General Relativity & Cosmology

It might be said that *special* relativity begins with Einstein's deceptively simple postulate that the speed of light is the same in all frames. *General* relativity, a more general theory, begins with another of Einstein's 'simple' postulates--that 'inertial mass' and 'gravitational mass' are the same. In introductory physics we learn Newton's universal law of gravitation: The force between objects of mass m and M separated by r is $F = \frac{G M m}{r^2}$, where G is the universal gravitational constant. Taking m to be an object near the earth's surface, with M and r being the earth's mass and radius, the force on m is:

$$F_{grav} = \frac{G M_{earth}}{r_{earth}^2} m \qquad 9.8\text{m/s}^2 \times m \qquad\qquad (1a)$$

In introductory mechanics we also learn the second law of motion: that the acceleration of an object of mass m is proportional to the net force on the object and inversely proportional to m.

$$\mathbf{a} = \mathbf{F_{net}} \frac{1}{m} \qquad\qquad (2a)$$

On the face of it, these *two* properties of the mass m are entirely different. There is no fundamental reason that the property governing how hard gravity pulls on a given object should have anything to do with the property governing its reluctance to accelerate when a net force is applied! (After all, the net force might be purely electrostatic, unrelated to gravity.) Accordingly, it might be safer to use an m_g in the first equation, signifying a gravitational property, and an m_i in the second, signifying an inertial property.

$$F_{grav} = 9.8\text{m/s}^2 \times m_g \quad (1b) \qquad\qquad a = F_{net} \frac{1}{m_i} \quad (2b)$$

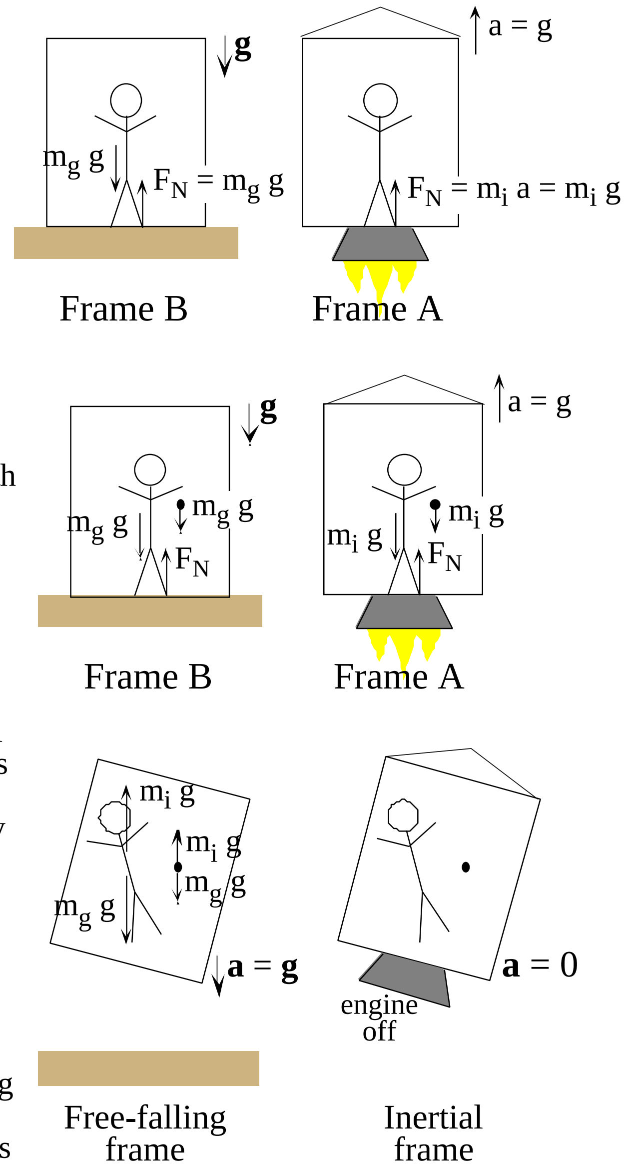
Now consider what happens when an object is dropped from shoulder height. In this case the net force is simply the gravitational force. Thus,

$$a = \left(9.8\text{m/s}^2 \times m_g\right) \frac{1}{m_i} = 9.8\text{m/s}^2 \frac{m_g}{m_i}$$

Were m_g and m_i truly different properties, there is no reason why we could not have an Object 1 with $m_i = 1.1\ m_g$ and an Object 2 with $m_i = 0.9\ m_g$, in which case they would accelerate at different rates. This is certainly not what we expect. In fact, the equivalence of m_g and m_i has been experimentally verified to better than one part in 10^{12}. It is natural to *assume* that the properties are the same. But with the unthinking assumption comes a certain blindness.

