A small simulation of a large phenomenon: Gravitational Lensing

Tanya Roussy, Jake Retallick, Angus Lim, Evan Goodacre, Adrien Emery
Department of Rocket Science, University of British Columbia
(Dated: April 8, 2012)

This article offers an introduction to General Relativity and one of the theory's many interesting consequences, gravitational lensing. This is aimed at any curious undergraduate student with a sufficient background in Newtonian Physics and Calculus. The derivations do involve tensor analysis, which we explain only briefly here. We recommend this [4] excellent, concise introduction to tensor analysis for students who are unfamiliar with the topic. The authors of this article have developed a gravitational lensing simulation, which you can find here[add link], and is available as an iPhone app : [name of app]. The reader should note that einstein summation notation is employed throughout this paper.

INTRODUCTION

The bending of light in a gravitational field predicted by General Relativity provided one of the first experimental verifications of Einstein's theory, in 1919 [7]. It was not until the 1960's that further confirmations were made. After Einstein's theory was published, it was noted that the light deflection by a stellar object would produce secondary images of the background sources, which could resemble binaries, if the lens is not perfectly aligned with the observer, or rings in the case of a perfect alignment [8] .

It was later shown that more massive deflecting objects, such as a galaxy, would lead to resolvable images [9]. This 'macrolensing' effect can provide information about the mass of the intervening galaxy, and could make very distant objects more observable due to the magnification of light from the lens . The first observation of a quasar lensed by a foreground galaxy was in 1979 [10], and a new era of research began. The lensing effect has a long and growing list of applications for cosmological and astronomical research.

The macrolensing effect can be used to measure the Hubble constant, via the time delay among multiple images, which may also lead to a determination of the expansion rate of the universe. Statistical frequency of of multiple imaging gives constraints on the cosmological constant, and the distribution of lensing events offers observations on the mass structure/distribution of the universe. Since lensing is the direct result of any massive object, lensing measurements offer the possibility of detecting and measuring baryonic and non-baryonic dark matter [11].

This brief overview of gravitational lensing's applications is in no way exhaustive. We see that gravitational lensing offers a direct way of measuring the amount and spatial distribution of all types of (conjectured) matter in the Universe, and allows us to obtain indispensable information on cosmological parameters.

FIG. 1. Dark Matter Map in Galaxy Cluster Abell 1689. Hubble Space Telescope image
[13]

FIG. 2. A background galaxy being lensed by dark matter in foreground cluster. Hubble Space Telescope image
[13]

SPECIAL RELATIVITY

We begin this introduction with a brief look at Special Relativity. Relativity theory (in general) examines relations between 'reference frames': observational perspectives. There is no better motivation for this theory than the one Einstein gave at the beginning of his paper on the subject. We reproduce it here:

"It is known that the application of Maxwell's electrodynamics, as ordinarily conceived at the present time, to moving bodies, leads to asymmetries which don't seem to be connected with the phenomena. Let us, for example, think of the mutual action between a magnet and a conductor. The observed phenomenon in this case depends only on the relative motion of the conductor and the magnet, while according to the usual conception, a strict distinction must be made between the cases where the one or the other of the bodies is in motion. If, for example, the magnet moves and the conductor is at rest, then an electric field of certain energy-value is produced in the neighbourhood of the magnet, which excites a current in those parts of the field where a conductor exists. But if the magnet be at rest and the conductor be set in motion, no electric field is produced in the neighbourhood of the magnet, but an electromotive force is produced in the conductor which corresponds to no energy per se; however, this causes (equality of the relative motion in both considered cases is assumed) an electric current of the same magnitude and the same course, as the electric force in the first case. Examples of a similar kind, as well as the unsuccessful attempts to substantiate the motion of the earth relative to the "light-medium", lead us to the supposition that not only in mechanics, but also in electrodynamics, no properties of the phenomena correspond to the concept of absolute rest, but rather that for all coordinate systems for which the mechanical equations hold, the equivalent electrodynamical and optical equations hold also, as has already been shown for magnitudes of the first order. In the following we will elevate this guess to a presupposition (whose content we shall subsequently call the "Principle of Relativity") and introduce the further assumption, an assumption which is only apparently irreconcilable with the former one that light in empty space always propagates with a velocity V which is independent of the state of motion of the emitting body. These two assumptions are quite sufficient to give us a simple and consistent theory of electrodynamics of moving bodies on the basis of the Maxwellian theory for bodies at rest. The introduction of a "luminiferous ether" will be proved to be superfluous in so far, as according to the conceptions which will be developed, we shall introduce neither a "space absolutely at rest" endowed with special properties, nor shall we associate a velocity-vector with a point in which electro-magnetic processes take place. Like every other theory in electrodynamics, the theory to be developed is based on the kinematics of rigid bodies; since in the arguments of every theory, we have to do with relations between rigid bodies (co-ordinate system), clocks, and electromagnetic processes. An insufficient consideration of these circumstances is the cause of difficulties with which the electrodynamics of moving bodies has to fight at present.[3]"

