstein's famous "elevator" thought experiment. Imagine you are in an elevator with no windows and no way of communicating with the outside (everyone's nightmare!). You are weightless, with no acceleration, like astronauts in space. Suddenly you feel a force between your feet and the floor of the elevator. Has some large gravitating mass appeared under your feet? Or has the elevator acquired a rocket engine that is accelerating you in the opposite direction? It's hard to tell without looking outside.

Einstein broadened the equivalence principle to assert that *no* local (i.e. inside a small region of space) physical experiment can distinguish between a gravitational field and an accelerating frame of reference. If the elevator were large enough, you might be able to measure tidal effects, due to the variation of the gravitational field over the extent of the elevator, but setting that aside, the assertion is that gravity and acceleration are indistinguishable. That is the heart of General Relativity.

The elevator experiment has an interesting corollary. Newton's theory didn't say much about light and gravity, but if a light beam crosses the elevator (of width w) as it is accelerating upward with acceleration g, then in the time it takes to cross the elevator, $t = w/c$, the light must move downward a distance $\Delta y = -\frac{1}{2}gt^2 = -\frac{1}{2}w^2/c^2$. The light beam in the accelerating beam appears curved. Therefore the same must be true in a gravitating field. Light is deflected by gravity.

7.11 The Metric

See Maoz §4.5

General relativity is a geometric theory of gravity. It relates gravity to the curvature of spacetime via the *Einstein equation*

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu},$$

where G is a 4×4 tensor, the *Einstein tensor*, describing the curvature of spacetime and T is another 4×4 tensor, the *stress-energy tensor*, describing the distribution of matter and energy. Don't worry too much about what a tensor is. It's a generalization of a vector, and we aren't going to have to deal with the details in this course. Just think of it as a 4×4 matrix in a 4-D spacetime. The left hand side of the Einstein equation is entirely to do with the geometry of spacetime. The right side is entirely about matter and energy in that spacetime. Particles move on geodesics (curves of minimum length) in spacetime—straight lines in a Euclidean space, or great circles on the surface of a sphere. The connection between matter and spacetime was succinctly summarized by relativist John Archibald Wheeler:

Spacetime tells matter how to move;
matter tells spacetime how to curve.

The geometric side of the Einstein equation consists of derivatives of the *metric tensor* $g_{\mu\nu}$, which defines the basic geometric properties of spacetime. The notion of a metric may sound abstract, but it is something you have seen many times before. It is simply a rule for calculating the distance between two neighboring points in a space. For example, in a two-dimensional Euclidean space, the squared distance (ds^2) between points (x, y) and $(x + dx, y + dy)$ is given by the theorem of Pythagoras:

$$ds^2 = dx^2 + dy^2.$$

In three dimensions, the result is

$$ds^2 = dx^2 + dy^2 + dz^2.$$

In spherical polar coordinates r, θ, ϕ, the same statement becomes

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2,$$

and so on. The metric tensor is nothing more than the coefficients of the "dx^2" coordinate pieces. The theory of *differential geometry* relates the metric to the geometrical properties of the space. For example, if we took the previous expression and confined it to the surface of a sphere, with $r =$ constant, or $dr = 0$, we would obtain

$$ds^2 = r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2.$$

Application of differential geometry to just this metric would immediately reveal that the two-dimensional space under study was curved, and that its curvature was the same everywhere—a sphere.

The *Minkowski metric* in special relativity incorporates time into a spatially flat space:

$$ds^2 \;=\; -c^2 dt^2 + dx^2 + dy^2 + dz^2.$$

In General Relativity, we generalize this further and write $dx_0 = dt,\ dx_1 = dx,\ dx_2 = dy,\ x_3 = dz$ and

$$ds^2 \;=\; \sum_{\mu,\nu=0}^{3} g_{\mu\nu}\ dx_\mu dx_\nu.$$

The indices conventionally range from 0 to 3, where 0 is time. Thinking of $g_{\mu\nu}$ as a matrix, we can see that all the previous metrics are diagonal, which is very common but not required. As a matrix, the Minkowski metric is

$$g_{\mu\nu} \;=\; \eta_{\mu\nu} \;=\; \begin{pmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

In general, $g_{\mu\nu}$ is not diagonal, although it is always symmetric.

7.12 The Schwarzschild Metric

See Maoz §4.5

Einstein published his theory of General Relativity in 1915. In 1916, Karl Schwarzschild published a solution for the metric due to a point mass—the equivalent of the Newtonian 1-body problem. The *Schwarzschild metric* is

$$ds^2 \;=\; -\left(1 - \frac{2GM}{rc^2}\right) c^2\,dt^2 + \left(1 - \frac{2GM}{rc^2}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2.$$

Here, r, θ, and ϕ are spherical polar coordinates, and t is time as measured by an observer at infinity (we called it t' earlier). Note that $2GM/rc^2 = r_s/r$. The interval ds can also be written as $ds^2 = -c^2 d\tau^2$, where τ is the *proper time* of a particle, that is, time measured by a clock moving with the particle.