Albert Einstein was the first to discover the startling possibilities that arise by *postulating* that m_g and m_i are the same. In particular, if $m_g = m_i$, it would be impossible to determine whether one is an inertial frame permeated by a uniform gravitational field, or in a frame in which there is no field, but which accelerates at a constant rate. Suppose Bob stands in a closet on earth, Frame B. He is in a frame of reference that is inertial and in which there is a uniform gravitational field of $g = 9.8\text{m/s}^2$ downward.[1] Anna is in an identical closet, Frame A, but one that is out in space, far from any gravitational fields. By means of a rocket engine, Anna's closet is accelerating in a straight line at 9.8m/s^2. For Bob to remain stationary, the

[1] Actually, the earth revolves, so the frame does accelerate, and the field is not exactly uniform, since all field lines point toward the earth's center and so are not exactly parallel. But it is inertial and of a uniform field to a reasonably good approximation.



Principle of Equivalence

floor must push upward on his feet with a force whose magnitude is equal to the downward force, $m_g \times 9.8\text{m/s}^2$. For Anna to accelerate along with her rocket-powered closet, the floor must push ‘upward’ on her feet with a force sufficient to give her an acceleration of 9.8m/s^2. By the second law of motion, this force is $m_i \times 9.8\text{m/s}^2$. Now, if m_g and m_i are equal, the forces in the two frames would be equal, and would provide no clue to distinguish whether an observer is in Frame A or Frame B. The normal force is only the simplest example of something one might use; the fact is that no mechanical experiment would be able to distinguish the frames. In the linearly accelerating Frame A, all things appear to be affected by a downward force just as they are in Frame B: the floor must push ‘up’ on objects; ‘dropped’ objects to appear to accelerate downward (because, once let go, they do not accelerate along with the frame). All these effects could be attributed to a fictitious ‘inertial force’ $-m_i$ **a** opposite the direction of acceleration. *Provided that m_i equals m_g*, this force would mimic a gravitational force m_g **g** in all respects.

By the same token, no mechanical experiment would be able to distinguish a frame that is accelerating in ‘free-fall’ in a uniform gravitational field from one that is inertial and without a gravitational field. An observer in the inertial frame would see all objects floating or moving at constant velocity, because no forces act. An observer in the free-falling frame would also see objects seemingly moving at constant velocity (though all would actually be accelerating with the frame), because the gravitational force m_g **g** is exactly canceled by the ‘inertial force’ - m_i **a**.

Einstein’s customary leap forward was to postulate that *all* physical phenomena, not just mechanical ones, occur identically in a frame accelerating in gravitational free-fall as in an inertial frame without gravity. No experiment could distinguish the frames. Accordingly, he generalized the concept of an inertial frame by defining a ‘locally inertial frame’: a frame that is falling freely in a gravitational field (which includes ordinary inertial frames--no gravity, no acceleration--as a special case.)[2] With this definition, Einstein’s fundamental postulate of general relativity, known as the **Principle of Equivalence**, is:

The form of each physical law is the same in all locally inertial frames

This postulate is only the *basis* of general relativity. Just as the Lorentz transformation equations follow from the postulates of special relativity, a mathematical framework follows from this postulate. Unfortunately, general relativity theory is too sophisticated to discuss quantitatively here (it involves the mathematics of **tensors** and **differential geometry**). Nevertheless, some of its astonishing predictions can be understood qualitatively just from the Principle of Equivalence. Three have attracted particular attention: (1) gravitational red-shift; (2) the deflection of light by the sun; and (3) the precession of the perihelion of Mercury.

[2] The acceleration of freely falling objects in a frame would not be the same unique value unless the gravitational field is uniform. Therefore, the frame must also be small enough that any non-uniformities in the field within it are negligible.