We assume that the reader is familiar with the Galilean transformation used in classical physics to relate coordinates in 2 reference frames moving with velocity v with respect to each other, and with the classical velocity addition formula. As we expect you also know, Newton's laws are invariant under the Galilean transformation. Maxwell's laws, on the other hand, presented a paradox: Maxwell's equations imply that light propagates through the vacuum in any direction with speed $c = 1/\sqrt{\epsilon_0\mu_0} = 3 \cdot 10^8 m/s$. The problem was, that if light travelled at this speed in some frame S, then in any other frame moving with velocity v with respect to S, we must see light moving at $v' = c \pm v$. This implied that there was only one frame in which Maxwell's equations could hold (if the Galilean transformation were the correct transformation). Many physicists were trying to find this frame, the so-called 'ether frame'. As we see above, it was Einstein who realized that we could do away with the 'ether' business altogether with some surprising insights.

Einstein's Special Relativity theory is founded upon two postulates:

The Principle of Relativity: The laws by which the states of physical systems undergo change are not affected, whether these changes of state be referred to the one or the other of two systems in uniform translatory motion relative to each other.

The Principle of Invariant Light Speed: "... light is always propagated in empty space with a definite velocity c, which is independent of the state of motion of the emitting body." [3]

Let's unpack these statements. The first postulate states that the *all* the laws of physics are the same in all inertial frames of reference. This means that every physical theory should look the same, mathematically, to every inertial observer. The second postulate says that the speed of light in free space is a constant (c) in all inertial reference frames.

These two postulates will lead us to the conclusion that we must abandon the Galilean transformation in favour of the Lorentz transformations. Combined with other laws of physics, the two postulates of special relativity predict the equivalence of mass and energy, as expressed in E $= mc^2$. Furthermore, Special relativity reveals that c is not just the velocity of a certain phenomenon, namely the propagation of electromagnetic radiation, but rather a fundamental feature of the way space and time are unified as space-time.

In this introduction, we will only examine some outcomes of Special Relativity: time dilation, length contraction, and relativistic momentum. We suggest that interested readers for more information begin with the original paper by Einstein himself. Translations can be found here [12].

We must introduce the notion of an 'event', which is the basic building block of the theory. By an event, we mean an idealized occurrence in the physical world having extension in neither space nor time [5]. For example, the explosion of a firecracker could be represented by this idealized notion of an event. This notion of an event is analogous to the mathematical definition of a point: Euclid defined a point as "that which has no part" [6]. In other words, a point has no spatial dimension, but a location. The concept of an event in relativity is similar, but it is not defined in the Euclidean framework, it is defined in spacetime.

Time Dilation

We begin, as is usually the case when discussing relativity, with a simple thought experiment. Suppose there were a sapceship flying past you at some speed v. On this ship there is a light source, pointed upwards to the ceiling at a mirror, such that it reflects the light pulse into some optical detector which is attached to a stopwatch. You are standing on the ground, with excellent vision and a stopwatch.

