Simple Special and General Relativity

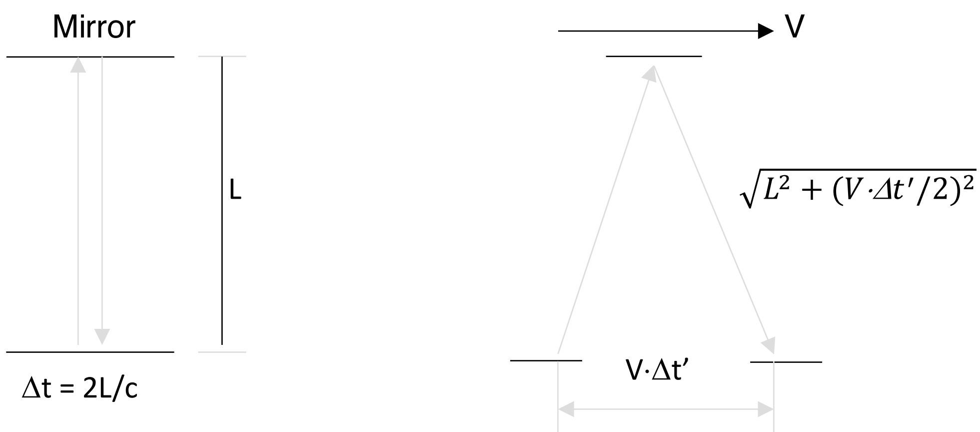
Virgil Elings

My introduction to special relativity was as an undergraduate engineering student. We were told that the speed of light was constant for all observers traveling at different velocities with respect to one another and then spent the rest of our time manipulating Lorentz transformations. Due to a combination of bad teaching and bad learning, I thought that moving clocks ran slow because it took time for the information to get back to the observer. Of course, this idea was wrong, but it was what I was left with. I will try to stay away from manipulating formulas but try to derive things as we go along (learning by thinking, I call it). I will try to do almost everything in special and general relativity using time dilation and length contraction. The aim is to get a good feel for relativity, without the complexity of tensors and all that fancy stuff the experts use. The serious reader needs to get out a pen and pencil and fill in some of the math I have left out.

Einstein's assumptions for special relativity were that in systems traveling with constant velocity with respect to each other (no acceleration), they are equal and the physics inside each of the systems is the same. He assumed that the speed of light in free space, with respect all observers, is a constant. This assumption was not well received because it really screws up space and time. This assumption came from Maxwell's equations in which one solution showed a traveling electromagnetic wave traveling at a velocity which seemed to be close to that of the speed of light, but it didn't say in what system this speed was measured. There were also experiments which showed that the velocity of light from a star was the same whether you were traveling, in the earth orbit, toward the star or away from the star.
General relativity is complicated. We will see how far we can get using what we learn in special relativity.
Other than some classical mechanics, everything will be done using only these assumptions. Relativity is not my field and so I'm afraid there's a lot of approaches that I haven't looked at and don't know about.

Special relativity

Let's start with the standard optical clock. This clock consists of an emitter/detector which emits pulses which go up to a mirror, spaced a distance L from the emitter, and then back down to the emitter/ detector where a new pulse is sent out, making one tick of the clock. There is an identical clock which is moving off to the right at a velocity V in another reference frame, which we will call the primed frame. So why the mirror? It turns out in relativity if you're going to measure a time interval, it is best to measure the start and end of the interval at the same place in whatever reference frame you are in. Time and space get mixed up so it is better if you can keep space out of it. We will see an example of this shortly.



What we want to calculate is the ratio of the time interval in the primed frame, as seen by an observer in the stationary frame, to the time interval in the stationary frame for one "tick" of the clocks. Now the velocity of light, c, is the same, 3x10^8 m/sec, in both systems whether the system is moving with a velocity to the right or not. In our stationary system, one tick is just 2L divided by c. In the primed frame one tick is clearly going to be longer because the path is now longer than 2L. During the time interval Δ t' (one tick of the moving clock) the detector has moved a distance V·Δt'. The distance the pulse must travel is, therefore, $2\sqrt{L^2 + (V{\cdot}\Delta t'/2)^2}$.

So, $\Delta t' = 2\sqrt{L^2 + (V{\cdot}\Delta t'/2)^2}$/c. With Δt = 2L/c, we get:

$$\Delta t' = \gamma \cdot \Delta t$$

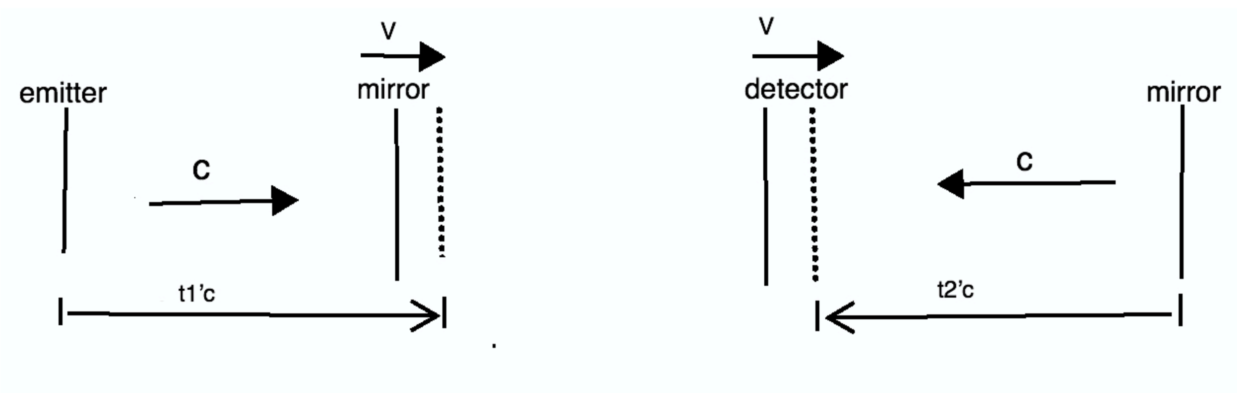
where $\gamma = 1/\sqrt{1 - V^2/c^2}$ a parameter greater than or equal to 1. In our everyday world where velocities are not much greater than a few hundred meters per second, this parameter is essentially 1 and so we rarely see any of this time dilation (GPS is the glaring exception-more on this later). If V$> c$, then the clock (and time) does not work anymore so the restriction is that V $\leq c$.

The takeaway from this is that the clock in the primed (moving) system is ticking slower than the stationary clock; Moving clocks run slow! time is slower in the moving system. By the way, if we were in the primed frame looking at the clock in the unprimed system, we would see it running slow by the same factor. Clocks moving by us run slower than our clock. Now Einstein said that without acceleration or gravity, these reference frames are equal. So, who is moving? The other guy, of course.

Back in my deep dark past I was part of an experiment at the Berkeley Bevatron to measure the lifetimes of the pi plus and pi minus mesons. If we could find a difference in the lifetimes, it would mean that a fundamental theory, CPT, was violated and therefore it would be a big deal. Alas (for us, but not for physics), we measured the lifetimes to be equal. The point of all this is that the pi mesons, which have an average lifetime at rest of 2.6×10^{-8}sec, travelling at the speed of light, without relativity, would only travel about 8 m on the average. But our spectrometer was about 30 m in length over which we measured the decay of the mesons. They were living about twice their “at rest” lifetime as they moved through the lab. Why? Moving clocks run slow and in our case the clock was the meson’s lifetime and γ was about 2.

Now let’s say that a particular pi meson decays halfway down the spectrometer. People in the lab frame and the meson in the meson’s frame both agree where this happens, halfway down the spectrometer. But in the meson’s frame it’s clock is not running slow and therefore the lifetime is the at-rest lifetime. In this case the meson, at best, can only see 8 m of spectrometer go by before it decays. How did it get to the 16 meter mark? The 16 meter mark in the lab frame is only 8 meters in the meson’s at rest frame. From the meson’s point of view, the spectrometer is contracted by about a factor of 2 so that in the meson’s normal at-rest lifetime, half of the spectrometer passes by. Objects in a reference frame moving by a stationary observer are length contracted in the direction of motion by γ i.e. $L' = L/\gamma$. This length contraction is only along the direction of motion. There is no transverse length contraction, as opposed to time dilation which occurs everywhere equally in the moving frame. So, the meson sees a distorted moving space where dimensions along the direction of the moving lab are contracted but the transverse dimensions are not. Time dilation and length contraction go hand-in-hand and the magnitude of both relativistic effects is the same.

With this length contraction idea, let’s go back and revisit our optical clock, but this time we will lay it along the X axis, the direction of motion, so that the pulses travels along the X axis. What we want to show is that length contraction is required in order to make the clock work properly. One might think offhand that, due to the length contraction of the clock, it will actually take less time for the pulses to go to the mirror and back than it does in the stationary frame and time would run faster in the moving frame. If that is true, we are in deep trouble. The rate time passes in the primed frame should not depend on the orientation of our clock.



Let’s break the round-trip of the path, as seen by the stationary observer, into two parts. The first part being from the emitter/detector to the mirror (time t1’) and the second part from the mirror back to the emitter/detector (t2’). In the first part the mirror is moving away from the pulse, which is travelling at velocity c, so the distance traveled is $(L/\gamma + V\cdot t1') = t1'\cdot c$, where the length of the clock has been length contracted to L/γ. We solve for t1′ and get t1′ = L/(γ·(c – V)). In the second part the emitter/detector is moving toward the pulse and so the distance traveled is (L/γ - V·t2′) = t2′·c and so we get t2′ = L/(γ·(c + V)). The total time for one tick of the moving clock is t1′ + t2′ which is 2L/(γc·(1 - V^2/c^2)) = 2Lγ/c. The time for one tick of the stationary clock is 2L/c and so the clock in the moving frame, as seen from the stationary frame, is still running slow by the factor γ. Things are OK, but we get the correct answer only because the moving clock was length contracted in the direction of motion. In the stationary frame the times t1 and t2 are identical, so the time between two events, spaced apart along the x-axis, are different as seen by the two frames. Even events that are simultaneous in one frame will not be simultaneous as seen in the other frame. Here is a simple example. Let's say we synchronize two clocks in the primed (moving) frame, spaced a distance L apart, by placing a pulse emitter halfway between the clocks and sending out pulses of light which start the clocks running when the pulses arrive. In the primed frame the clocks are now synchronized. As seen from the stationary frame, the left clock is running into the pulse and therefore gets set before the right clock, which is moving away from the pulse, and so the left clock measures time which is ahead of the right clock as seen from the stationary frame. If we calculate how much it is ahead, it involves the same calculation we did for our clock which was laying down.

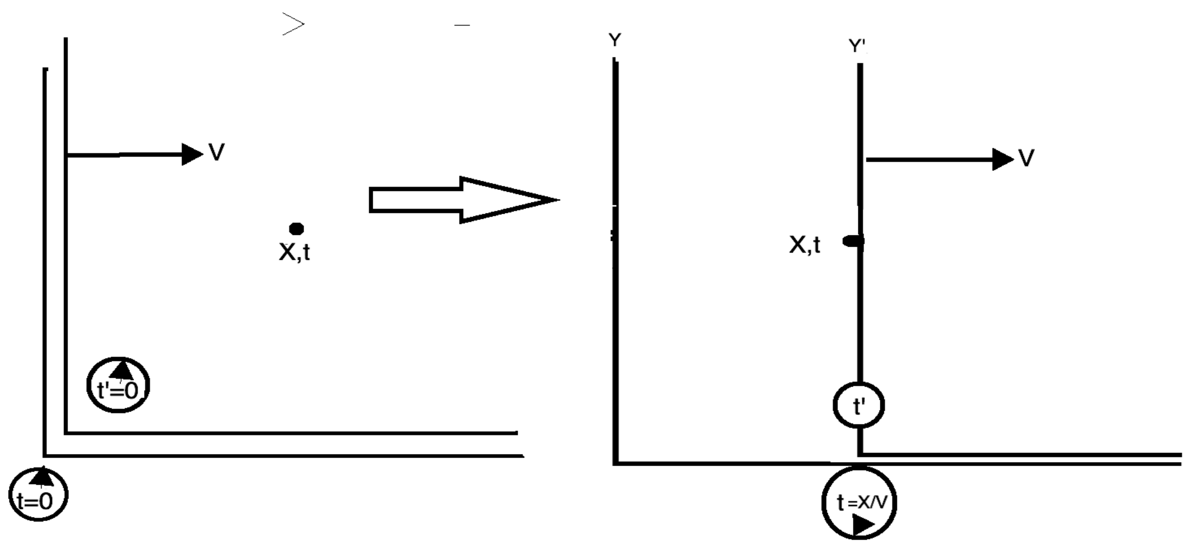
Lorentz Transformations

(I said I was going to stay away from this but some think it is needed)

Again, we consider two reference frames, one the primed frame moving with a velocity V to the right of the stationary frame. We have a point in the stationary frame with coordinates X and t and we now want to calculate the coordinates X′ and t′ in the primed frame. At time zero, all clocks are synchronized in each of the respective frames and set to zero and the origins of both coordinate systems are superimposed.