Gravitational Red-Shift

According to the equivalence principle, light *emitted* at one point in a gravitational field will have a different frequency if *observed* at a different point. We see this by analyzing not a fixed light source and observer in a gravitational field g, but the *equivalent* case of a source and observer in a frame accelerating at g (without gravity). In figure 1.1(a), a source in an accelerating frame emits a wavefront when the frame has zero speed. An observer a distance H 'above' the source observes the wavefront after a time $^{H}/_{c}$. But by this time the observer is moving 'upward' at speed $v = g\,^{H}/_{c}$. Thus, the light was *emitted* in a frame that moves at velocity $g\,^{H}/_{c}$ *away* from the frame in which the light is observed. According to the observer, the light will be red-shifted. The observed period will be longer than the source period by the factor $1 + {}^{v}/_{c} = 1 + {}^{gH}/_{c^2}$.[3]

$$T_{obs} = \left(1 + {}^{gH}/_{c^2}\right) T_{source}$$

Inverting to obtain frequencies,

$$f_{obs} = \left(1 + {}^{gH}/_{c^2}\right)^{-1} f_{source} \qquad \left(1 - {}^{gH}/_{c^2}\right) f_{source}$$

The observed frequency is lower by the factor $^{gH}/_{c^2}$. The fractional change in frequency is given by:

$$\frac{\Delta f}{f_{source}} \equiv \frac{|f_{obs} - f_{source}|}{f_{source}} = \frac{gH}{c^2}$$

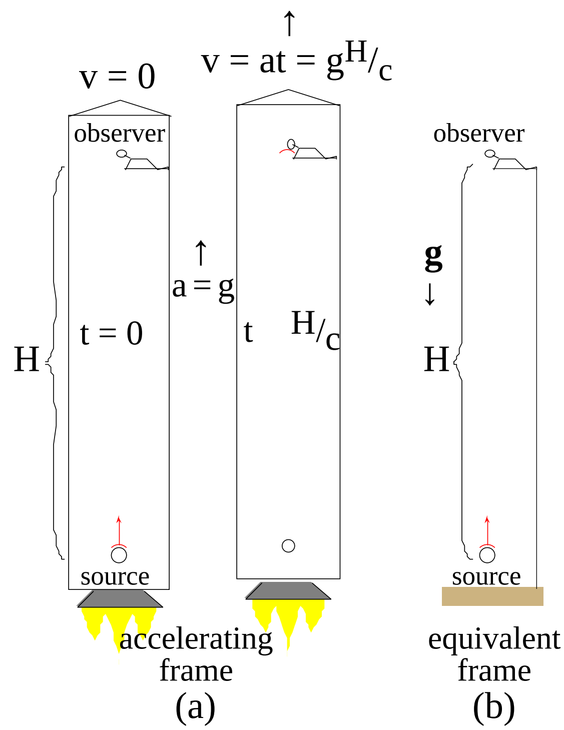


Figure 1.1

Since a gravity-free frame accelerating at g is equivalent to a fixed frame in a gravitational field g, the same conclusion must apply to the latter case, figure 1.1(b). Thus, as light moves upward, its frequency becomes smaller and its wavelength longer. Two interesting conclusions follow. First: This is a time dilation effect, but it has nothing to do with being in different reference frames. In figure 1.1(b) there is no relative motion between source and observer. Assuming light of frequency 5×10^{14} Hz (600nm), as one second passes at the source, 5×10^{14} wavefronts are emitted. It might be said that 5×10^{14} distinct events in the life of the source pass. But in a given second, the observer receives fewer wavefronts, because he observes a smaller frequency. To witness all 5×10^{14} events in the life of the source, the observer has to wait more than one second. Relative to the observer, the source, 'deeper' in the gravitational field, is aging slowly.

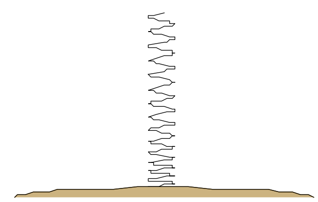


Figure 1.2

Does time really pass more slowly on the surface of the earth than at some altitude above? The weakness of most gravitational fields ($^{gH}/_{c^2} \ll 1$) had long made verification of gravitational time dilation difficult. But modern 'atomic' clocks of extremely high precision have changed this. Atomic clocks use as their basic unit of time the very short period of certain atomic oscillations. By comparing the frequency of such a clock on the surface to that of one aboard a high-altitude rocket, strong confirmation of Einstein's equivalence principle has been obtained.[4]

---

[3] This is a 'lowest-order' result, correct only when the acceleration and distance traveled are small.