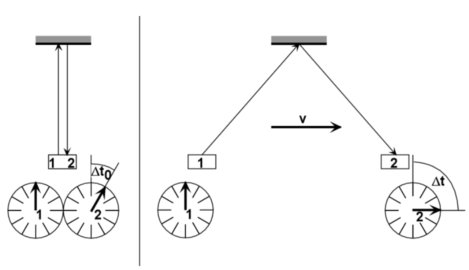

FIG. 3. The experimental set-up from different reference frames

When the light pulse is sent, the stopwatch begins counting, and is triggered to stop when it is received again by the optical detector. So, as the ship flies past you, a person on the ship flashes the light and measures the time the light takes to travel to the mirror and back. The point of view of an observer on the ship is on the left, and what you observe standing on the earth outside the ship is on the right. In the reference frame of the observer, the light travels a distance 2D at a speed c, thus the time measured is $\Delta t_0 = 2D/c$. You observe the same process, from the earth's reference frame. To you, the ship is moving. Since light travels at the same speed in all reference frames, you see the light travel the diagonal path shown in 3. Thus, the light travels a longer distance in your reference frame, so the time required for the light to hit the optical detector is greater than the time measured by the observer on the spaceship.

We calculate the time interval Δt which you record, as follows: In the time Δt, the spaceship travels $2L = v\Delta t$. Thus, the light travels a distance $d = 2\sqrt{D^2 + L^2}$, where $L = v\Delta t/2$, thus:

$$c = \frac{2\sqrt{D^2 + L^2}}{\Delta t} = \frac{2\sqrt{D^2 + (v\Delta t/2)^2}}{\Delta t}$$

We square both sides and solve for Δt to find:

$$\Delta t = \frac{2D}{c\sqrt{1 - v^2/c^2}}$$

And upon substituting in the time found in the spaceship, namely $\Delta t_0 = 2D/c$, we find:

$$\Delta t = \frac{\Delta t_0}{\sqrt{1 - v^2/c^2}} \geq \Delta t_0 \qquad (1)$$

Now, since the sending of light is one event, and the reception of the light is another event, we see that the time between the two events is different in each reference frame. Since $\sqrt{1 - v^2/c^2}$ is always less than one, you observe a longer time interval. Time is measured to pass more slowly in any moving reference frame as compared to your own. This result is referred to as time dilation.

the factor $1/\sqrt{1 - v^2/c^2}$ occurs so often in relativity, it is often given the shorthand symbol $\gamma = 1/\sqrt{1 - v^2/c^2}$. We note that γ is never less than one, and at most human speeds that γ is very, very close to 1. Only at speeds close to the speed of light is γ larger than 1. There remain a few things to clarify: The equation is only true when Δt_0 represents the interval between the two events in a reference frame where the two events occur at the same point in space. Δt_0 is referred to as the proper time. It follows that Δt represents the time interval between the two events in a reference frame which moves with speed v relative to the first reference frame. Furthermore, it is clear that the proper time is the shortest interval that any observer can measure between two events. One may infer many, many applications of time dilation, not the least of which is the ability to travel interstellar distances, if we were able to build propulsion systems which would produce relativistic speeds. Furthermore, we see that putting $v = c$ or $v > c$ would lead to absurd results. This suggests that v must always be strictly less than c.

Length Contraction

So far, we have established that time is relative, and dependent on one's reference frame. We now examine another fundamental quantity: length. Let us engage in another thought experiment. Suppose you are again standing on Earth watching a spacecraft fly by at speed v. This time, the spacecraft is flying to Mars, but you are able to observe it's journey with your superhuman vision. You have measured the distance between Earth and Mars to be L_0. The time you expect the trip will take, then, is $\Delta t = L_0/v$. Your friend is on the spacecraft. Since our two 'events'; arrival at Mars and departure from Earth, occur at the same point in space, this reference frame gives the proper time Δt_0. Thus, the time observed between the two events on the spaceship is $\Delta t_0 = \Delta t/\gamma$. We note that both observers measure the same relative speed: you observe the spacecraft moving away at v, while your friend on the ship observes Earth moving away from the ship at v. Since you both measure the same speed but different times, (notably the craft measures less time), the craft must measure a shorter distance. Let L be this distance. Then, $L = v\Delta t_0$. Thus, $L = v\Delta t/\gamma$. Or:

$$L = L_0/\gamma \leq L_0$$

This is a general result of the special theory of relativity and applies to lengths of objects as well as distances between them.

As I'm sure you expect, the length L_0 is called the proper length. This length is measured in the frame which is as rest with respect to the object. More concisely, when some object of proper length L_0 travels past some observer at speed v, the observer measures it's length to be $L = L_0/\gamma$. It is important to note that length contraction occurs only along the direction of motion.