Length Contraction First we will derive the Lorentz transformations, for transforming the space and time coordinates from a stationary frame to a moving frame, by simply using only length contraction. Classically we have, as seen from the stationary frame, X′ = X – Vt. With relativity, the moving frame is length contracted and, therefore, we see, from the stationary frame, that X′ is now X′/γ . So, X′ = γ(X – Vt), and looking from the primed frame we see an equal relation, X = γ(X′ + Vt′), except that V is now in the other direction. Eliminating X′ between these relations we get; t′ = γ(t - X·V/c^2) and, similarly, t = γ(t′ + X′·V/c^2) . The term X·V/c^2means that the time on the X′ axis, as seen from the stationary frame, depends on the position in the stationary frame. Therefore, the Lorenz transformation for transferring the coordinates in one frame to a moving frame are the following.

X′ = γ(X – Vt) and t′ = γ(t - X·V/c^2)
X = γ(X′ + Vt) and t = γ(t′ + X′·V/c^2)
Y′ = Y



Time Dilation

We will now derive the Lorentz transformations using only time dilation. We already found that it's dangerous to compare the time measured on two clocks in the moving frame that are separated in space, so we will use only one clock in the moving frame. The clocks in the stationary frame all read the same time.

Consider a clock at the origin of the moving frame and count the time on this clock, as viewed from the stationary frame, from time zero until the Y' axis of the moving frame is coincident with the point of interest, X,t. We have, $t' = t/\gamma$ (the moving clock is running slow).
Multiplying the top and bottom by γ, we get:

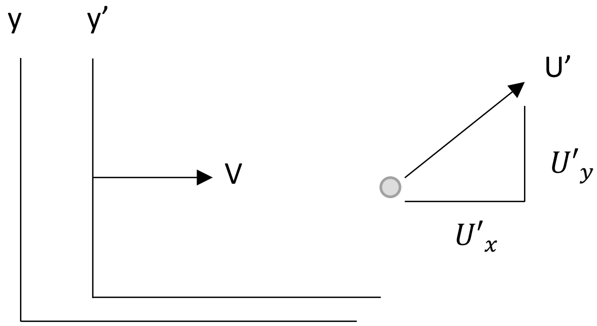
$$t' = \gamma/\gamma^2 \cdot t = \gamma t \cdot (1 - V^2/c^2) \text{ but, } t = X/V \ (V = X/t) \text{ so: } t' = \gamma(t - X \cdot V/c^2)$$

As before, the rest of the transformations follow from this. The Lorentz transformations can be derived from either length contraction or time dilation. A big difference is that length contraction is directional and time dilation is not. In general relativity, we will see that we have to use both to get the right answers, and the directional aspect of length contraction will be a big deal.

Velocity Addition

We now will calculate how to add velocities in relativity. We will do this two ways, one by thinking, and one using the Lorenz transformations. Let's do thinking first.
Consider a particle moving in the primed frame with a velocity U' and with the frame moving with a velocity V in the X direction. What we want to calculate is the velocity U of the particle in the stationary frame ie. we want to add U' and V to get U. This velocity U must satisfy several boundary conditions. Two are that when U' or V increases to the velocity of light(c), U also goes to c and cannot go above it. Also, if U' and V are very small compared to the velocity of light, we

should get the classical result, which is just U' + V. Whatever relation we derive, it must be symmetrical in U' and V.



Consider first only the x component of U', the component in the direction of V. In the low velocity case we have:

U_x = U'_x+V, but as V ⇒ c $U_x = \alpha \cdot (U'_x + c)$ ⇒ c where α is some factor.

α therefore must have a form something like c/$(U'_x + c)$ but similarly must have the form c/(V+ c) when U'_x⇒c. To include all these conditions and the symmetry between U and V, we wind up with α= c/(c + V·$U'_x/c) = $ 1/(1 + V·U'_x/c^2).

So: U_x =(U'_x+V) /(1+V·U'_x/c^2) ⇒c when V, U'_x ⇒c and ⇒U'_x+V when V/c ⇒0

Another way to find U_x is to calculate dX/dt from the two Lorentz transformations; X = γ(X' + Vt') and t = γ(t' + X'·V/c^2) dX/dt = γ(dX'+Vdt')/γ(dt'+dX'V/c^2)
Dividing by dt' top and bottom we get, as before

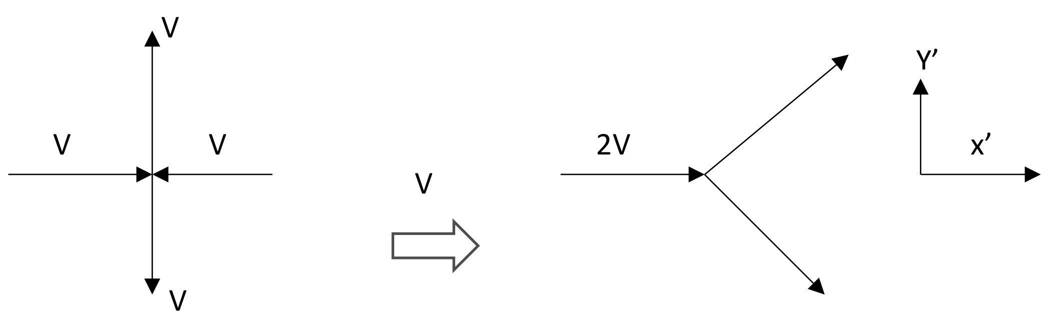
dX/dt = U_x = (U'_x+ V) / (1 + V·U'_x/c^2).

Now consider the Y component of U. You might assume that the velocity in the Y direction should not be affected by a transformation between two coordinate systems which have a relative velocity in the X direction. Let's see how that works out. Consider a particle travelling only along the Y' axis. When transformed to the stationary system, we will assume the Y velocity to be unchanged and the X velocity will just be V (U'_x=0). The resultant velocity will be $\sqrt{U'_y{}^2 + V^2}$. Well, you can see the problem. As V ⇒ c, this resultant velocity is clearly going to be larger than the velocity of light, and that is a not allowed. U_y must go to 0 as V⇒ c. The Y velocity measured in the stationary frame, U_y, must be smaller than the Y' velocity measured in the primed frame, U'_y, to avoid this problem. Let's take U_y to be κ·U'_y and see if we can figure out κ. We have U= $\sqrt{(\kappa \cdot U'_y)^2 + V^2}$. If this is to be c as V goes to c, we have U^2 = $\kappa^2 \cdot U'_y{}^2$ + c^2 = c^2. So κ must go to zero as V goes to c. Let U'_y go to c and the result must be the same. $\kappa^2 c^2$ + V^2 = c^2. Or: κ^2= 1 - V^2/c^2 = 1/γ^2. So, κ = 1/ γ. Now if time were running slower in the primed frame, by the factor γ ,that is, t' = t/γ, then dY/dt would be slower than dY'/dt' (velocity is inverse time). But, we already knew that time runs slower in the moving frame. So, you can derive time dilation, and therefore all the Lorentz transformations, by only insisting that things not go faster than the speed of light.

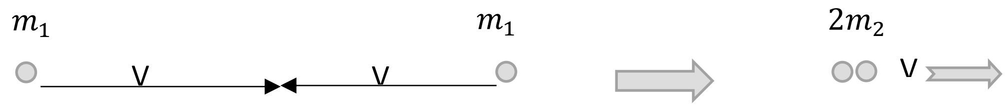
Conservation of Momentum

So far, special relativity has made a mess out of space and time. Let's see if we can save Newton's most famous law of motion, action equals reaction, also known as conservation of momentum. Some people feel that F = ma is his most famous law of motion but, to me, it is just the definition of force, and besides, the force is measured in Newtons! We will find that unless we do something with mass, momentum is not conserved because of our new velocity addition rule.

Consider the collision of two particles of mass m and velocity V travelling in opposite directions, colliding, and then going off at 90° with velocity V. In this frame, momentum is conserved because both the X and Y components of momentum are zero both before and after the collision. Call this the primed frame. We now let this frame move to the right with velocity V and observe the collision in a stationary frame. Is momentum still conserved?



Let's consider the classical case first. In this case we have a particle moving with the velocity 2V striking a stationary particle and then the particles go off at a 45 ° angle with each particle having an X component of velocity V. We have an incoming X momentum of 2V·m and outgoing momentum of mV·2. The total Y momentum is still zero, and so classically momentum is conserved in the collision. Let's now use our velocity addition formula to see what relativity thinks of this. The incoming velocity is <u>not</u> 2V but rather 2V/(1+V^2/c^2) which is less than 2. The X component of momentum after the collision is still 2·mV, and so momentum is ***<u>NOT</u>*** conserved. To make this work we need to have the mass of the incoming particle be larger than either of the masses of the 2 outgoing particles. The likely suspect is, of course, to multiply the rest mass of the particles by γ so that particles with a higher velocity would have a larger mass. This would also be convenient because it would mean that the transverse, (Y), momentum would stay constant when we transform between the moving frame and the stationary frame, the larger mass making up for the smaller transverse velocity. If you do this fudge to the mass <u>of all the mass's</u> by multiplying the rest masses by γ, then you'll find out, if you do the algebra, that in this example the momentum is conserved. Try not to make any mistakes because the algebra is sort of a mess. And be sure that the mass fudge factor involves the TOTAL velocity of the particle. We now want to show, with our correction to the mass, that we can create mass from energy.



Let's take a simpler example, an inelastic collision where we have two particles of rest mass $m_1$travelling with equal and opposite velocities V, colliding with each other and sticking together, making a rest mass 2·m_2. We call the masses different, because they may be different. Momentum is conserved in this frame because both the X and Y total momentum are zero before and after the collision. Let's again move this frame off to the right with a velocity V so that in a stationary frame we now have only one particle moving with a velocity given by relativity of 2V/(1+V^2/c^2) which we call V_{in}. After the collision, we have a particle of mass 2·m_2 moving to the right with a velocity V. Conserving momentum, we have.

$$2\cdot m_1 \cdot V_{in}/\sqrt{1-{V_{in}}^2}/c^{\,2} = 2m_2 \cdot V/\sqrt{1-V^2}/c^{\,2}$$

You can do the algebra (there's a lot of cancellation) and what you end up with is that the <u>rest mass</u> m_2 is NOT equal to the rest mass m_1 but rather:

$$m_2 = m_1/\sqrt{1-V^2}/c^{\,2}$$ this is the REST MASS of m_2, not the moving mass!

We now have a problem. We either give up our idea of conservation of momentum or we need to do something about this thing we call mass(again!). We will stick with conservation of momentum.

Classically the loss in kinetic energy goes into heat of the final particles that are stuck together. Let's see what this has to do with the increase in rest mass. In the stationary system the kinetic energy before the collision is 1/2 $m \cdot (2\,V)^2$ and after the collision is $2\cdot 1/2mV^2$ so the loss in kinetic energy in the collision is, <u>classically</u>, 1/2 $m \cdot (2\,V)^2 - 2\cdot 1/2mV^2$ = mV^2 since the input and output masses are the same.