For a clock at rest, $dr = d\theta = d\phi = 0$, so

$$d\tau \;=\; \left(1 - \frac{2GM}{rc^2}\right)^{1/2} dt,$$

so, expanding for $GM/rc^2 \ll 1$,

$$\frac{d\tau}{dt} \;\approx\; 1 - \frac{GM}{rc^2} \;=\; 1 - \frac{\Delta\phi}{c^2},$$

as before. Note that, in this terminology, $d\tau$ was Δt previously, and dt was $\Delta t'$.

The *gravitational redshift* is

$$\frac{\nu(r)}{\nu(\infty)} \;=\; \frac{dt}{d\tau} \;=\; \left(1 - \frac{2GM}{rc^2}\right)^{-1/2}.$$

In our earlier terminology, the observed frequency is $\nu' = \nu(\infty)$, so

$$\frac{\nu'}{\nu} \;=\; \left(1 - \frac{2GM}{rc^2}\right)^{1/2} \;\approx\; 1 - \frac{GM}{rc^2} \;=\; 1 - \frac{\Delta\phi}{c^2},$$

again as before. Note that the observed $\nu' \to 0$ as $r \to r_s$. Light emitted from the Schwarzschild radius is infinitely redshifted.

7.13 Light Rays

According to relativity, light rays move along *null geodesics*, with $ds^2 = 0$. For example, in a simple 1-D Minkowski space,

$$ds^2 \;=\; -\,c^2dt^2 + dx^2,$$

so if $ds = 0$ we have

$$dx^2 \;=\; c^2dt^2,$$

or

$$\frac{dx}{dt} \;=\; \pm\, c,$$

and

$$x \;=\; x_0 \pm ct,$$

describing light rays in a flat space.

In the Schwarzschild metric, we set $ds = 0$ for a light ray and assume radial motion, so $d\theta = d\phi = 0$, to find the radial coordinate speed of light

$$\frac{dr}{dt} \;=\; \pm\, c\left(1 - \frac{2GM}{rc^2}\right) \;=\; \pm\, c\left(1 - \frac{r_s}{r}\right).$$

Note that for $r \gg r_s$, $dr/dt = \pm c$, as expected. However, $dr/dt \to 0$ as $r \to r_s$. The Schwarzschild radius represents an *event horizon* from which light (or any information) can't escape.

Now let's consider infalling matter with speed high enough that we can regard it as relativistic, and hence described by the above equation:

$$\frac{dr}{dt} \;=\; -\,c\left(1 - \frac{r_s}{r}\right).$$

If the matter starts at radius r_0 at time $t = 0$, this differential equation is easily solved (Homework 6.2) to find

$$\begin{aligned} ct \;&=\; (r_0 - r) - r_s \ln\frac{r - r_s}{r_0 - r_s} \\ &=\; -r_s \ln(r - r - s) \;+\; \text{constant} \end{aligned}$$

for $r \sim r_s$. Hence, as the matter approaches the event horizon, $r = r_s$,

$$r - rs \;\sim\; e^{-ct/r_s}.$$

Not only is radiation infinitely redshifted (as seen from infinity) by the time it reaches the event horizon, it takes an infinite amount of time (as seen from infinity) to get there. However, although we haven't shown it here, from the point of view of the infalling matter, the event horizon is reached and crossed in an (all too finite) amount of proper time.

7.14 Hawking Radiation

Although particles and radiation moving in the Schwarzschild metric cannot escape from $r \leq r_s$, in 1974 Stephen Hawking found a novel semi-classical mechanism that effectively allows a black hole to *radiate* energy into space. The Schwarzschild metric is a classical vacuum solution to the field equations, and the classical vacuum is empty, but the quantum vacuum is a very different thing. It constantly creates *quantum fluctuations*, consisting of virtual particle-antiparticle pairs. These violate conservation of energy, but that is allowed by the uncertainty principle $\Delta E \Delta \geq h/4\pi$, so long as the particles recombine and the books are balanced in a sufficiently short period of time.