The Lorentz Transformation

Our observations in the previous two sections have demonstrated that the Galilean transformation between two inertial reference frames is not valid. We use what we have learned to derive the Lorentz transformation. We will do this in three-dimensional space, and we will assume that the two reference frames are moving with velocity v with respect to each other only in the x direction. Let S' be the frame in which two events happen in the same location, and S be the frame in which they happen some distance apart. Then, we see immediately that $y' = y$ and $z' = z$. Using the length contraction formula, we see that $x - vt = x'/\gamma$, or $x' = \gamma(x - vt)$. We can repeat this argument in reverse, where the event is now at the same location in S, and we would get: $x = \gamma(x' - vt')$. Substituting this result into the previous and solving for t', we find: $t' = \gamma(t - vx/c^2)$. Putting these all together we arrive at the Lorentz transform:

$$\begin{aligned} y' &= y \\ z' &= z \\ x' &= \gamma(x - vt) \\ t' &= \gamma(t - vx/c^2) \end{aligned}$$

The Lorentz transformations express the properties of space and time that follow from Special Relativity.

Space-time

Let us digress for a moment to consider our current notions regarding space and time. We consider the above equations and note that x and t are inevitable coupled. Simply, relativity has shown that space and time considerations are not independent of one another. Contrary to popular belief, Einstein did not draw the conclusion that space and time could be seen as components of a single four-dimensional space-time fabric. That insight came from Hermann Minkowski (1864-1909), who announced it in a 1908 colloquium with the dramatic words: "Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality". [14] Put simply, Minkowski suggested that the three spatial coordinates should be combined with a time coordinate to form a four-dimensional space-time. The Lorentz transformations act as a kind of rotation on these four-vectors. If we let the fourth component $q_4 = ct$, then all components of our space-time vector have the same dimensions. General four-vectors are defined to be and vectors such that the values in two frames S and S' are related by the Lorentz transform matrix. We recognize, then, that the language of space-time is tensor mathematics.

Relativistic Momentum

We seek a definition of relativistic momentum that agrees, at non-relativistic speeds, with our classical concept of momentum, as we know the classical concept is an excellent approximation and enormously useful. We begin by defining the invariant mass: the invariant mass is that mass which is measured in any reference frame in which the object is at rest. We say that this is the mass, m that we will use, whatever the object's speed. We then define the four-velocity as follows: [15]

$$u = \gamma(d\mathbf{x}/dt, cdt/dt) = \gamma(\mathbf{v}, c)$$

Where $\mathbf{x}$ denotes the object's position vector. The reasons for this definition are not discussed here. We are now ready to define relativistic momentum, with the goal that it agrees with our old definition, $\mathbf{p} = m\mathbf{v}$ at nonrelativistic speeds, and that the total momentum of a closed system is conserved. Of course, our new definition must also be consistent with the postulates of Special Relativity. One can try to use the old definition, using the three-velocity $\mathbf{v}$, in two reference frames, applying the Lorentz transformations as appropriate, and they will find that this definition does not hold. However, it is an equally simple exercise to show that conservation of momentum does hold if we modify our old definition to use the four-velocity, $\mathbf{u}$. Thus: [15]

$$p = mu = (\gamma m\mathbf{v}, \gamma mc)$$

The spatial part, $\gamma m\mathbf{v}$, is referred to as the three-momentum. It is clear that is $v << c$, then γ is very close to 1 and we have our classical definition again. We note that this is a four-dimensional vector, so the law of conservation of momentum in component parts leads to four equations. The first three are the conservation of spatial momentum, which we are already familiar with. We will see that the fourth component of the four-momentum is actually the energy of the object, expressed as E/c, so we now have a compact way to express both the conservation of energy and momentum.