To first order in V^2/c^2 we have, from above, $2m_2 = 2m_1(1 + 1/2(V^2/c^2)) = 2m_1 + m_1V^2/c^2$, or in English; the outgoing particles (which are stuck together) have a total rest mass equal to the two incoming particle's rest mass + the loss in kinetic energy in the collision divided by c^2. This collision has turned kinetic energy into mass! Let's see what it looks like in the moving frame where it is easier to analyze things. Here we have two particles headed toward each other with the kinetic Energy of 1/2· mV^2. The final kinetic energy is zero. If this is turned into mass, then we get an increase in the rest mass/c^2 of mV^2, the same as see from the stationary frame.

We need to be careful. We may be adding apples and oranges. We combined "mass is energy", a relativistic concept, with kinetic energy, a Newtonian concept. It looks like the fudge factor gamma which we multiplied the mass with to make conservation of momentum work, may be a way to include in the mass, it's energy of motion. Rest energy (mc^2)+ energy of motion = γmc^2 =

E. This certainly works where V/c is much less than one. We get E= $\gamma mc^2 \cong$ (1+ $1/2 \cdot V^2/c^2$) mc^2 = mc^2+$1/2 \cdot mV^2$. It also works at higher velocities.
With total energy E = γmc^2 and momentum, P = γmV, we have that E = Pc^2/V. For a particle travelling at the speed of light (a photon) this gives E = Pc. In the limit of V going to 0 (P=mV), we have E = mc^2, the rest energy. Everything is consistent.

Work, energy and conservation of energy

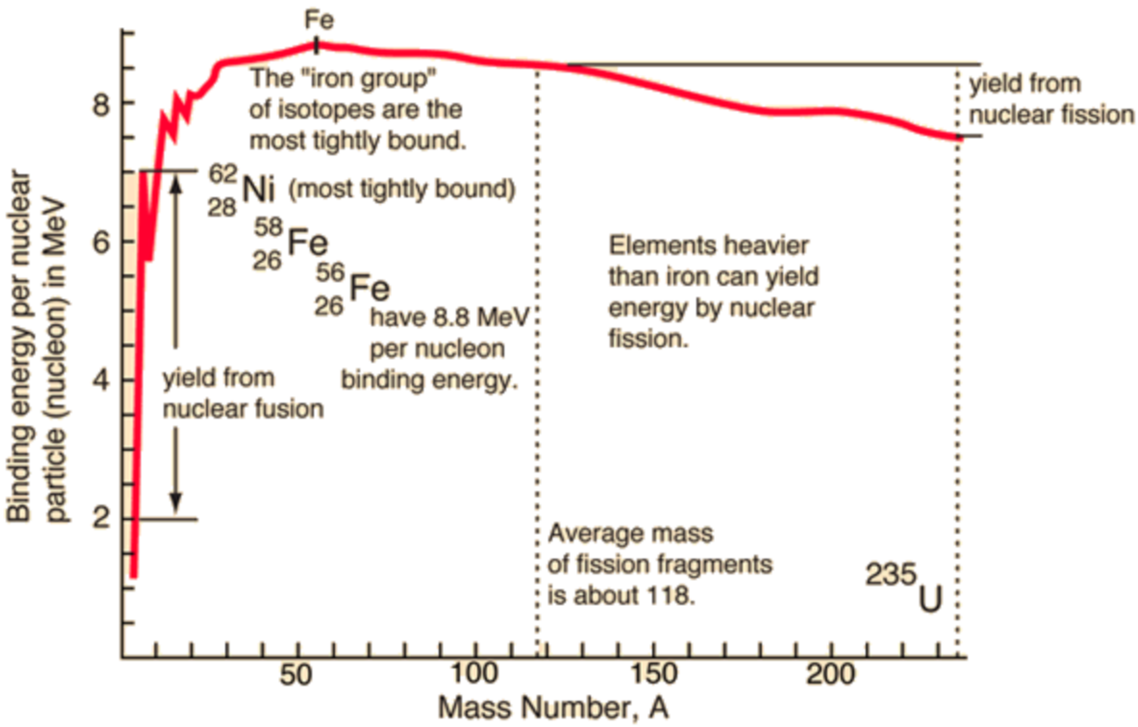
We can derive this mass-energy relationship by using the age-old concept of work (force times distance) equals the increase in energy. We have dE = Fdx = dP/dt$\cdot$ $dx =$ dP/dV$\cdot$ VdV. Integrating from 0 to V we have:

$$\text{Change in Energy} = \Delta E = \int_0^V dP/dV \cdot VdV.$$

The difference with relativity is that P involves γ and so the integral becomes more involved. You can convince yourself that dP/dV is just $\gamma^3 m$ and that the integral, then, is $\gamma mc^2]_0^V$ = γmc^2 - mc^2. The increase in energy from rest to velocity V is the total energy (γmc^2) minus the rest energy (mc^2) .
Since energy is now mass, then our concept of conservation of energy is the conservation of total mass, γmc^2. Let's try this in our inelastic collision example. In the moving frame, the total mass before the collision is just 2γm and therefore that is the mass after the collision, which is just the rest mass of the final particles stuck together. This is the same answer we got when we used conservation of momentum.
Are there other examples of turning energy into mass? Of course, that's what they do at particle accelerators. I mentioned before the measurement of π^+ and π^- meson lifetimes. But if you want to measure those lifetimes, first you need to make the mesons. This is done with a proton-proton inelastic collision, p+p $\Longrightarrow$ p + p + $\pi^+ + \pi^-$, where the first proton comes out of a proton accelerator and the second proton is stationary in a liquid hydrogen target. In the center of mass system, it is clear that the threshhold kinetic energy of each proton must be such to create the rest energy of one pi meson, about 135 Mev (million electron volts). So, in the center of mass we have $\gamma m_p = m_p + m_\pi$ for each proton. Transforming to the laboratory frame where one of the protons is stationary, using our velocity transformation formula, we get that the total energy of the proton from the accelerator is E = $\gamma' m_p$ = m_p+4m_π+2m_π^2/m_p. The kinetic energy required, 4m_π+ 2m_π^2/m_p, is over twice that needed in the center of mass frame. Not very efficient. If we were making a proton- antiproton pair, each with a rest energy of 1000 Mev, instead of a pi meson pair in the collision, then the kinetic energy required in the lab frame would be three times that required in the center of mass. This ratio gets bigger the larger the masses that are created in the collision.
In order to create more massive exotic particles, the problem was solved by making the center of mass frame and the lab frame the same using colliding beam accelerators. The first proton colliding beam accelerator was the Intersecting Storage Rings (ISR) built in Geneva in 1971. This machine stored protons from a synchrotron into two counter-rotating beams and then steered them with magnets to collide with each other, making the lab frame the center of mass frame. The current (2018) most powerful colliding beam accelerator is the large hadron collider (LHC),

also at the CERN lab in Geneva. This machine has two counter rotating beams each with kinetic energy of 6.5 TeV, yes, that is with a T. This accelerator is huge and has a diameter of 27 km. Yes, that's kilometers. This accelerator was used to confirm the existence of the famous Higgs boson which was found to have a mass of 125 GEV, about 125 times the mass of each of the colliding protons.

We've shown how mass can be created with kinetic energy, so how about the reverse where we create kinetic energy from mass. Well, it happens every day in your (non-electric) automobile. In this case carbon and oxygen combine to make CO_2 plus 4.1 eV of energy. The reactants contain a total of 44 nuclei which have a total rest energy of about 44 billion electron volts. It would take a very good scale to notice that 4.1 eV of rest energy is missing in the final product. This reaction doesn't convert much of the rest energy into kinetic energy. These atomic reactions involve energies of the order of electron volts, the binding energy of electrons in atoms. Nuclear reactions, on the other hand, involve energies of millions of electron volts, the binding energy of nucleons in the nucleus. The method would be to take one or more nuclei with a small binding energy and somehow convert them into one or more nuclei that have a much higher binding energy. This binding energy might be thought of as a negative energy. If two things bind together then it takes energy to separate them so, therefore, when they bind together, they give off kinetic energy equal to the binding energy. From the diagram below we see that the most tightly bound nuclei are the iron group and so we could create kinetic energy by taking something very heavy, like uranium, and splitting it to make smaller nuclei or, we could take very light nuclei like hydrogen and deuterium and fuse them together to make heavier nuclei like helium. The first is easier to do but unfortunately some of the products in the fission reaction are very radioactive and last for tens of decades. The second fusion process is very hard to do because you need to get the light nuclei up to a high enough temperature so that their kinetic energy can overcome the repulsive Coulomb force between the positively charged nuclei, so they can react. This is usually done by containing the particles in a magnetic field. I worked one summer at the Oak Ridge National Laboratory and had a chance to tour the fusion facility. They told me that they were just turning the corner to success. It sounded impressive. The problem is, that was 1961. It has been a big corner but progress is still being made.



This brings special relativity to an end. Let's see what we have learned based on just two bold assumptions, one that the speed of light is the same for all observers and two, that reference frames travelling at a constant velocity with respect one another are equal.

1. Clocks moving with respect to us run slow
2. Objects moving with respect to us are length contracted but only in the direction of their motion.
3. Velocities don't add as easily as they used to.
4. Momentum can only be conserved if mass increases as a function of velocity.
5. The mass of a body contains both its rest energy and kinetic energy.
6. Conservation of energy is conservation of mass.
7. Energy can be turned into mass.
8. Mass can be turned into energy.

Quite a bit from simple assumptions. If this all seems strange, stranger things are yet to come.

General relativity

OK, I'll be honest, I don't know much more about the formalism of general relativity than the man on the street, which is essentially next to nothing. It looks like a difficult subject, and so I'm going to stay away from the guts of general relativity. We have come pretty far so far, and it would be a shame to give up now. What we are going to do is to use what we learned about special relativity, along with an analysis of free fall, and see how far we can get. The object is to see if we can derive some of the predictions made by general relativity. The point of the freefall analysis is that it creates a coordinate system where there is no gravity, by canceling the gravity with acceleration, much as in an orbiting satellite. From this free- fall system, using special relativity, one can analyze other coordinate systems, even ones that have gravity. From that frame we can analyze other frames which do have gravity so that we can reproduce some of Einstein's most famous calculations.

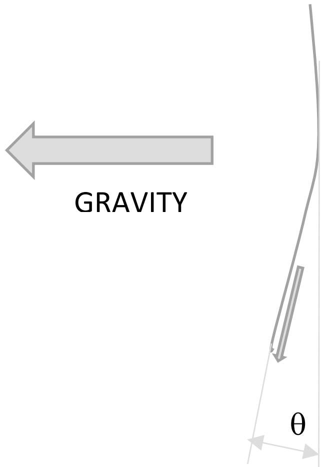
For our simple relativity, we are going to replace gravity with acceleration. Actually, this is not exactly a new concept. In classical mechanics, we go into a rotating coordinate system, such as a merry-go-round, and then replace what was centripetal acceleration in the stationary frame with a mystical force called centrifugal force in the opposite direction. Acceleration replaced with a force. For our simple gravity, we are going to replace a force with acceleration.

The question is, does gravity affect time and distance like velocity does in special relativity? Our approach to general relativity is going to be quite simple so that even I can understand it. Our approach is robust enough that we can reproduce some of Einstein's most famous calculations. Our approach avoids a lot of the formalism in general relativity and so, therefore, our approach is not the end all, but is simple enough that we can get a good feel for general relativity

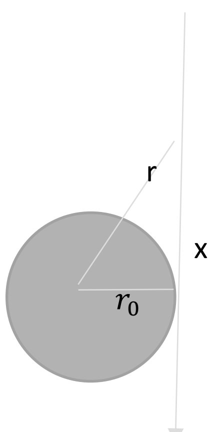
The bending of light by gravity

First a classical calculation

Before we get into the relativity calculation, let's calculate classically the deflection of a massive particle, travelling at the speed of light, passing near the surface of the sun. This is a classical calculation, so a particle travelling at the speed of light is OK. This is a calculation that could've been made by Newton, because he thought light was a particle, but apparently, he never made it.   Consider the imaginary gravitational field shown below which is directed toward the left. As a particle passes through this field, the angle of deflection, for small angles, is simply the amount of horizontal momentum given to the particle by the gravitational force, divided by the total momentum of the particle. $\theta = \Delta P_t$/P where ΔP_t is the change in transverse momentum.