[4] Nevertheless, a philosophical question remains: Is time really passing slower on the surface, or is it just that gravity interferes with the behavior of light sources, clocks, *etc.*? In introductory physics we learn that time is based upon an accepted standard unit. Today we *define* time in terms of the period of apparently regular oscillations of cesium-133 atoms. There does seem to be a predictable relationship between this unit of time and the rate of all other microscopic physical processes, such as those that govern the affairs of the human body. If all processes occur at a slower rate on the surface than at altitude, as it appears that they do, it certainly seems reasonable to say that time passes more slowly at the surface. Of course, if we cannot equate 'real' time with the rate at which physical processes occur, the question remains unanswered.

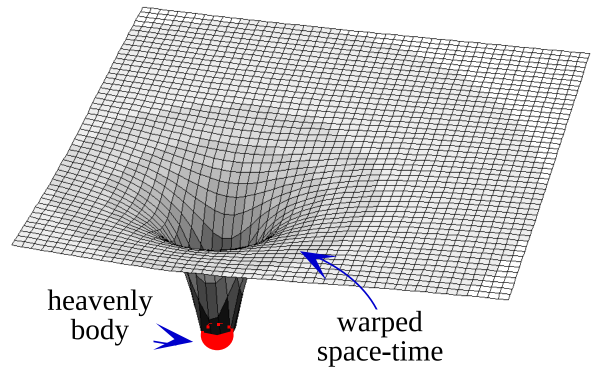


Figure 1.3

Before we discuss the second conclusion, let us venture a little further with the present train of thought: If a gravitational field somehow 'warps' time intervals, even when there is no relative motion, why not space intervals? Indeed, one of the tenets of general relativity is that a massive heavenly body warps space-time nearby. Representing warped space-time in three dimensions is difficult. It is easier in two dimensions, in which space is *area*. Figure 1.3 shows a massive heavenly body disturbing the regularity of a two-dimensional space. Inhabitants of this two dimensional universe *expect* all 'cells' to be of equal area. We 'outside observers', however, can see that the cells near the heavenly body are really larger--but only from our 'extra-dimensional' viewpoint. Similarly, the warpage of real three-dimensional space is not apparent to human beings; we are creatures of our space of three dimensions, and are not able to stand back and view our universe on *four*-dimensional axes. Nevertheless, even 'reduced-dimensional' views such as figure 1.3 help to provide a qualitative understanding of some features of general relativity, as we will soon see.

The second conclusion arising from the gravitational red-shift is that gravity has an effect on light! Now light has no mass--its energy is all kinetic--so it must be that gravity pulls on forms of energy besides mass (or internal) energy. As it moves away, the light's kinetic energy must decrease as potential energy increases. But this leads to the question: How can the universal law of gravitation $F = \frac{G\,M\,m}{r^2}$ be used to account for the effect of gravity on light without an *m* for light? The answer is that Newton's 'universal law' is really a special case, correct only when the gravitational field is very weak-- the so-called 'classical limit'.[5] In this limit, the effect on light is negligible and there is no need to consider an *m*. For a strong field, however, general relativity comes into play and Newton's law is replaced by a different view--warped space-time. Near a massive heavenly body, the regularity in space and time intervals is disturbed. Light 'naturally' changes frequency as it passes through.

Most heavenly bodies produce only a very small gravitational red-shift. They are not dense enough to warp nearby space-time significantly. The gravitational red-shift of our sun is only about two parts in 10^6. (The earth's is much smaller still.) It has been measured, though, and agrees quite well with what Einstein's principle of equivalence predicts. On the other hand, it is theoretically possible for an object to be so dense that light simply could not escape its gravitational potential energy at all. We discuss such objects, known as 'black holes', later in the section.

---

[5] Newton's law of gravitation $F = G\,\frac{m_1\,m_2}{r^2}$ bears somewhat the same relationship to General Relativity theory that Coulomb's Law $F = \frac{1}{4\pi\varepsilon_0}\,\frac{q_1\,q_2}{r^2}$ does to Maxwell's equations. Coulomb's law is a special case, correct only for *static* electric fields. An understanding of electromagnetic waves must come from Maxwell's equations. Similarly, *gravitational waves*, one of the predictions of General Relativity, cannot be analyzed via Newton's static law. Great efforts are now being made to detect gravitational waves. We expect that the accelerations of certain immense heavenly bodies send gravitational waves out into space. But they are hard to detect. The difficulty is in sorting out their small effect from the 'noise' inherent in earth-bound experiments.