We define relativistic energy E of a freely moving object with four-momentum p as [15]

$$E = p_4 c = \gamma mc^2$$

Let's examine E for an object moving at non-relativistic speeds to see if this new definition agrees with our old one. Since v¡¡c, we can expand γ using the binomial series:

$$\gamma = [1 - (v/c)^2]^{-1/2} = 1 + 1/2(v/c)^2 + ...$$

So we find

$$E \simeq mc^2 + 1/2mv^2$$

Classical physicists would have thrown out the first term as an 'ignorable constant', since we can always define the zero of energy as we please. We shall not, however. We can see that in an elastic collision that $T^{fin} = T^{in}$, using this definition, so it follows that $M^{fin} = M^{in}$ as well. In *inelastic* collisions, where the total kinetic energy is not conserved, we still have a conservation of relativistic energy. This leads to the surprising result that $M^{fin} \neq M^{in}$. Thus, if relativistic energy is conserved, then mass is not conserved. [15].

We note that if an object is at rest in some reference frame, $\gamma = 1$, and we arrive at the famous equation, $E = mc^2$.

We would like to point out a few useful relations before we conclude our introduction to special relativity. First, note that $p \cdot p = -(mc)^2$, in an object's rest frame ($\mathbf{v} = 0$). Since both sides of this equation are invariant, we can conclude that this relation holds in *any* (inertial) reference frame. We can use this to re-write the relativistic energy:

$$E^2 = (mc^2)^2 + (\mathbf{p}c)^2$$

where $\mathbf{p}$ is the three-momentum. Also, since $p = mu = (\gamma m\mathbf{v}, \gamma mc) = (\mathbf{p}, E/c)$, we can deduce that [15]

$$\mathbf{v}/c = \mathbf{p}c/E$$

For some particle with no mass (say, a photon), $E = \mathbf{p}c$, and since $\mathbf{v}/c = \mathbf{p}c/E$, we have the satisfying conclusion that $v = c$. With this we will conclude our introduction to Special Relativity. There is much, much more that the interested reader can learn, and we would always encourage you to do so.

AN INTRODUCTION TO GENERAL RELATIVITY

Gravitational lensing is the deflection of light by matter. The computation of this deflection must be done within the theory of General Relativity. After we examine some basics of the theory, we will derive the Schwarzschild metric describing the space-time around a massive point-like object. This metric is what we use in our simulation to obtain the trajectories of light-rays.

Einstein used the idea that gravity is the same as any acceleration, and applied it to light. This was his insight. We will briefly review our notions of classical mechanics, and examine the notions Einstein introduced with his theories. We recall that Newton's laws are invariant when transforming from one inertial reference frame to another. In classical mechanics, it is assumed that there is a single universal time, and it is an elementary exercise

FIG. 4. A background galaxy (blue) being lensed by dark matter in foreground cluster. Hubble Space Telescope image

to find the coordinate transformation between two inertial reference frames moving with velocity V with respect to each other.

In order that the speed of light be the same in all reference frames, measurements of time and distance must depend on the motion of the observer. What is independent of the reference frame used is the interval between nearby events. For example, at spacetime coordiantes (t, x, y, z) and (t+dt, x+dx, y+dy, and z+dz), the differential interval is defined through:

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

Any other observer, using a different coordinate system (t', x', y', z'), will find the same interval.

The equivalence principle

We begin with Newton's gravity. Of course, we know: $\vec{F} = m\vec{a}$, for any force. In his studies of falling objects, Galileo observed $\vec{F_g} = m\vec{g}$, which Newton reformulated as $\vec{F_g} = \frac{mMG}{r^2}\hat{r}$. We note, upon equating the two equations : $\vec{F} = m\vec{a} = \frac{mMG}{r^2}\hat{r}$, that the mass of the object we are examining cancels out. It seems surprising, but the acceleration the object experiences is completely independent of it's own mass. In other words, a brick and a feather, in the absence of other forces, will fall to the surface of the earth (in a vacuum), at the same rate. Previously, it was considered that there were two 'types' of mass: inertial and gravitational. It was Einstein who formulated the (weak) equivalence principle:

"A little reflection will show that the law of the equality of the inertial and gravitational mass is equivalent to the assertion that the acceleration imparted to a body by a gravitational field is independent of the nature of the body. For Newton's equation of motion in a gravitational field, written out in full, it is:

(Inertial mass) (Acceleration) = (Intensity of the gravitational field) (Gravitational mass).