We will now calculate this deflection for a particle passing close to the surface of the sun at a distance r_0 from its center. Since the deflection is small, there is no reason to worry about the shape of the path. We will use just a straight path and calculate the horizontal momentum given to the particle. You can pretend the deflection is there, but it's so small you can't see it in the diagram. We integrate half the path.



$$\Delta P_t = \int_0^\infty dP = \int_0^\infty dP/dt \cdot dt/dx \cdot dx = \int_0^\infty F_t/\mathrm{c}\cdot \mathrm{dx}$$

$$F_t = \mathrm{mMG}/r^2 \cdot \cos\theta \text{ where } \cos\theta = r_0/r$$

$$\text{and } \mathrm{r} = \sqrt{x^2 + {r_0}^2}$$

m is the mass of the particle, M is the mass of the sun and G is the gravitational constant.

$$\Delta P_t/\mathrm{P} = \mathrm{mMG}/\mathrm{m}c^2 \cdot \int_0^\infty r_0/r^3 \cdot dx$$

$$= \mathrm{MG}/c^2 \int_0^\infty r_0/\sqrt[3/2]{x^2 + {r_0}^2} = \mathrm{MG}/r_0 c^2$$

We must multiply by 2 since we only integrated half of the path, so the total deflection is $2\theta = 2\mathrm{MG}/r_0 c^2$

The product $2\mathrm{MG}/c^2$ is called the Swartzchild radius, r_s, and has dimensions of length.

It will come up again and again. The deflection, then, is r_s/r_0, a dimensionless number. This angle, which is only 0.87 arc seconds, or about $4\cdot 10^{-6}$ radians, is <u>half</u> the correct angle for deflection of light by the sun. So, we need to do better.

A word of caution

Given the parameters at hand; the mass of the sun M, the gravitational constant G, the radius of the sun r_0 , and the speed of light c, we need to construct a dimensionless parameter, because the angle of deflection is dimensionless. The only dimensionless parameter that you can create is MG/$r_0 c^2$, or powers of it. So, no matter what your theory, if you keep track of units, you are going to get "something" times this parameter. The important point is, what is the "something"? That is the calculation.

Einstein's first calculation of the bending of light - the dropping of a photon

Einstein's first idea for the bending of light by gravity was in 1911 and the idea was to give the photon a mass and then see how this would affect the bending of light. From special relativity, we know that energy is equal to mass and so we can set the mass equal to E/c^2. The easy way out in the above calculation would be to replace the mass of the particle with E/c^2 and replace the momentum P with E/c.

ΔP_t/P = E/c^2MG/E$\cdot \int_0^\infty r_0/r^3 \cdot dx$ = MG/$c^2 \cdot \int_0^\infty r_0/r^3 \cdot dx$ = MG/$r_0 c^2$

The energy cancels out and you, of course, get the same answer as the classical calculation.

But Einstein took a different approach. He took the energy of the photon to be hf (why not, he discovered that) where h is Plank's constant and f is the frequency of the photon, and the mass therefore is m= E/c^2 = hf/c^2. As the photon falls in the gravitational field, the energy would increase and therefore the frequency would increase. Since the speed of light is constant, this means that the wavelength would decrease. If the frequency is f_0 at infinity then, as the photon falls, we get:

E = E_0 + mMG/r = hf_0 + hf_0 MG/rc^2 = hf_0 (1 + MG/rc^2) where r is the distance to the center of the sun. Similarly we have: $\lambda = \lambda_0$(1 - MG/rc^2) where we have used, and will use a lot, $(1+x)^n$ = 1 + nx for x$\ll$ 1 and for all values of n. Remember, MG/rc^2 is very small, even for r as small as the radius of the sun. Einstein then used the concept that light travels from one point to another on the path of the minimum number of wavelengths and proceeded to calculate the angle of bending. If you have not heard about the path of minimum number of wavelengths, let me give you my thoughts and use some crude diagrams

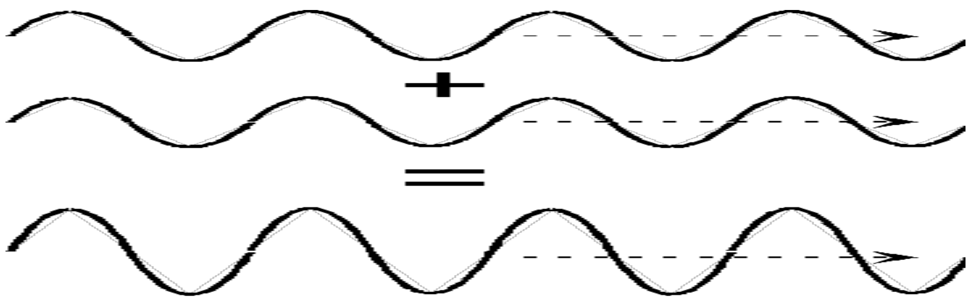
The assumptions:

1. Light is a wave

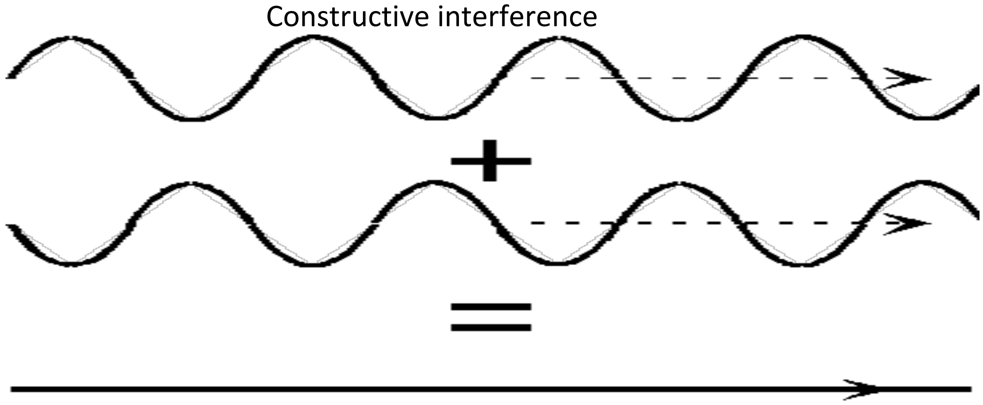
2.Light can take any and all paths from A to B.

I am going to blame this second assumption on Richard Feynman, which is where I think I ran across it (summing over paths). Probably, though, it was Huygens several centuries ago. By the way, you can do essentially all of optics using only these two assumptions. See the Optics section.

Now people usually think light travels in a straight line, and so what is this business about taking any path? But light is a wave and, as you know, two waves out of phase can cancel each other, so it may turn out that a lot of these paths cancel each other.

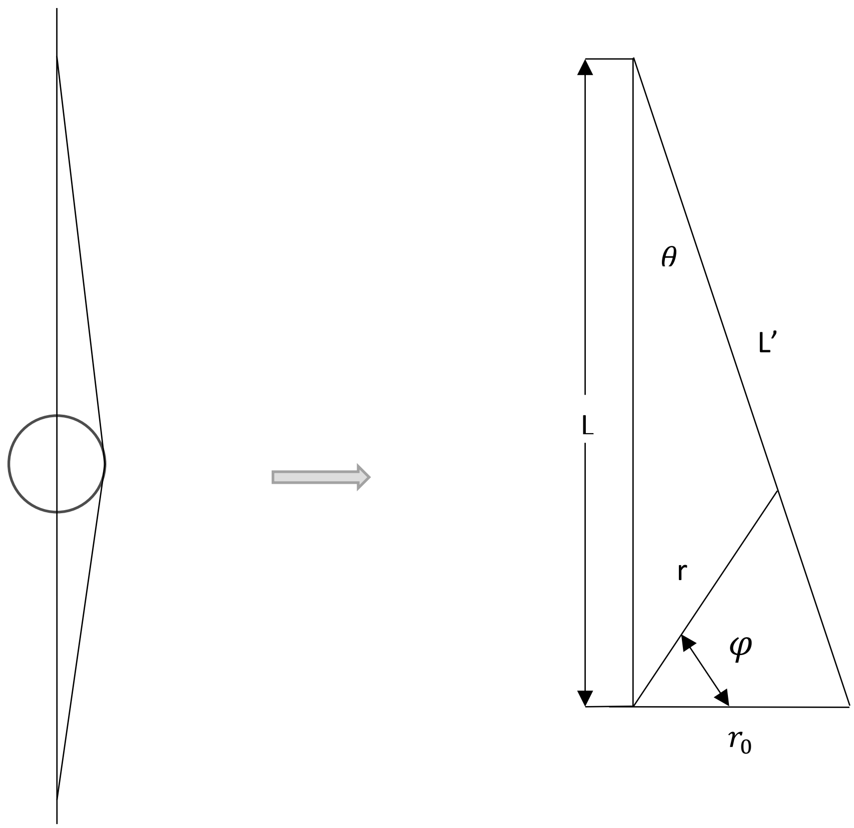


Constructive interference



Destructive interference

If we take the long path from A to B, there is a shorter path, shorter by one half a wavelength, that cancels that path and so on down until we get to the path of the minimum number of wavelengths from A to B. When we get to that point, the paths can no longer cancel each other and around one quarter of a wavelength of the path of minimum wavelengths are the paths on which the light travels. Absent other things, this path will be the shortest distance between A and B, a straight line. If the wavelength varies in space, as in this case, then the path will not be a straight line. A similar thing happens when light goes through a medium with an index of refraction that varies in space and the light does not follow a straight line. Graded index (GRIN) lenses work this way. The “slop” in the path of about 1/4 of a wavelength is what causes all of the diffraction effects in classical optics. See the Optics section of this website.
The path is shown below. The bending is small, so there is no reason to worry about the exact shape of the path. We just use two straight lines but integrate over only half the path.



We now calculate the average wavelength over the path L' by integrating over φ from 0 to π/2, with dL' = r/cos($\varphi-\theta$) dφ. Remember, $\lambda = \lambda_0(1 - MG/rc^2)$

$$\lambda_{avg} = \lambda_0 /L' \cdot \int_0^{\pi/2}(1-a/r)dL' \quad \text{where a is } MG/c^2 \quad \text{This gives:}$$

$$\lambda_{avg} = \lambda_0 - \lambda_0 a/L' \int_0^{\pi/2} 1/\cos(\varphi-\theta)d\varphi = \lambda_0(1 - a/L' \cdot \ln(\tan(\varphi/2-\theta/2+\pi/4))_0^{\pi/2}$$

We need to be careful at the upper limit because the tangent is very large. For small θ, tan(π/2−θ/2) is just 2/θ. So, for λ_{avg} we get:

$$\lambda_{avg} = \lambda_0(1-a/L' \cdot (ln2/\theta - \ln(\tan(\pi/4-\theta/2))$$

The first ln term is large and sensitive to θ whereas the second ln term is small, essentially ln1, which is zero. Notice that all of the "action" is at the upper limit of the integral. To first order we have:

$$\lambda_{avg} = \lambda_0(1 - a\cdot\ln(2/\theta)/L')$$

This average wavelength is less than the wavelength far from the sun, λ_0, because the closer we get to the sun, the higher the frequency of the photon, therefore the shorter the wavelength. Notice that λ_{avg} increases as θ increases.

We now want to vary θ to find the value of θ when the path is the minimum number of wavelengths. This path will favor distances away from the sun because there the average wavelength is longer. The number of wavelengths on the path L' is L'/λ_{avg}. We minimize this by taking the derivative wrt θ and set it equal to zero.

$\#\lambda = L'/\lambda_{avg}$ and $d\#\lambda/d\theta = dL'/d\theta/\lambda_{avg} - \lambda_0 a/\theta/\lambda_{avg}{}^2 = 0$. Notice that the 2 in the ln makes no difference in the derivative.

