For more details, see

http://www.astro.umd.edu/∼miller/teaching/astr688m/ lectures 6–9 on general relativity, and lectures 10–13 on black holes. For a much more comprehensive introduction, the classic text "Gravitation" by Misner, Thorne, and Wheeler is still the best in my opinion.

General Relativity

When massive stars die in a supernova, they leave behind a compact remnant: either a neutron star or a black hole. For such objects, Newtonian gravity is no longer accurate enough, and general relativity must be used. Note that black holes, in particular, have had and continue to have a major influence on the universe. It is thought that radiation from the first supermassive black holes may have reionized the universe, accelerating structure formation. In addition, it now appears that light from active galactic nuclei (powered by black holes) may compose 1/3 to 1/2 of all light ever emitted in the universe.

We'll start off with qualitative statements of some of the most important concepts of general relativity, then sketch some of the mathematical formulation.

(1) As in special relativity, space and time are both considered as aspects of spacetime. However, whereas in special relativity spacetime is "flat" (in a sense to be defined later), in general relativity the presence of gravity warps spacetime.

(2) The natural motion of objects is to follow the warps in spacetime. "Matter tells space how to curve and space tells matter how to move." An object that is freely falling (i.e., following spacetime's warps) does not "feel" force, meaning that an accelerometer would measure zero. The path of a freely falling particle is called a geodesic.

(3) The only "force" in this sense that can be exerted by gravity is tidal force. That is, if an object has finite size, different parts of it want to follow different geodesics, and these deviate. Geodesic deviation is the GR equivalent of tidal forces.

(4) Because of this deviation, global spacetime is not flat and there is no coordinate transformation that will make it look flat everywhere.

(5) However, THE most important principle of GR is that in a sufficiently small region of spacetime (small spatial scale, small time interval), the spacetime looks flat. This means that there is a local inertial frame that can be defined in that small patch of spacetime. In that local inertial frame, all the laws of physics are the same as they are in special relativity (electrodynamic, hydrodynamic, strong+weak nuclear, ...)!! This is called the *equivalence principle* (elevator analogy). It means that in practice one of the best ways to do calculations in GR is to do them in the local inertial frame and then use well-defined transformations between the local and the global frame.

(6) All forms of energy gravitate. In the Newtonian limit, rest mass is overwhelmingly the dominant component, but in ultradense matter other forms can be important as well.

(7) A rule of thumb about the importance of GR uses the quantity GM/rc^2, where M is the total mass-energy (possibly including negative contributions from potential energy) and r is some distance. The fractional deviations from Newtonian answers are often of order GM/rc^2 (not always true, e.g., deflection of light in weak-field limit is twice Newtonian). For the Sun, $GM/c^2 \approx 1.5$ km.

Now let's go into some more detail. Here I'm following Shapiro and Teukolsky, chapter 5, which summarizes some of the astrophysically relevant aspects of general relativity (much of this is taken verbatim from them).

When we say "matter tells space how to curve", what does that mean? It means that the geometry of spacetime is altered by the presence of mass. We therefore need to think a bit about geometry.

To start, let's stay with Euclidean geometry. Suppose we have nice, flat, Cartesian coordinates in two dimensions (x and y). Say we have two points A and B, whose x coordinates differ by dx and whose y coordinates differ by dy. **Ask class:** what is the distance between them? Distance is given by the Pythagorean theorem, so if the distance is ds then $ds^2 = dx^2 + dy^2$. **Ask class:** what happens to the distance if we keep the points fixed but change the coordinates, say to $x' = x - y$ and $y' = x + y$? Distance is unchanged. What about if we have polar coordinates, $r = \sqrt{x^2 + y^2}$ and $\theta = \tan^{-1}(y/x)$? Again, distance is unchanged. That means that the distance is independent of the coordinate system. **Ask class:** what happens to that conclusion in three dimensions? Nothing, it's still true. **Ask class:** if you consider special relativity, is the distance ds between two points independent of how fast you move? No. Apparent length contraction occurs in the direction of motion, so this is *not* invariant. We have to do something else. **Ask class:** any suggestions? Must involve time in some way.

In special relativity, one considers *events*, which are labeled by the time and location of occurrence. Two nearby events in spacetime have, in special relativity, an interval (or pseudo-distance)

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 \tag{1}$$

as measured by a local inertial observer (i.e., one in a Lorentz frame). There are different choices for the sign of this distance, but that's what I'll use. Unlike the Euclidean distance $dx^2 + dy^2 + dz^2$, this interval ds is not only coordinate-independent, but is a Lorentz invariant, meaning that all inertial observers will agree on ds even though they may disagree on dt, dx, dy, or dz. **Ask class:** what is the interval measured if the two events are the locations and times of a given photon? Zero. In fact, $ds^2 = 0$ if and only if the world line of the particle in question is *null*, so that it's a photon or some other massless particle. If we use $x^0 = ct$, $x^1 = x$, $x^2 = y$, and $x^3 = z$, we can write this more compactly as

$$ds^2 = \eta_{\alpha\beta} x^\alpha x^\beta \ , \tag{2}$$

where $\eta_{\alpha\beta} = (-1, 1, 1, 1)$ and there is an implicit sum over repeated indices (the Einstein summation convention). The tensor η is the Minkowski tensor, and in general this way of defining distances is a metric approach (no, not cm, kg, etc.!). The Minkowski space is a flat space.