Hawking realized that, if this process happens just outside the event horizon of a black hole, it is possible for one of the virtual particles to cross the event horizon and be lost forever, allowing the other to escape to infinity (slides8.2/07). The result is that the black hole radiates matter and energy into space. What's more, Hawking showed that the radiation has a blackbody spectrum, with temperature

$$T = \frac{hc^3}{16\pi^2 GMk} = 6.2 \times 10^{-8}\mathrm{K} \left(\frac{M}{M_\odot} \right)^{-1}.$$

Because of Hawking radiation, a black hole slowly radiates its energy away into space like any other object of radius r_s:

$$\frac{d}{dt}(Mc^2) = 4\pi r_s^2 \sigma T^4 \propto M^{-2}.$$

This equation is easily solved to show that the mass of the black hole goes to zero (the black hole *evaporates*) after time

$$t_{evap} = 2 \times 10^{67} \text{ yr} \left(\frac{M}{M_\odot} \right)^3.$$

This is a very long time, except for low-mass *quantum black holes* of masses less than $\sim 10^{-19}\, M_\odot$, which might have formed in the very early universe, and would be exploding today. None has yet been observed.

7.15 Observational Evidence for Black Holes

Several lines of reasoning have been used to argue for the existence (and observations) of black holes.

1. The X-ray source Cygnus X-1 (slides9.1/01) appears to be associated with a B-type star called HDE 226868. Stellar evolutionary models suggest that HDE 226868 has a mass of around $30\, M_\odot$. The X-ray emission is thought to come from an accretion disk (Maoz §4.6) around a companion compact object. The two orbit one another with a 5.6 day period. The calculation of the compact object mass is similar to HW1.4, except that here the inclination is uncertain. Best estimates indicate that the compact object has mass $\sim 15\, M_\odot$—too massive to be anything else but a black hole. Note that the argument "it's too massive to be anything else" comes up a lot in this topic. About a dozen known X-ray binaries have properties that suggest they contain a black hole.

2. X-ray emission from several globular star clusters also suggest that they may contain one or more black holes of stellar origin.

3. Studies of the orbits of stars in some globular star clusters suggest that the clusters may harbor *intermediate-mass black holes* of masses $1000 - 20000\, M_\odot$, at their centers (slides9.1/02). The

observations have been controversial, but in some cases the massive black holes can't be ruled out. They may have formed by collisions of massive stars early in the cluster's lifetime, or when a large stellar black hole collided with and accreted multiple other stars in the cluster.

4. Many galactic nuclei, including our own, are thought to contain *supermassive black holes.* Slides9.1/04-07 show observations of the orbits of about 20 so-called *S stars* orbiting near the center of the Milky Way. They have been intensively studied for more than 2 decades, and the orbital elements of many of the stars have been determined. One (S2) is well on its way around its second orbit since it was discovered in 1992. The orbits are consistent with all the stars orbiting the same massive compact object (called Sgr A*), of mass $4 \times 10^6\, M_\odot$.

5. Finally, the recent LIGO observation of gravitational radiation also represents strong evidence for binaries consisting of two black holes. Such radiation, even from very violent events, are extremely weak. When they pass, they cause a tiny change in the distance between two points in space — a fractional shift of just 10^{-20} for two $10\, M_\odot$ black holes merging in a galaxy 10 Mpc away. Yet remarkably, using interferometric techniques, that shift can be measured. The Laser Interferometer Gravitational-Wave Observatory (LIGO) uses an interferometer with 4-km long arms, each traversed thousands of times by laser beams, to increase the total distance to the point where the change can be detected. The LIGO instruments are located in Hanford, Washington, and in Livingston, Louisiana. Subsequently, they have been joined by the Italian VIRGO instrument, and numerous additional instruments are under construction.

   The first gravity-wave observation, GW150914, was made on September 15, 2015. It revealed a merger between black holes of masses $29\, M_\odot$ and $36\, M_\odot$. The waves (slides8.2/11) observed by the two LIGO detectors agree uncannily well with the waveform predicted for this event on the basis of numerical simulations (slides8.2/12). It should be noted that numerical simulations of these events are extremely important, as they allow the LIGO/VIRGO team to interpret the observational data and extract important properties from the signal, such as the individual black hole masses, their spins, and the mass of the merger product. In the case of GW150914, the mass of the product was $62M_\odot$ — three solar masses worth of energy were emitted in the form of gravitational radiation!

   Since then, 9 additional black-hole mergers have been detected, and 1 neutron-star-neutron star merger (GW170817). Most of the black holes have masses significantly above the range previously predicted for the result of stellar evolution, a puzzle whose resolution remains controversial in the community. The GW170817 event was observed in both gravitational and electromagnetic radiation, the first example of a new field of astrophysics called *multi-messenger astronomy.* With ongoing improvements in sensitivity, observers soon expect to see hundreds of merger events per year. As with the arrival of all new observational windows, we can expect our knowledge of the universe to increase in important and unexpected ways as a result of these new observations.