Now $L' = L/\cos(\theta) = L(1+\theta^2/2)$ so, $dL'/d\theta = L\theta$: to first order we can use λ_0 for λ_{avg} in the denominators

So: $L\theta - a/\theta = 0$ or, since $L\theta = r_0$, $\theta = a/r_0$

total deflection = $2\theta = 2a/r_0 = 2\,MG/r_0c^2 = r_s/r_0$ The Something is 1

Seems like a lot of work to get the same answer, but a step in the right direction. No time dilation or length contraction, though. The dropping of a photon was basically a shot at time dilation. The effect is the same but what is really going on is that the frequency of the photon is increasing as it drops toward the sun because it is entering regions of slower and slower time. The longer the seconds, the more cycles per second. This dropping of a photon is used to describe things like the gravitational red shift and the rate of clocks that are orbiting the earth. In the red shift, many say that the photon loses energy, and therefore gains wavelength, as it comes out of the gravitational pull of a distant star where it was created. The real explanation is that the photon was created in a region of slow time, that is, in the gravitational field of the distant star, and has now entered a region of faster time, where we see it. Shorter seconds means lower frequency.

Now what about length contraction? If we include that, you can guess where the correct answer for the bending of light is going to come from. So, let's do this thing with time dilation and length contraction.

Free Fall Toward a Point Source of Gravity

In special relativity, we dealt with reference frames which we call inertial frames, they travelled with constant velocity and had no acceleration. Einstein wanted to include acceleration.

One of the results of classical mechanics was that the mass attracted by gravity (weight) was exactly the same as the mass involved in collisions, that is, the mass involved in acceleration (F=ma, if you will). Could it be that gravity and acceleration are the same thing? It is a little hard to see how gravity, this mystical force from a massive body, could be the same as acceleration which has to do with motion. They don't seem to be the same and they are not. It is more complicated than that and gravity requires a lot of mathematics to really understand it. What we are going to do is see how far our special relativity can get us into general relativity.

For our simple relativity, we are going to replace gravity with acceleration. Actually, this is not exactly a new concept. In classical mechanics, we go into a rotating coordinate system, such as a merry-go-round, and then replace what was centripetal acceleration in the stationary frame with a mystical force called centrifugal force in the opposite direction. Acceleration replaced with a force. For our simple gravity, we are going to replace a force with acceleration.

The question is, does gravity affect time and distance like velocity does in special relativity? We are going to analyze a reference frame which has gravity from a massive body, like the sun, and then cancel gravity with acceleration so we can analyze the frame from the viewpoint of motion. Getting rid of gravity is not that hard, it's done every day in the space station. The occupants accelerate toward the earth just the right amount to eliminate the force of gravity. They are in freefall. Einstein's idea was an elevator in freefall, but that was before the space station. A reference frame in freefall is an inertial frame, there is no gravity and no acceleration, so we can use special relativity to analyze other frames like the "gravity" frame, the frame with gravity in it.. From the standpoint of this freefall frame, it is the rest of the world that is accelerating past it, moving faster and faster as time goes by. But remember, we have simply replaced gravity with acceleration. As the world moves by, we see that three things happen in that world.

The first is that if there were a row of clocks stretching to the sun, as the clocks go by they are running slow with respect our clock, but the closer the sun gets to us (or we get to the sun) the slower they run. Time dilation causes clocks nearer the source of gravity (the sun) to run slower. The second is that if there was a line of rulers stretching to the sun, we would notice that the rulers would get shorter and shorter the closer we got to the sun due to length contraction. There is no length contraction perpendicular to the direction of motion and therefore space becomes warped. We now have that the circumference around the source of gravity is no longer two pi times the radius, but is greater than that. Space is no longer "flat". The third, which is disturbing to me, is that as the sun gets closer and closer to us, its' mass increases because its' velocity is greater and greater. If mass is energy, where did the energy come from? If we are going to use this approach, though, we are stuck with this increase in mass.

For $V \ll c$, we have $\gamma = 1 + 1/2 \cdot V^2/c^2$.

The freefall velocity, by the time we get to the surface of the sun (or it gets to us), will be equal to the escape velocity from the surface of the sun (freefall run backwards) which is about 620 km/sec, 1/500 the speed of light, so we can use the approximation that $V^2/c^2 \ll 1$. Turning gravitational potential energy into kinetic energy, our freefall velocity squared, starting with zero at infinity, will be 2MG/r at a distance r from the center of the sun. The relativistic corrections, therefore, to time, radial length, and mass of the sun, as seen in the gravitational field are, multiplied by $\gamma = 1 + 1/2 \cdot V^2/c^2$ Relativistic corrections = $\gamma = 1+MG/rc^2$

The distance to the source of gravity, in the gravity frame, is $r(g) = r(1-MG/ rc^2)$. In our examples, r is always much larger than MG/ c^2, so r is essentially equal to r(g) except for a small difference, that is, $r(g) = r - MG/c^2$.The gravity shifts the radial axis and this shift does not depend on the radius. How did we get such a strange result? The difference does depend on r times a factor which varies inversely as r and therefore the effect is a constant, but small, difference between the two radii. We will calculate the bending of light in two ways. The first will be to calculate the bending in the warped system, the gravity system, where the speed of light is constant but time and distance are warped. The second calculation will be to calculate in a Newtonian coordinate system where time and distances are not warped, but where the speed of light is not constant.

In the first case, we will look at the effect of time dilation with no length contraction, and then we will look at the effect of length contraction with no time dilation, and then add the results. You may say that we should multiply the results, but for a small effect, multiplying and adding are the same. For a, b, $c \ll 1$, (1+a)·(1+b)/(1+c) ≅1+a+b-c.

In the second case we will calculate how the variation of the speed of light causes the wavelength of light to vary (just as in the classical case), and will then calculate the bending of the path of minimum number of wavelengths.

The warped system

Time Dilation

With time dilation, the frequency of the light will be greater as it passes nearer the sun. Since the velocity of light is constant, the wavelength of light will be shorter near the sun. We have: $\lambda = \lambda_0(1 - \text{MG}/\text{r}c^2)$ where λ_0 is the wavelength at infinity, just as in the case of dropping a photon. The rest of the calculation to find the path of the minimum number of wavelengths is the same as what we have already calculated and therefor the answer is that the angle of deflection is;

$2\theta = 2\text{MG}/r_0c^2 = r_s/r_0$, deflection caused by time dilation

Length Contraction

Our distant observer now sees dimensions near the sun to be distorted somewhat like the map of the northern hemisphere shown on a flat page. Let's say you had such a map with distorted grid of lines 100 km apart and were asked to calculate the shortest path in km from say Paris to New York. You would calculate, in your flat world, a curved path which favors the northern latitudes where the kilometers in the distorted grid are quite large. Using the dimensions distorted by gravity, we want to calculate the path of minimum length (distance), which is the path that light will take (more on that later). This path will not be a straight line, somewhat like the curved airline routes shown on a flat map of the northern hemisphere. This calculation will be like our time dilation integral, calculating the path of the minimum number of wavelengths, except with one difference, only the radial dimensions are distorted, not the tangential dimensions. The bending is quite small and so we choose the "curved" path to be two straight lines as shown in the previous figure with only θ as a variable.

Consider a meter stick laying along L'. The radial component of this meter stick is m_0·sin(φ–θ) and the tangential component is m_0·cos(φ–θ) , essential all radial component at large r and all tangential component small r. m_0 is the length of a meter stick with no gravity . The stick is very short considering the scale of things, that is, the stick is local. Let's calculate the length of this meter stick with distorted dimensions. The length contraction is only in the radial direction, in the amount a/r where, as before, a=MG/c^2. Therefore:

m = $m_0 \cdot \sqrt{\{sin^2(\varphi - \theta) \cdot (1 - a/r)^2 + cos^2(\varphi - \theta)\}}$
where m is the length of the meter stick in the distorted frame. This all reduces, in first order, to m = $m_0 \cdot (1 - a/r \cdot sin^2(\varphi - \theta))$. We now integrate along L' to get the average length of a meter along that path under the influence of gravity. The integral is similar to what we have already done for wavelength except for the $sin^2(\varphi - \theta)$.

$$m_{avg} = m_0/L' \int_0^{\pi/2} (1 - a/r \cdot sin^2(\varphi - \theta))\, dL'$$
$$= m_0/L' \int_0^{L'} dL' - m_0/L' \int_0^{\pi/2} (a\ \ sin^2 (\varphi - \theta)) d\varphi/\cos(\varphi - \theta)$$

The first term is L' and the second term is, from the integral tables,
$a[\sin(\varphi - \theta) - \ln(\tan(\varphi/2 - \theta/2 + \pi/4)]\,{}_0^{\pi/2} = a[(1 + \theta) - \ln(2/\theta) - \ln 1]$.
So $m_{avg} = m_0[1 - a/L' \cdot (-(1+ \theta) + \ln(2/\theta))]$.

#meters along L' = L'/m_{avg} = L'/$m_0[1 - a/L' \cdot (-(1+ \theta) + \ln(2/\theta))]$
We minimize the number of meters in the grid distorted by gravity, by taking the derivative of # with respect to θ and setting it equal to zero (it is probably easier to maximize 1/# and so we will do that).

$$d/d\theta(m_{avg}/L') = m_0 d/d\theta[1 - a/L' \cdot (-(1+ \theta) + \ln(2/\theta))]/L' = 0$$

as before, with dL'/dθ = Lθ and L$\theta = L'\theta = r_0$, we get, after some algebra

$$r_0 - a(1/\theta - 1)=0$$

Since 1/θ is about 10^5, the 1 is second order, so: $\theta = a/r_0$

total deflection due to length contraction = $2\theta = 2\ MG/r_0c^2 = r_s/r_0$

A few observations. The effects of time dilation and length contraction are long range effects, they drop off only as 1/r. We noticed in the integrals that the major contribution is from the high-end of the integrals whereas at the low-end, with small values of r, there is essentially no contribution. The $sin^2(\varphi - \theta)$ term, which kills off contributions only at the smaller values of r, made little or no difference.(see below, depends on your point of view)
The above calculation becomes very simple if, instead of contracting the radial dimensions, we expand the circumferential dimensions and leave the radial dimensions unchanged. It's the same warped space. The expression for the length of a meter stick becomes:

m = $m_0 \cdot \sqrt{\{sin^2(\varphi - \theta) + cos^2(\varphi - \theta)(1 + a/r)^2\}} = m_0(1 + a/r \cdot cos^2(\varphi - \theta))$

with dL'= r/cos(φ–θ)dφ and L' ≅ r_o/θ we have

$$m_{avg} = m_0/L' \int_0^{L'} dL' + m_0 a/L' \int_0^{\pi/2} (\cos^2 (\varphi-\theta))d\varphi/\cos(\varphi-\theta)$$

$= m_0(1 + a\theta/r_o)$ (to first order, the integral is ≅ 1).

Minimizing L'/m_{avg} (the number of meters on the path) quickly gives

θ = a/r_o = MG/$r_0 c^2$ as before

The ease of calculating depends on your view. So why was one calculation much easier than the other? In the first calculation the effect of length contraction only affected light in the radial direction and that is at very large radii, giving a small affect integrated over large distances, a nasty integral. In the second calculation, the effect of tangential "length expansion" was only in the tangential direction which affects light only at small radii. A larger affect integrated over a smaller distance, giving a reasonable integral. You can ignore some of the comments in "a few observations

The total deflection due to both effects is 4 MG/$r_0 c^2$, 1.75 arcseconds

This whole light bending thing could've been done in one fell swoop by just calculating the path of minimum number of wavelengths and taking into account both the effects of length contraction and time dilation on the wavelength of the light. The integral would've been

$$\lambda_{avg} = \lambda_0 / L' \int_0^{\pi/2} (1 - a/r - a/r \cdot \sin^2(\varphi-\theta)) dL'$$

where a is MG/c^2 and dL' = r/cos($\varphi - \theta$)dφ.

$$\lambda_{avg} = \lambda_0 (1 - a/L' \int_0^{\pi/2} (1 + \sin^2(\varphi-\theta))/\cos(\varphi-\theta) \cdot d\varphi)$$

$$= \lambda_0 (1 - a/L' \int_0^{\pi/2} (2 - \cos^2(\varphi-\theta))/\cos(\varphi-\theta)\, d\varphi)$$

The first term in the integral gives a factor of two times what we got for time dilation and the second factor is equal to 1 + θ, but gives only a second order contribution when we calculate the path of minimum number of wavelengths.