In addition to the distance ds, you can define a *proper time*, which is the time measured by an observer who travels from one event to the other. That is, if you have a particular path from event A to event B, you can integrate $d\tau = (-ds^2)^{1/2}/c$ along that path to find the time the observer would measure.

It is of course possible to define other coordinates, such as spherical coordinates, where the diagonal elements are not -1 or 1, and it is also possible to define noninertial coordinates. If the space is Minkowski, however, it is always possible to find a coordinate transformation that brings it back to the Minkowski form (and, therefore, the space is still flat; "If you call a tail a leg, how many legs does a dog have? Four. Calling a tail a leg doesn't mean it is one." A. Lincoln).

Is it always possible to bring a spacetime into the Minkowski form by some change of coordinates? As an analogy, think of the surface of a sphere. Topologically, it is different from a flat plane. Put in a coordinate transformation way, you can't take the metric $ds^2 = R^2 d\theta^2 + R^2 \sin^2\theta \, d\phi^2$ and turn it into $ds^2 = dx^2 + dy^2$ by a change of coordinates. Another way to think of the difference is to imagine how a vector would change if it were transported parallel to itself in a closed curve. On a flat plane it will return to its original orientation, but on the surface of a sphere it won't in general.

Analogously, in curved spacetime the metric can be expressed as

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta \tag{3}$$

but *no* change of coordinates can bring it back to the Minkowski form. However, ds^2 is still an invariant. As in special relativity, the proper time is $d\tau = (-ds^2)^{1/2}/c$.

The principle of equivalence means that we can choose any event in spacetime to be the origin of coordinates, and at that point can diagonalize the metric to the Minkowski form. We can also eliminate the first derivatives of the metric components, so that the interval can be written

$$ds^2 = \left[\eta_{\alpha\beta} + \mathcal{O}(|x|^2)\right] dx^\alpha dx^\beta \ . \tag{4}$$

In this system of coordinates, all laws of physics proceed (within this small region of spacetime!) as if no gravity were present.

Once again, there is a nice analogy to a sphere. If we are on a sphere, then at a given point there is a unique flat plane (the tangent plane) that is the best approximation to the sphere at that point, meaning that small deviations are $\mathcal{O}(|x|^2)$.

The principle of equivalence guides us on how to transform laws of physics in the presence of gravity. The precise transformations are well-defined but more involved than we want at this point. Suffice it to say that partial derivatives ∇_α need to be replaced by something that involves curvature terms, called the "covariant derivative". S+T liken this to the fact that when we use non-Cartesian coordinates, e.g., spherical, the gradient isn't simply $\partial_r A^r + \partial_\theta A^\theta + \partial_\phi A^\phi$. You have to take into account what the coordinates are doing as well.

When particles fall freely, i.e., under the influence of gravity alone, they feel no force. In general relativity, as in Newtonian physics, this is represented by the equation $a^\alpha = f^\alpha/m$, where a^α is the α component of the acceleration, f^α is the α component of the force, and m is the rest mass. Something similar exists for zero rest mass particles (e.g., photons). The acceleration involves partial derivatives, so one must make the curvature corrections.

One of the consequences of general relativity is that there is a gravitational redshift. However, one can derive the redshift from energy conservation. Suppose you have two particles, an electron and a positron. You drop them in a gravitational field, some distance, and at the bottom have them annihilate to produce two photons. Using mirrors, you reflect the photons up to the original place, where you convert them into an electron and positron again to restart the cycle. **Ask class:** from energy conservation, what condition must hold? The photons must lose as much energy going up as the electron and positron gained in falling down. From this, one can figure out the weak-field redshift of gravity: $\nu(\infty) \approx (1 - GM/rc^2)\nu(r)$ in a spherically symmetric gravitational field around mass M. The exact solution in general relativity is $\nu(\infty) = \sqrt{1 - 2GM/rc^2}\nu(r)$.

When writing equations in general relativity one usually has factors of G and c all over the place. As a nod towards tired writing hands, the community has adopted a solution similar to that in QM, where $\hbar = 1$ is a standard substitution. Here we can set $c = G \equiv 1$; when we need units, we can figure out how many factors of G and c go in that way. Incidentally, since gravity is so weak, G is very poorly determined experimentally. One consequence of this is that $GM_\odot$ is known to about eight significant figures, whereas $M_\odot$, in grams, is known to only about four significant figures!