It is only when there is numerical equality between the inertial and gravitational mass that the acceleration is independent of the nature of the body." [1]

Since the resulting motion is the same for any object, independent of it's mass, Einstein postulated that gravitational effects can be ascribed to the intrinsic properties of the space-time geometry, rather than the action of a force in the Newtonian formulation. Einstein showed in general relativity that space-time becomes curved in the presence of matter, and through it's curvature determines the motion of bodies.

One of the most important of these is the Principle of Equivalence, which can be used to derive important results without having to solve the full equations of General Relativity. There are several ways to formulate the Principle of Equivalence, but one of the simplest is Einstein's original insight: he suddenly realized, while sitting in his office in Bern, Switzerland, in 1907, that if he were to fall freely in a gravitational field (think of a sky diver before she opens her parachute, or an unfortunate elevator if its cable breaks), he would be unable to feel his own weight. Einstein later recounted that this realization was the "happiest moment in his life", for he understood that this idea was the key to how to extend the Special Theory of Relativity to include the effect of gravitation. We are used to seeing astronauts in free fall as their spacecraft circles the Earth these days, but we should appreciate that in 1907 this was a rather remarkable insight.

Importance of the Equivalence Principle

An equivalent formulation of the Principle of Equivalence is that at any local (that is, sufficiently small) region in space-time it is possible to formulate the equations governing physical laws such that the effect of gravitation can be neglected. This in turn means that the Special Theory of Relativity is valid for that particular situation, and this in turn allows a number of things to be deduced

The fundamental principle of most direct physical relevance to general relativity is the equivalence principle, which asserts that gravitation is locally equivalent to acceleration. In practical terms this means that different falling bodies should follow the same trajectory in the same gravitational field, independent of their mass or internal structure, provided they are small enough not to disturb the environment or to be affected by tidal forces. To test this principle, one drops objects of different mass or composition in the same gravitational field and looks for differences in rate of fall. Such experiments have a long and fascinating history.

The Equivalence Principle

Einstein's happiest thought (1907): "For an observer falling freely from the roof of a house, the gravitational field does not exist" (left). Conversely (right), an observer in a closed boxsuch as an elevator or spaceship-cannot tell whether his weight is due to gravity or acceleration.

Soon after completing his special theory, Einstein had the "happiest thought of his life" (1907). It came while he was sitting in his chair at the patent office in Bern and wondering what it would be like to try to drop a ball while falling off the side of a building. Einstein realized that a person who accelerates downward along with the ball will not be able to detect the effects of gravity on it. An observer can "transform away" gravity (at least in the immediate neighbourhood) simply by moving to this accelerated frame of reference no matter what kind of object is dropped. Gravitation is (locally) equivalent to acceleration. This is the principle of equivalence.

To understand how remarkable the equivalence principle really is, imagine how it would be if gravity worked like other forces. If gravity were like electricity, for example, then balls with more charge would be attracted to the earth more strongly, and hence fall down more quickly than balls with less charge. (Balls whose charge was of the same sign as the earth's would even "fall" upwards.) There would be no way to transform away such effects by moving to the same accelerated frame of reference for all objects. But gravity is "matter-blind" it affects all objects the same way. From this fact Einstein leapt to the spectacular inference that gravity does not depend on the properties of matter (as electricity, for example, depends on electric charge). Rather the phenomenon of gravity must spring from some property of space-time.

Gravity as Curved Space-time Einstein eventually identified the property of space-time which is responsible for gravity as its curvature. Space and time in Einstein's universe are no longer flat (as implicitly assumed by Newton) but can pushed and pulled, stretched and warped by matter. Gravity feels strongest where space-time is most curved, and it vanishes where space-time is flat. This is the core of Einstein's theory of general relativity, which is often summed up in words as follows: "matter tells space-time how to curve, and curved space-time tells matter how to move". A standard way to illustrate this idea is to place a bowling ball (representing a massive object such as the sun) onto a stretched rubber sheet (representing space-time). If a marble is placed onto the rubber sheet, it will roll toward the bowling ball, and may even be put into "orbit" around the bowling ball. This occurs, not because the smaller mass is "attracted" by a force emanating from the larger one, but because it is traveling along a surface which has been deformed by the presence of the larger mass. In the same way gravitation in Einstein's theory arises not as a force propagating through space-time, but rather as a feature of space-time itself. According to Einstein, your weight on earth is due to the fact that your body is traveling through warped space-time!