This is analogous to classical optics. In our flat world, the path of minimum number of wavelengths in free space turns out to be a straight line i.e. the shortest distance. But if there is something else going on, like a variable index of refraction (or time dilation), then the path is not a straight line. Time dilation and length contraction, each cause an increase in the wavelength of light (but in somewhat different ways) with the distance from the sun, causing the path to bend. In classic optics, there is something called Fermat's Principal that says light goes on the path of least time. In the warped system, with time dilation, the path of minimum number of wavelengths prefers paths away from the sun where the wavelengths are longer (fewer wavelengths), whereas in Fermat's case the light prefers paths near the sun where the seconds are longer (fewer seconds). The answer would have the same magnitude but the opposite sign, so light would bend in the opposite direction for the path of minimum time from what we get for the path of minimum number of wavelengths. In Einstein's first calculation of the bending of light, neither time nor light velocity varied in space and therefore the path of minimum time would've just been a straight path for the beam of light, no bending. Why is this, because Fermat's principal works in classical optics? In classical optics, with a variable Index of refraction, the velocity of light changes but the frequency stays the same. If the frequency stays the same, then the wavelength varies and regions of higher velocity are regions of longer wavelength. Paths of least time (high velocity) are equal to paths of least number of wavelengths (longer wavelengths). In our warped system, with time dilation, the speed of light is constant and so in areas of slower time (higher frequency) the wavelengths are shorter and therefore paths of least time are not paths of the least number of wavelengths. Fermat's principal works in classical optics because it is the same as the path of minimum number of wavelengths, but without the important "slop" in the path. In the next example, where we will calculate the light bending in an unwarped frame, we will see that Fermat's principal works.

The unwarped system

We now want to calculate the bending of light as seen by a distant observer and calculated in his flat, unwarped, reference frame. The speed of light is a constant, c, locally in the gravity frame, but because of length contraction and time dilation, the speed will not be the same as calculated in the flat reference frame. We must translate the constant speed of light in the gravity frame to the speed of light in the flat reference frame. It will depend on both the position and direction of the light.

We need to calculate dr/dt and dC/dt (circumference) in the flat frame so let's think about differentials. Because the radius is shorter in the gravity frame, for a given length dr in the flat frame there is a larger number of meters in the gravity frame, ie, dr_g = dr(1+a/r). Because time is slower in the gravity frame, for a given time increment in the flat frame, dt, there are fewer seconds in the gravity frame so, dt_g=dt(1-a/r). The circumference is the same in both frames.

That said, we get dr/dt = c(1-a/r)/(1+a/r) = c(1-2a/r) and dC/dt = c/(1+a/r) = c(1-a/r). Time is the same everywhere in the flat frame, so this is analogous to the classical case where the frequency of the light is constant and the wavelength varies directly with the speed. In fact we could write the following, λ_r = λ_0(1-2a/r) and λ_c = λ_0(1-a/r) where λ_0 is the wavelength with no gravity present. If one wants a complete analogy with the classical case, then one has an index of refraction which is somewhat strange in that it depends both on position and direction of the light.

As before, we want to calculate the average wavelength along our path, so we calculate the number of wavelengths along the path and then minimize it in order to get the angle of bending. The wavelength in the integral will be

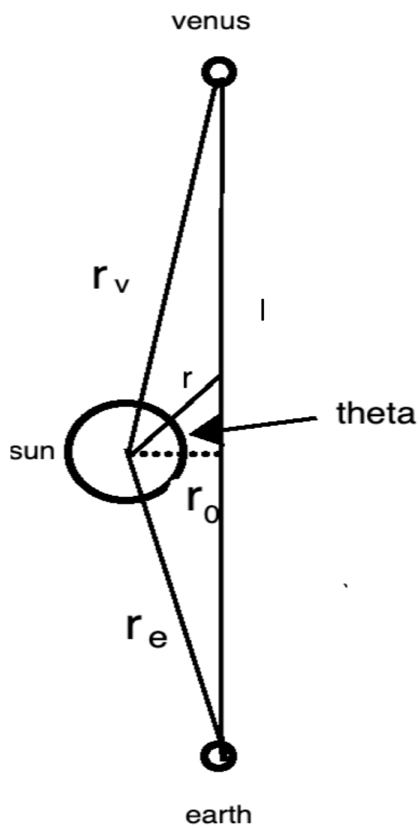
$$\lambda = \lambda_0 \cdot \sqrt{\{sin^2(\varphi-\theta)(1-2a/r)^2 + cos^2(\varphi-\theta)(1-a/r)^2\}}$$

This reduces to; 1 - a/r - a/r$\cdot sin^2$(φ–θ) (remember $(1+x)^n$ = 1 + nx), which is exactly the same as we had above when we combined time dilation and length contraction in the warped (gravity) frame, and so the light bending will be as before; 4 MG/r_0c^2. For the distant observer, it makes no difference whether he calculates the bending of light as it goes through the warped gravity frame where the velocity of light is constant, or whether he calculates the bending in his flat frame where the velocity of light varies

The Shapiro time delay

In 1964, Irwin Shapiro proposed a test of general relativity which consisted of reflecting a radar beam off of either Venus or Mars, and back to earth, while the planet was moving behind the sun so that one could observe the change in transit time as the radar beams came closer and closer to the surface of the sun and, therefore, into a stronger and stronger gravitational field. Shapiro calculated and measured a time delay of about 200 μs as the beams moved from far away to near the surface of the sun.

From what little I have read, there seem to be a fair number of people who think that the time delay is caused by the fact that the radar beam is bent around the sun, and the longer path causes a time delay. Let's put some numbers into this idea. Due to bending, the path will be multiplied by 1/cosine(θ) = 1 + $\theta^2/2$ where θ is half the angle of the bending calculated in the previous section, which is 4×10^{-6} radians. This gives a fractional path increase of 16 x 10^{-12}. If we consider radar bouncing off of Venus, the path in one direction is just the radius of the orbit of Venus plus the radius of the orbit of earth, which is equal to about 2.5×10^{11}m. Multiplying by two and dividing by the speed of light gives a transit time out and back of about 2000 sec. If we multiply by the fractional increase in the path length, we get a time delay of about 30 nanoseconds, nowhere near Shapiro's 200 microseconds. It's fair to say, this idea for the time delay does not work.



Let's calculate the time delay similar to our last example of bending in the flat frame. The times and distances of the measurement of course are in the flat frame and so the speed of light will not be a constant in this frame. Consider the figure above. We have a beam of radar which goes from earth to Venus and back again passing at a distance r_0 from the center of the sun. We will integrate along this path to get the time. We integrate first along a straight line from a level of the center of the sun to the planet Venus. In our drawing that would be an integration from theta equals zero up to an angle which is almost 90°. We want to integrate dt = dl/velocity. Now the velocity of light along the path at the position theta is;

V = c ($\sqrt{}$(1-2a/r)$^2 sin^2\theta$+(1-a/r)$^2 cos^2\theta$) = c(1 -a/r - a/r$sin^2\theta$) to first order

We want to calculate the integral T = $\int$ dl/v = 1/c$\int$ dl(1 +a/r + a/r$sin^2\theta$)

dl = dr/sinθ , r = r_0/cosθ, and dr =$r_0$sinθ/$cos^2\theta$dθ

The integral becomes

T =r_0/c$\int$ (1/$cos^2\theta$ + a/$r_0$cosθ + a·$sin^2\theta/r_0$cosθ)dθ

A trip to the integral tables gives;
T =r_0/c· tanθ + 2a/c·ln(tanθ + secθ) - a/c·sinθ
evaluated from 0 to arctan $\sqrt{(r_v^2 - r_0^2)}/ r_0$
Now $r_0^2 \ll r_v^2, r_e^2$, so the distance along the straight path is essentially $r_v + r_e$
Most of the contribution comes from the upper limit where tan = sin/cos ≅ 1/cos = r_v/r_0 so:
T = r_v/c + 2a/c·ln(2 r_v/r_0) -a/c   for the Venus leg,

T = r_e/c + 2a/c·ln(2r_e/r_0) -a/c   for the earth leg;

Adding the two together and multiplying by 2 for the return trip

$$\text{Transit Time} = 2(r_v + r_e)/c + 4a/c\cdot(\ln(4r_e\cdot r_v/r_0^2) - 1)$$

The first term is just the Newtonian value of distance/c and the second term is the Shapiro time delay, which is about 200 microseconds for r_0 equal to the radius of the sun . Since one does not know the transit time well enough, only the ln ($r_e \cdot r_v/r_0^2$) term is important because it gives the variation of the transit time with the distance of closest approach, r_0, which is what the measurement does.

Precession of the Perihelion of Mercury

First, a little history. In the mid-1800's it was calculated that the precession of the orbit of mercury should be 532 arc seconds/century due to the gravitational force from the other planets, but unfortunately the measured value was 575 arc seconds/century. This leaves about 43 arc seconds to account for. Can relativity be the answer? Let's see. Now the angle of precession per revolution of the planet in its orbit is a dimensionless number. As we have already mentioned, the answer will be "something" times MG/rc^2 where r is now the radius of the orbit. We need to find the "something". As usual the radius of the orbit will be length contracted and will be equal to r_g =r (1-MG/rc^2) .

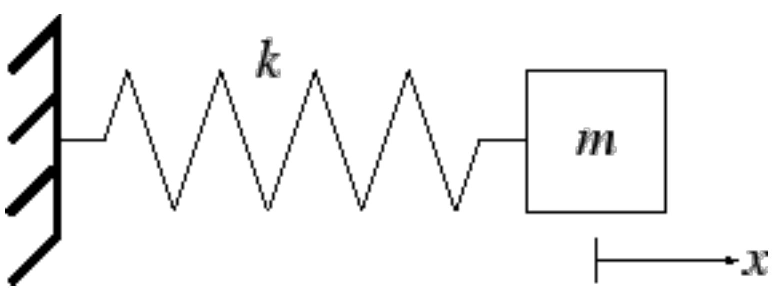
Let's consider the parameters involved in the orbital motion. When relevant, the parameters will be per unit mass of the planet. We want to consider their deviation from their non-relativistic values, as seen by our distant observer. The parameters will be something like A = A_0·(1 + nMG/rc^2) = A_0·(1 +n2r_s/r) where n is a positive or negative integer. Remember that multiplying is the same as adding. We will list the n values for each parameter. For instance, r will be -1 because of length contraction in the radial direction. Also, 1/r is +1 and r^2 is -2. The orbital velocity V is ≪ c.

Parameter	Designation	n	Comment
Radius	r	-1	length contraction
Time	t	+1	time dilation
Mass of the sun	M	+1	freefall analysis
Velocity in orbit	V	+1	time dilation, no tangential length contraction
Angular momentum	rV, h	0	a constant of motion that stays constant
Centripetal accel	V^2/r, h^2/r^3	+3	
Gravitational force	MG/r^2	+3	
Orbit circumference	C	0	no tangential length contraction

If one doesn't like the increase in M, then one can argue that the centripetal acceleration increases by three units and, therefore, there must be a gravitational force caused by "something" which causes this acceleration. The increase in M, though, is part of the freefall analysis, so, like it or not, we keep it. The angular momentum contains the following, length contraction in the radial direction, no length contraction in the tangential direction, and time dilation in all directions.

Orbital Mechanics and Harmonic Oscillators

Let's start simple and work our way up. Consider the mass on a spring shown in the following diagram. We will do things per unit mass, that is, m = 1.