Deflection of Light by the Sun

We have seen that gravity 'pulls' on light moving directly away from a heavenly body. But what if the light is traveling laterally--does it curve? It must! Since a laterally-moving light beam would appear to curve toward the floor in Anna's rocket-powered closet, it must curve toward the floor in Bob's earth-bound closet.

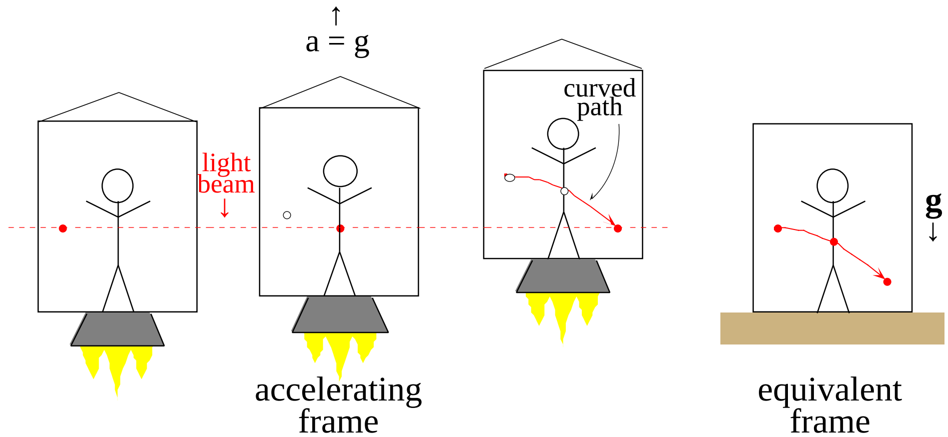


Figure 1.4

Again, however, we need not attempt to reconcile light's curvature with Newton's law of gravitation; rather, the light simply moves in the most natural way given the warped space through which it travels. A guiding 'classical' principle applies even in warped space-time: light always takes the minimum time to travel from one point to another. If we combine this with the idea of warped space-time, it becomes clear why light should 'bend' near massive heavenly bodies. Figure 1.5 shows two possible paths of a light beam originating at one point in space, passing through the warped space-time near a large heavenly body, then observed at another point. The darker path follows what inhabitants of the flat, two dimensional space might *believe* is a straight line, one of the lines in a 'grid' that would be regular if space were not warped. Smugly observing from our greater-dimensional perspective, however, we see that this path is rather long. The other path, though appearing not to follow a straight line from the 'flat perspective', is actually shorter. Since this path is the one that the light actually takes, an observer with a 'flat perspective' sees a beam that curved as it passed near the heavenly body and thus *seems* to have originated at some other point in space.

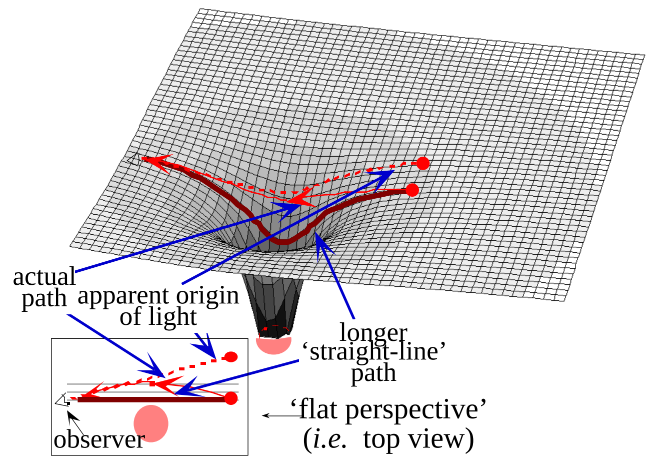
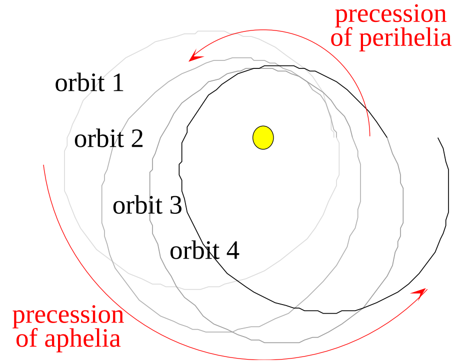