While intuitively appealing, however, the rubber-sheet picture has its limitations. Mostly, these have to do with the fact that it allows us to visualize the spatial aspect of Einstein's theory, but not the temporal one. To see this, we need only remember that Newtonian gravity must be approximately valid, whatever Einstein says, and Newton tells us that bodies move in straight lines unless acted upon by a force. Why, then, do the orbits of planets around the sun on the rubber sheet appear so far from straight, if there is no attracting force reaching out through space-time to tug on them? The answer is that planetary trajectories are very nearly straight in space-time, not space. The worldline of the earth, for example, resembles a stretched-out spiral whose width in space is only one astronomical unit, but whose length in the time direction is measured in lightyears! Another way to appreciate the importance of the "time" in "space-time" is to apply the equivalence principle and ask whether the fact that we experience a gravitational field on the earth's surface is "equivalent" to stating that the earth's surface is continually accelerating outward. Obviously not, for we do not observe the earth to grow larger! The trouble is that, in speaking of the earth's surface, we have again lapsed into thinking of acceleration in spatial terms. On earth, where speeds are small compared to the speed of light and the gravitational field is weak, it turns out that nearly all of our weight arises due to the warping of time, rather than space. What this means in practice is that gravity on earth is "equivalent" to acceleration mostly in the sense that clocks on the surface run more slowly than clocks in outer space.

General Relativity General relativity is based physically on the equivalence principle, but the theory also has a second, more mathematical foundation. Known as the principle of general covariance, it is the requirement that the law of gravitation be the same for all observers even accelerating ones regardless of the coordinates in which it is described. (It is for this reason that Einstein named his new theory "general", as opposed to "special" relativity he dropped the earlier restriction to uniformly moving observers.) This proved to be the most difficult challenge that Einstein ever faced. As he later said, to express physical laws without coordinates is like "describing thoughts without words". Einstein was obliged to master the abstract mathematics of surfaces and their description in terms of tensors, a field pioneered by Carl Friedrich Gauss (1777-1855) and generalized to higher dimensions and more abstract spaces by Georg Friedrich Bernard Riemann (1826-1866). In this labor he was aided above all by his friend the mathematician Marcel Grossmann (1878-1936). Another mathematician named David Hilbert (1862-1943) nearly beat him to his final equations.

But general relativity is above all Einstein's achievement, and the phrase "Einstein's space-time" is entirely appropriate. No theory of comparable significance before or since is more nearly due to the struggle of a single scientist. At the end of 1915 Einstein wrote to a friend that he had succeeded at last, and that he was "content but rather worn out". He later described this period as follows: "The years of searching in the dark for a truth that one feels but cannot express, the intense desire and the alternations of confidence and misgiving until one breaks through to clarity and understanding, are known only to those who have themselves experienced them".

tidal forces

The equivalence principle is limited by tidal forces, in that one cannot truly and simply replace a gravitational field with a uniform acceleration. We can only do so locally, in space and time, to replicate the effect of being small enough and local enough in time to avoid sensing the vraiation in the gravitational field. ie we can completely replace the gravitaitonal field with an accelerated frame of reference in small localities in space time. Mass corresponds to a divergence in gravitational field, and it is this feature which inhibits us from replacing a whole gravitational field with an accelerated fame of reference.

geometry

Mathematically, we can specify a geometry by specifying the distance between any two points on our object in question. On a flat surface (in 2d), we have $ds^2 = dx^2 + dy^2$, in cartesian coordinates. What about in other coordinates? We would, like in a familiar cartesian grid, have points of constant x and y, but these lines would no longer be straight, orthogonal lines, they would form a curvy grid. We express the distance formula in these general coordinates as : $ds^2 = g_{11}(x)dx^2 + 2g_{12}(x)dxdy + g_{22}(y)dy^2$, where the term dxdy reveals that the coordinate axes are not perpendicular. The coefficients g_{ii} are position dependent, and refer to the 'spread' of the coordinates. The functions gii describe the metric. We note that if we can find some coordinates which simplify to the form $ds^2 = dx^2 + dy^2$, we have a flat geometry. There are many geometries which cannot be simplified to this form. These are, obviously, curved geometries. How do we know if a given metric describes something flat? We find it's curvature. Curvature is the obstruction to flattening out the coordiantes. Tidal coordinates are the obstruction to replacing a gravitational field by coordinate transformation to an accelerated frame of reference. We will find the curvature of space time is the same as tidal forces.