For small movements Δx from the equilibrium position, the restoring force F is –k·Δx where k is the spring constant. The equation of motion is:

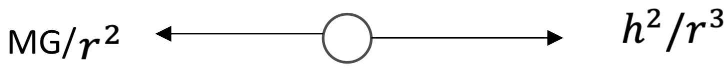
d^2(Δx) $/dt^2$ + kΔx = 0 and the solution is Δx = a·sin(ωt)

where ω = $\sqrt{k} = \sqrt{dF/dx}$ and a is an amplitude

If k is a function of x, the analysis is still OK for small Δx. Our springs are nonlinear and so we will consider only small deviations from the equilibrium(null) position with the derivative dF/dx being evaluated at the equilibrium position (zero net force).

Classical Calculation

Consider a planet orbiting the sun. We go into the orbiting frame of the planet (you don't have to, but it is easier to visualize) and add a centrifugal force, h^2/r^3 outward from the sun, to replace the centripetal acceleration. We now have two forces on the planet, a centrifugal force and a gravitational force as shown below.



A few things to notice. As r increases, we notice that the centrifugal force decreases faster than the gravitational force and therefore there is a restoring force to decrease r. If the gravitational force were a $1/r^3$ force, then there would be no restoring force and the universe would consist of things floating around randomly. $1/r^4$ is even worse. Luckily for us, the force is $1/r^2$.

The total force is $F_t = h^2/r^3 - MG/r^2$

For oscillations in the radial direction:

$-k = dF_t/dr = -3h^2/r^4 + 2MG/r^3 = -h^2/r^4$ (at the equilibrium(null) position)

Therefore $\omega_r = \sqrt{k} = h/r^2$, but, because $h = \omega_o r^2$, this is the same as the frequency of the planet orbiting the sun, ω_o (r for radial, o for orbital). This means that the point of maximum radius (perihelion) occurs at the same point every revolution of the planet around the sun. This slightly elliptical orbit is, therefore, stationary and has no precession.

Classically: radial frequency = orbital frequency

"General Relativity" Calculation

We start with the classical expression for the spring constant along the radial direction.

$k = -dF_t/dr = 3h^2/r^4 - 2MG/r^3$ for the no gravity(inertial) frame.

We then write this in terms of the gravity frame parameters using $r = r_g(1+a/r)$.

$a = MG/c^2$ and r_g is the orbit radius in the gravity frame as measured in the inertial frame. Remembering that $(1 + x)^n = 1 + nx$ for small x, we get , first by transforming r (the subscript g denotes the gravity frame)

$$k_g = 3h^2/(\mathrm{r}_\mathrm{g}^4(1+4a/r)) - 2MG/(\mathrm{r}_\mathrm{g}^3(1+3a/r))$$

$$k_g = \omega_{rg}^2 = 3h^2/\mathrm{r}_\mathrm{g}^4 - 2MG/\mathrm{r}_\mathrm{g}^3 + (-12h^2/\mathrm{r}_\mathrm{g}^4 + 6\,MG/\mathrm{r}_\mathrm{g}^3)\cdot a/r$$

We must now evaluate this at the null point in the gravity frame.

In the no gravity frame the null point is where h^2/r^3 = MG/r^2. We naïvely showed earlier that the changes in the gravitational force and the centripetal acceleration were equal when we moved to the gravity frame, as long as we included the increase in the mass of the sun, M. (I wish I had a better explanation). If we do this, then the null point is where, $h^2/\mathrm{r_g^3}$ = M_gG/$\mathrm{r_g^2}$, so

In the term with MG/c^2, it makes no difference, to first order, which mass and which radius you use, gravity or no gravity, since it is such a small term to begin with.
We can write the change in k_g due to a shift in r of MG/c^2, with dr = dr_g, as;

k_g = k + dk/dr_g ·Δr = 3$h^2/\mathrm{r_g^4}$- 2MG/$\mathrm{r_g^3}$+ d/dr_g (3$h^2/\mathrm{r_g^4}$- 2MG/$\mathrm{r_g^3}$)· MG/c^2

k_g = 3$h^2/\mathrm{r_g^4}$- 2MG/$\mathrm{r_g^3}$ + (- 12$h^2/\mathrm{r_g^4}$ + 6 MG/$\mathrm{r_g^3}$)· MG/rc^2 as before

ω_{rg}^2= $h^2/\mathrm{r_g^4}$ (1– 6a/r) = ω_{og}^2(1 – 6MG/c^2r)
Taking the square root,
ω_{rg} = ω_{og}(1 – 3MG/c^2r)

Implied in all of this is that time dilation affects the two frequencies the same and therefore does not affect the ratio. How did all this happen? Well, as seen from the inertial frame, the shift in r decreased k, and therefore the radial frequency. There was no corresponding change in the circumference and therefore the orbital frequency stayed the same. Gravity has caused a difference in the two frequencies, which we can probably blame on radial length contraction.

This answer is less than the frequency of orbital rotations, ω_{og}, and, therefore the orbit has a precession, and is not stationary. The point of maximum radius (perihelion) occurs later each revolution of the planet around the sun (the orbit <u>advances</u> each revolution by the time the maximum radius occurs). The difference per revolution, in radians per second is:
ω_{rg} - ω_{og}= -ω_{og}· 3MG/c^2r. If we multiply by the time of one orbit, 2π/ω_{og}, we get the number of radians the radial oscillation lags the orbital oscillation per revolution.

$$\Delta\theta/\text{revolution} = 6\pi\ \mathrm{MG}/rc^2$$

This is the answer. 6π is the “something” we were looking for.

For light bending, 4 MG/r_0c^2 was equal to 1.75 arcseconds , so we replace 4 by 6π and the radius of the sun r_0 (7x10^5 km) with the radius of mercury's orbit. It is a little sloppy, but we will just use the mean radius of 57x10^6 km. The answer is 0.102 arc seconds/ revolution. The period of mercury's orbit is 88 of our days, so we multiply by 365/88·100 to get 42 arc seconds/century compared to the 43 that we needed. Treating the amplitude of oscillation of Mercury's orbit radius as being small is pushing it a little bit, so a more careful analysis would give a better answer. This whole analysis is for any planet going around a large body like the sun. If you do the calculation for Earth you get about 3.8 arc seconds per century .

Comparison with the Swartzchild Metric

It is interesting that we were able to calculate the bending of light by gravity, the Shapiro Time Delay and the shift of the Precession of the Perihelion of Mercury
without reference to something called the Swartzchild metric, which is the exact solution to Einstein's field equations for a spherical symmetric radial gravitational field. There must be some relationship between our approach and the Swartzchild metric. Here is what I found.
In classical mechanics the distance between two points, ds is such that ds^2 = dx^2+dy^2+dz^2 and is invariant to the choice of coordinate system. In special relativity, space and time got mixed together and the invariant now is ds^2 = dx^2+dy^2+dz^2- $(cdt)^2$ where c is the speed of light and dt is the time difference between the points, and again, ds is invariant to the choice of coordinate system. In spherical coordinates this is ds^2= dr^2+ $r^2(d\theta^2$+ $sin\theta^2 d\varphi^2)$ - $(cdt)^2$. In this case the time difference between two points depends on the coordinate system and their spacial positions, but we already found that to be true This metric is called, I believe, the Minkowski metric for flat space-time.
Now gravity further mixes up space and time and the metric now is the Swartzchild metric. The invariant is, for a spherically symmetric gravitational field:
ds^2= dr^2/(1-r_s/r) + $r^2(d\theta^2$+ $sin\theta^2 d\varphi^2)$ - $(cdt)^2$·(1-r_s/r) where r_s is the Swartzchild radius, 2MG/c^2. It is fairly obvious that the 1-r_s/r in the first term is related to length contraction and the 1-r_s/r in the last term is related to time dilation.

Results from our freefall analysis were :

1. The radial dimension was contracted, making the distance between two points measured on that axis to increase such that dr was altered to dr(1+r_s/2r).
2. The dimensions in the tangential direction were not changed.
3. Because of time dilation, the time difference between two points decreases such that dt is altered to dt(1-r_s/2r)

If we plug these results into the Minkowski metric to account for gravity, we get

ds^2= dr^2(1+r_s/2r)^2+ r^2($d\theta^2$+ $sin\theta^2 d\varphi^2$) - $(cdt)^2(1-r_s$/2r)^2 The second and third terms are not altered because they are tangential. Using $(1+x)^n$= 1 + nx for x$\ll 1$ and expanding the squares, this can be written as:

ds^2= dr^2/(1-r_s/r) + r^2($d\theta^2$+$sin\theta^2 d\varphi^2$) - $(cdt)^2$·(1-r_s/r) which is the Swartzchild metric.

So, our analysis agrees with the exact solution for r_s/r$\ll 1$, which is the case for the two examples we have calculated and which is the assumption we made at the beginning.

Black Holes and the Photon Sphere

So far, our simple approach to general relativity has done well as far as calculating goes. We have taken examples only where the radius from the large gravitational object is much larger than the Swartzschild radius. We should probably quit while we're ahead but let's forge into black holes.
We revisit our free fall example, but this time falling towards a very small but very massive object. We want to calculate at what radius our free fall velocity, with respect to this object, is equal to the speed of light. The first calculation will be done assuming that the massive object accelerating towards us doesn't change its mass as it's velocity towards us increases.

dV = acc·dt = acc·dt/dr·dr = MG/r^2·1/V·dr So VdV = MG/r^2·dr acc is acceleration

integrating with V equals zero at infinity, we get, for the radius where V equals c,

r = 2MG/c^2, which is the Swartzschild radius r_s.
This radius is twice as big as the radius you would get for a black hole using our approximation of low velocity and so don't be disturbed by the size of this radius, besides, with our approximation, we are not supposed to be working down at these small radi.
For our sun, this radius is about 3 kilometers compared to its actual radius of 0.7 million km. We will never get to the Schwartzschild radius by free-falling toward the sun. For it's mass, it is not small enough. An object smaller than this radius is called a Swartzschild black hole.
In the gravity frame things are pretty grim. The clocks have slowed to a halt and the radial dimensions have now decreased to zero due to length contraction. This radius is called the event horizon. Anything that occurs inside of this event horizon will not make it out. Massive particles would have to be traveling at the velocity of light just to make it out. Photons would have to have infinite frequency and, even then, they would be red shifted down to zero frequency if they escaped.

The Photon Sphere

In classical optics, there is an interesting example I will call the light merry-go-round. Let's say we have a transparent disk which has an index of refraction which decreases with increasing radius. For instance, let n = n_0(1 –r/r_0) where n is the index of refraction. Because of this index, the velocity and wavelength of light will increase with increasing radius. Here comes the path of

minimum number of wavelengths again. If the index were constant, then for a circular path of light, an increase in the circumference of the path would always cause an increase in the number of wavelengths and so there would be no path of minimum number of wavelengths. But in our case the index decreases as the radius increases which means that the wavelength increases. There may be a radius at which the percent increase or decrease in circumference of the path and the increase or decrease in the wavelength, as the radius increases or decreases, are equal and, therefore, the number of wavelengths does not change. This would be the path of minimum number of wavelengths. The light would orbit around in a circle at this radius. Let's see if there is such a radius. The number of wavelengths is just the circumference divided by the wavelength.
$\lambda = \lambda_0/n = \lambda_0/n_0(1 - r/r_0)$

wavelengths = $2\pi r/\lambda = 2\pi r\cdot(1 - r/r_0)\, n_0/\lambda_0$.