Figure 1.5

As noted, our sun does not warp space much. To have any hope of detecting measurable deflection of light by our sun, observations have to be made of light rays passing very close to it, where its field is strongest. This light of course comes from other stars. Their positions should appear to shift slightly as the sun passes between them and earth. The problem is that the stars would ordinarily be completely obscured by the brightness of the sun. Therefore, such observations have to be made when the sun is 'darkened'--during a solar eclipse. One of the first, during an eclipse in 1919, showed a deflection of about 2 seconds of arc for light barely grazing the sun. As in the case of gravitational red-shift, agreement with the prediction of general relativity is very good.[6]

[6] The warpage of space caused by an immense mass allows for *multiple* shortest distances between two points. Light from a star may curve around opposite sides of an intervening massive object, producing a double-image at earth. Astronomers have indeed observed large galaxy clusters that produce this effect. They are referred to as 'gravitational lenses'.

Precession of the Perihelion of Mercury



The orbits of the planets about the sun are not exactly circular, but slightly elliptical. At one point in its orbit, called aphelion, a planet will be slightly farther than average from the sun, and at another, called perihelion, slightly closer. So long as a system is simply one object orbiting another, it is a direct prediction of classical Newtonian gravitation that the same path in space is retraced indefinitely. But if anything interferes with the simple interaction, the orbit will *precess*; the points of aphelion and perihelion progressively 'creep' around in a circular fashion. One source of interference is the presence of other heavenly bodies. The slightly elliptical orbit of Mercury, for instance, precesses due to the perturbing effects of the other planets. Now Newtonian theory can account for these effects, predicting a rate at which precession should occur. But, much to the consternation of early astronomers, the predicted rate did not agree precisely with observation. The problem was that Newton's law of gravitation is correct only for a weak gravitational field. A strong field might well cause observation to deviate from the 'classical' expectation. If any planet shows deviation, it should be Mercury, for it orbits where the sun's field is strongest. Using general relativity, a correction to the classically expected precession rate of Mercury may be calculated. The result of 43 seconds of arc per century is in good agreement with observation.

We have seen how successful general relativity has been in explaining some slight discrepancies between classical expectation of actual observation. Though this evidence took decades to accumulate, its weight was irresistible, leading to near-universal acceptance of Einstein's 'new' approach to gravitation. Let us now explore an important topic in which general relativity plays a significant role.

Cosmology

Cosmology is the study of the behaviors of heavenly bodies, both individually and collectively. The topic is invariably coupled to gravitation, because gravitation is the only fundamental force in nature that is both long-range and only attractive. Thus, it is by far the most important force in the evolution of and interaction between huge heavenly bodies.[7] Much of cosmology can be understood via classical Newtonian gravitation, but certain behaviors require general relativity.

Stellar collapse is the term describing the fate of individual stars. The energy source of stars is nuclear fusion. Early in the 'life' of a star, the enormous gravitational attraction of its mass is balanced by the constant generation of energy, which would tend to scatter the material--a constant size is maintained. When a star's fuel begins to 'burn out', the gravitational attraction predominates and gravitational collapse begins. If the star is not much more massive than our sun, the final result is a **white dwarf**, and ultimately a cold, dead chuck of matter typically no larger than the earth! For stars several times the sun's mass, the extra gravitational pressure is able to force protons and electrons to combine, forming neutrons and tiny particles known as **neutrinos**. The neutrinos carry away a huge amount of energy in a cataclysmic explosion known as a **supernova**. What remains is a cold **neutron star** of fantastic density--a typical radius is only tens of kilometers. The density of a star collapsed so small would be roughly 10^{17}kg/m^3, approximately equal to that of the atomic nucleus, 10^{14} times that of lead! For even larger stars there is a third possible fate--becoming a **black hole**, and here is where general relativity comes into play.

---

[7] Large objects tend to be electrically neutral, and so do not exert significant electrostatic forces on one another. The other fundamental forces (the "weak" and the "strong") are significant only for microscopic particles at very close range, typically 10^{-15}m and less.