Taking the derivative and setting it equal to zero gives $r = r_0/2$. One could make n_0 equal to 3 and therefore the orbit would be where the index is 1.5. Such a thing would be interesting on the scale of several microns such as a section of a graded-index multi-mode optical fiber. This effect would cause a spiral mode for light going down the length of a graded-index multi-mode fiber. The index does not need to vary linearly with r. For any index n(r), all that is needed is that there be some radius where dn/n = -dr/r.
In the gravitational field around a black hole, we have a similar situation. Time dilation causes the wavelength of light to increase the further away you get from the black hole, giving
$\lambda=(1-a/r)\lambda_0$ where λ_0 is the wavelength at infinity. The wavelength, in a circular orbit, is not affected by length contraction. The situation is a little more complicated than the previous example, because the circumference is not equal to 2π times the contracted radius. The circumference is equal to two pi times the uncontracted radius, r/(1-a/r), because the circumference is not affected by length contraction, as is the radius. We then have that the number of wavelengths on a path of radius r (contracted) is:

wavelengths = $2\pi r/(1-a/r)\lambda = 2\pi r/(1-a/r)^2\lambda_0$ where $a = MG/c^2$

Here we must be careful and not use $1/(1-a/r) \cong 1+a/r$ since a/r is not $\ll 1$.
The constants drop out when we set the derivative equal to zero.
$d\#/dr = d/dr(r/(1-a/r)^2 = 0$: This gives $r^2 - 4ar + 3a^2 = 0$. Solving for r:

$r = 3a = 3/2\cdot r_s$. This is the radius of what is called the photon sphere,

a radius where light "can" orbit around the black hole. The continuous orbiting is probably not stable and the light would probably leave this orbit. But, this radius is outside the bounds of where our approximations are reasonable. In "reality", we should not be calculating anything down at this radius with our simple approach. None the less, $r = 3/2\cdot r_s$ is the accepted value for the radius of the photon sphere. Go figure.

The twin paradox

A famous paradox of special relativity is the twin paradox. Here we have two twins, one which stays home and the other who travels in a spaceship and then turns around and returns to earth.

Both twins see the other twin's clock as running slow and so which one is younger when the second twin returns? The rule is to analyze the situation from the point of view of the inertial frame. This is the frame of the twin that stays home, so the answer is that the twin who took the ride in the spaceship comes back younger. How does he see this? The explanation is in the turnaround. Here the twin in the spaceship experiences an acceleration which is equal to a gravitational field. This field will be in a direction from earth to the spaceship. During turnaround then, the clock on earth will be higher in the gravitational field and therefore the twin in the spaceship will see earth's clock running faster than his. This speed up in earth's clock is enough to more than compensate for the fact that the earth clock was running slow before and after turnaround and when the spaceship twin gets back to earth, he will find that he is younger than his earth twin. Let's calculate. Say that the travelling twin starts his turnaround at a distance L from earth. If he has been traveling with a velocity V, then the turnaround acceleration, in a time T2, which we take to be much less than the travel time, will be g = 2V/T2. If you do the freefall exercise in this constant gravitational field, you will find the difference in the rate between two clocks that are spaced a distance X apart, assuming V$\ll c$: % difference in clock rates = γ-1 = gX/c^2. You can throw in length contraction but it's a second-order effect. During the turnaround time T2, the earth's clock, will gain gX/c^2·T2 = 2LV/c^2 seconds on the traveler's clock. But, during the out trip and the back trip, the earth clock lagged behind the traveler's clock by (γ-1)·travel time = $1/2V^2/c^2$·2L/V= LV/c^2, again assuming V$\ll c$. The net effect is that the earth clock, according to the traveler, advances LV/c^2ahead of his clock during the entire trip and so the traveler comes back younger. That is what we got when we look at the problem from the point of view of the earth twin. Paradox resolved.

A more realistic twin paradox would be one where the travelling twin accelerates from the earth up to a velocity V, turns around, and then decelerates down to zero velocity to land on earth. During the acceleration and deceleration near earth, the traveling twin sees the earth clock as running slow compared to his clock due to gravity. During turn around he sees the earth clock as running fast. These effects do not cancel. The effect of the acceleration on the clocks depends on the distance between the clocks and therefore the turnaround outweighs the effect of acceleration and deceleration near earth. The net effect is that the traveler comes back younger. For those who like to calculate, one might try the following exercise. Consider the traveler leaving earth with an acceleration a to a velocity V, and immediately decelerating and turning around until his velocity is V toward the earth. He then lands on the earth at the zero velocity. One can calculate to see whether the traveler and earth twins both agree on who is younger and by how much. A simpler exercise is to show that if a constant velocity section is added between the acceleration and deceleration sections, that they still agree, although, on a different number.

Do our orbiting astronauts come back younger? No, they come back older. More on that in the next section.

Another thing to think about with the twin paradox, is to consider the following. If the stationary twin looks at the moving twin's x axis as it moves by, which is marked off in meters, he notices that it is length contracted with respect to his. From his vantage point, out to the position of his twin, there are more meter marks on the moving twin's axis than on his axis. During turn around, when the travelling twin slows down to zero velocity, his x-axis becomes the same scale as that of the stationary twin. Where did all those meter marks go that had passed by? Try to figure out how this all evolves. The keyword here is "when". "When" is a tricky word in relativity.

Clocks and GPS

One occurrence of relativity in our everyday life is the Global Positioning System (GPS), although we probably don't know it (or care about it). GPS consists of 31 satellites (as of 2017) orbiting the earth at a distance of 20,000 km above the earth and with an angular velocity of twice that of the earth, so that the satellites pass over every 24 hours. The idea is to have every point on earth have a line of sight to at least 4 satellites at any given time. The satellites give out coded signals such that an Earth-based receiver, like your cell phone, can determine the time-of-flight of a radio signal from the satellite to your receiver. This time-of-flight gives the distance from the satellite to the receiver. There is a sphere around the satellite with this particular distance. If the receiver knows where the satellites are at any given time, then three spheres around three separate satellites will intersect at your position on earth (they also intersect at another point which is not near the surface of the earth). Radio waves (and light) travel at a speed of about 1 foot/ nano second. So, if you want to determine your position to a few feet, this timing needs to be done to a few nanoseconds.

There are two problems with the timing. The first one is that relativity causes clocks that are moving with respect to one another and clocks separated in a gravitational field, to run at different rates. To do the position calculations, the earth-based receiver must be synchronized to the satellites' atomic clocks to a few nanoseconds. The second is that your cell phone does not have an atomic clock in it and is not nearly a good enough clock to do the job.

Let's consider the first. We want to calculate the difference in rates between the satellite clock and our earth clock. If we calculate, say, in microseconds per day, this is a dimensionless number, time divided by time. As before, the answer will be "something" times MG/rc^2, where M now is the mass of the earth and r is the radius of the satellite orbit. Oddly enough, if you guessed 4, like the bending of light by gravity, you would be pretty close. OK, just a coincidence.

We will calculate things in terms of the radius of the satellite orbit, r_{sat}, the angular velocity of the satellite, ω_{sat}, and the ratio of the satellite's orbit radius to the earth's radius, which we will call x. Remember that the angular velocity of the earth is half that of the satellite. We will calculate the difference in the rate of the two clocks as seen by a distant observer. First, the effect of velocity as seen on the distant observer's clock.

The % difference in rates due to the velocity with respect to the distant observer's clock is just γ -1, which, with V$\ll c$, is equal to 1/2$\cdot V^2/c^2$. For the satellite, V = $\omega_{sat}\cdot r_{sat}$ and for the earth clock, $\omega_{sat}\cdot r_{sat}$/2x. The difference in rate between the two clocks, therefore, is $\omega_{sat}^2\cdot r_{sat}^2/2c^2(1 - 0.25/x^2)$. The centripetal acceleration of the satellite, $\omega_{sat}^2 r_{sat}$ is equal to the gravitational force per unit mass on the satellite, MG/r_{sat}^2. Therefore, $\omega_{sat}^2\cdot r_{sat}^2$= MG/$r_{sat}$, and the rate difference of the clocks due to their velocity with respect to the distant observer is MG/$2c^2r_{sat}$ (1 - 0.25/x^2), with the satellite clock running slow.

We are not done because we still need to consider the effect of gravity. We showed before that in a central gravitational field, time dilation causes time to be a function of the radius, T = T_0(1-MG/ rc^2). The difference in the rate of time at the earth and the satellite due to gravity is, therefore equal to MG/c^2 ($1/r_e$-$1/r_{sat}$) = MG/$r_{sat}c^2$(r_{sat}/r_e-1) = MG/c^2r_{sat}(x -1) with the satellite clock running fast.

Putting this all together:

The net difference in clock rates is: MG/c^2r_{sat}(x-1.5+0.12/x^2)

Now x = 26.4/6.4 = 4.1, so the difference in clock rates is 2.6MG/c^2r_{sat}.

And again, the answer is "something" times the dimensionless parameter MG/c^2r.

Putting in numbers, gravity causes 45 µs per day and relative velocity, 6 µs per day, the net difference in rate being 39 µs per day with the satellite clock running fast. If an astronaut were on the satellite, he would be aging faster than people on earth.

So, that is the standard approach, with gravity causing an effect which is about 7 ½ times that of velocity, that is, "general" relativity contributes 7 ½ times that of "special" relativity. But are they really separate effects? It depends on your point of view. Our approach to general relativity is that gravity is equal to acceleration, acceleration leads to velocity, velocity leads too length contraction and time dilation. They seem to be all mixed together.

To show that, let's go into a frame rotating with the satellite with the earth at the center. Here, the satellite is at a point with no acceleration and no gravity because gravity at the satellite has been canceled by the centrifugal force. An inertial point. We must then subtract the effect of the centrifugal force, $\omega_{sat}^2 r_{sat}$, integrated all the way down to zero radius, from our previous calculation of the effect of gravity. We need to integrate all the way down to zero because the centrifugal force also affects the earth clock. If you like the concept of dropping a photon, then it is the dropping of a photon in a force field which contains the force of gravity and a centrifugal force in the opposite direction.

The effect of centrifugal force = $\int_0^{r_{sat}} F_{cen} dr$ = $\omega_{sat}^2 \int_0^{r_{sat}} r dr$ = $\omega_{sat}^2 \cdot r_{sat}^2/2$.

As before, $\omega_{sat}^2 \cdot r_{sat}^2$= MG/$r_{sat}$, so when we subtract this from our previous gravity calculation we get the result: net "gravity" effect = MG/c^2r_{sat}·(x-1.5)

There is a small velocity of the earth clock in this frame of $\omega_{sat} \cdot r_{sat}$/2x. In this case the effect now causes the earth clock to run slow with respect to the satellite clock. The effect on rates is MG/c^2r_{sat}·0.12/x^2 and must be added because the effect is in the same direction as gravity.

The net difference in clock rates in the rotating frame is: MG/c^2r_{sat}(x-1.5+0.12/x^2)

It is the same as before, but now, with x= 4.1, "general" relativity contributes 99.7% of the entire effect. It depends on how you look at it. One man's acceleration is another man's gravity.

This difference in rates is not a big deal because the clocks in the satellites can be corrected with electronics. If they weren't corrected, the GPS system would lead you off course by about 11 km a day, but the correction is to include another satellite.

As mentioned earlier, the other problem is that your cell phone doesn't have a clock good to nanoseconds and therefore the distances from all the satellites is probably wrong, but by the same amount, the same effect as relativity except probably larger and dependent on the user. This, of course, gives the wrong position. If we include information from one more satellite, then we can dither the time on the cell phone until the distance from all 4 satellites coincide at some point. This is the correct position. One could also do this with uncorrected satellite clocks and just solve 4 equations and forget about the satellite clocks being off with respect to the cell phone. So, the whole relativity argument is essentially irrelevant, except that there may be small differences in the satellite clocks due to relativity that need to be corrected. Oh well, it was an excuse to think about relativity but it is a good test of relativity.

It looks like special relativity plus freefall is a good approximation to general relativity. It doesn't give us a gravity field plus gravity waves but it gives us a few things to hang our hat on. In writing this thing up, I've learned quite a bit about relativity and some wrong ideas I had.

Well, that's about all for relativity. It's strength, if you want to call it that, is that no fancy formalism is required and so any undergraduate engineer or science student, if they wanted, could understand it fairly easily. To me, relativity, along with optics, are examples of where simple assumptions can get you a long way just by thinking.

This writeup evolves in time because of more thinking or poor thinking. The next time you look it may have changed a little bit but the answers will stay essentially the same. Please excuse the formatting and formulas but I did this thing in Word.

If you have comments or can suggest changes, please send them to relativity@virgilelings.com.

Thanks for reading.