One of the most startling features of the theory of general relativity is the possibility of singularities in space.[8] A black hole would be such a singularity. A singularity would occur if a body of mass M becomes so compact that its radius drops below the so-called **Schwartzchild radius**: $r_S = \frac{2\,G\,M}{c^2}$. (Assuming that the critical radius depends only on the mass of the body and the fundamental constants G and c, simple dimensional analysis yields this equation to within a multiplicative constant.) The singularity would be separate from the universe as we know it. It would simply be a silent 'hole' in space. Moreover, its gravitational field would be so strong that not even light could leave it, thus the name '*black* hole'. Its presence would be betrayed only by its warpage of space, *i.e*. its external gravitational field, and its ability to gobble up things from the outside. A black hole would be uncharted territory, not on the map, and an unlucky passerby would be in for quite a ride. It should be noted that not even the atomic nucleus--trillions of times denser than the densest of ordinary materials--comes close to qualifying as a black hole. The diameter of a typical nucleus is 10^{-15}m, but its Schwartzchild radius would be roughly 10^{-52}m. Only under conditions such as the tremendous gravitational pressure in immense heavenly bodies might a black hole occur. Since black holes would be light-years away and dark, with gravitational effects serving as our only clues, we expect them to be difficult to detect. Though there are several candidates, there is as yet no conclusive proof of their existence.

The Evolution of the Universe

The most apparent change in the universe as a whole is that it is expanding. Furthermore, the way in which it is expanding suggests that it began with a 'bang', from the explosion of an initial, supremely compact body of energy. Let us view the universe as consisting of galaxies, each galaxy being a distinct frame of reference.[9] *If* all galaxies began moving away from the origin at the same time, after an arbitrary time they would be spread out according to their speeds; speed should increase linearly with distance from the origin. After one unit of time, those moving a unit distance per time would be one unit away, those moving two units of distance per time would be two units away, and so forth. Perhaps surprisingly, this would also appear to be the case *no matter which* frame-of-reference / galaxy an observer happened to occupy. Thus, if the universe did begin with a 'big bang', any galaxy (including our own, of course) should serve as a valid base from which to observe, and the speeds of the galaxies should increase linearly with their distances from that frame of reference.

---

[8] Mathematically, a 'singularity' is a point at which a discontinuity or divergence occurs.
[9] The reason for this arbitrary choice is that galaxies are the smallest units that move *more or less* independently. Stars within a given galaxy definitely do not. Our sun, for instance, *orbits* about the center of its galaxy, the Milky Way.

This is just what we observe! This evidence comes from the Doppler shifts in the light spectra from distant galaxies; the farther a galaxy is from earth, the greater is the red-shift in its spectrum, which implies a greater recessional speed. We find that the speed increases linearly with distance:

$$v = H_o\, r$$

This relationship is known as Hubble's Law, and the constant H_o as **Hubble's constant**. Since on object moving a distance r at constant speed v would travel for a time $^r/_v$, and this is the same constant $^1/_{H_o}$ for all galaxies, $^1/_{H_o}$ is often referred to as 'the age of the universe'. Modern doppler evidence gives a value of approximately $2{\times}10^{-18}s^{-1}$ for H_o, or an age of the universe on the order of ten billion years. Unfortunately, Hubble's constant is not known with great precision, mostly because it is difficult to be certain of the distances r to faraway galaxies.

Even if distances could be known precisely, a guess at the age of the universe based on Hubble's constant is approximate at best. Among the problems is that galaxies have really *not* moved out independently at constant velocity since an initial 'big bang'. All forms of energy in the universe share a gravitational attraction, which necessarily decreases the expansion rate. Indeed, one of the most compelling questions in cosmology is whether this gravitational attraction is sufficient to cause the universe at some time in the future to stop expanding and then reconverge. The answer hinges on the energy density in the universe. General relativity predicts a **critical density** ρ_c. If the actual density ρ is less than ρ_c , the universe will expand forever; if greater, it will eventually reconverge. The big question is: How does ρ compare to ρ_c? Unfortunately, the energy density in the universe is not known well enough to answer the question. One of the most important factors is the amount of 'dark matter'. It is fairly easy to estimate the energy density of things we can see (*i.e.* stars), and this suggests that $\rho < \rho_c$. But there is much out there that we cannot see. This dark matter may take several forms: intergalactic dust, collapsed galaxies, black holes, background neutrinos, or unknown particles or radiation. The density of neutrinos in space is difficult to determine, but of particular interest. Though once thought to be massless (like light), recent experiments indicate that neutrinos do have a small mass, and this may greatly affect the value of ρ. Much research has to be done before we can say whether the universe will or will not expand forever.