Curved Spaces

Supposing we have some n-dimensional space to which we attach any set of general coordinates: $(q_1, q_2, ...q_n)$. Now suppose there is some field which is a function of position, which varies in space. This variation corresponds to a non-zero divergence.

We have already abandoned the notion of space and time being separate in the special theory of relativity. We adopt general curvilinear coordinates to express 2:

$$ds^2 = g_{ik}dx^i dx^k \tag{3}$$

Where Einstein summation notation is employed. The metric g_{ik} describes how distances are measured in the spacetime. On the basis of the weak equivalence principle, we can transform a free-falling reference system to one in which the (local) spacetime reduces to the Minkowski one (a flat spacetime). This is the spacetime of special relativity. In the case of a gravitational field, there exist no coordinates for which the metric can be reduced to the Minkowski form, so we conclude that spacetime is curved. (In General Relativity, this curvature of spacetime is related to the energy momentum tensor of the matter, and turns out to depend on the second derivatives of the metric. The presence of matter requires that some of these second derivatives be non-zero.) In General Relativity, we must also abandon the concept of privileged inertial reference frames where we can synchronize clocks at different locations. In the presence of gravitational fields, clocks will run at different rates in different locations, and all physics must then be defined locally.

Tensors in curved spacetime

In General Relativity, we write all physical laws in such a way that they take the same form in any reference system, reducing to the expressions of inertial frames of Special Relativity. Thus, we state the laws using generally covariant expressions, such that they retain the same form after a general coordinate transformation. We do this using tensors, as a relation (written as an equality) between two tensors will hold in any reference system.

The simplest tensor is a scalar, which is any function which under a coordinate transform (t, x, y, z) $\rightarrow$ (t', x', y', z') retains it's value. In general any set of four quantities A^i transforming as the differential of coordinates, ie. $A'^k = \partial x'^k/\partial x^i A^i$, constitues the components of a contravariant vector. Any set of four quantities A_i transforming as $A'^k = \partial x/\partial x'^k A_i$ constitute the components of a covariant vector.

Scalars are tensors of rank zero, vectors are tensors of rank one. Tensors of higher order can be defined in an analogous way. For example, a contravariant tensor of rank two is a set of 16 quantities transforming as the product of two contravariant vectors, ie:

$$A'^{ik} = \frac{\partial x^i}{\partial x^l}\frac{\partial x'^k}{\partial x^m}A^{lm}$$

We can generalize these transformation laws to a tensor with n contravariant indices and m covariant ones: There is a factor $\partial x'^i/\partial x^k$ for each contravariant index and a factor of $\partial x^k/\partial x'^i$ for each covariant one.

Coordinate Transformation

Say we have some set of coordinates x and another set for the same space y. We know the x_i, and we transform to y according to some relation. We write the y's as such, since they are the result of a transformation: $y^n = y^n(x)$. Now, suppose we have some field Φ (any position-dependent function) defined in this space. We know that: $d\Phi = \partial\Phi/\partial x^m dx^m$. Then, if we know each partial derivative wrt x, it follows that:

$$\frac{\partial\Phi}{\partial y^n} = \frac{\partial\Phi}{\partial x^m}\frac{\partial x^m}{\partial y^n}$$

, where einstein summation notation is employed.

tensors continued

Tensors have components, like vectors. Their components change when you go from one coordinate system to another. We note that when we have some vector in a particular coordinate system, and we want to express it in some other coordinate system, we must transform it's components. The components always depend on the coordinate frame. We will look at the relation of vectors and tensors when changing reference frames. In a d-dimensional space, a vector has d components. A vector is a special case of a tensor: it only has one index (position). It is a tensor of rank one. Now say we want to include both position and velocity. To hold all of this information, it is most convenient to use a matrix or size 2d. We can think of this as a starting place of tensors.

Now, say we have some field $\Phi(y)$ where y is some general set of coordinates. The scalar is invariant under any coordinate transform, as the field still has all the same values at each location, no matter how wach location is now related.

Let's examine the displacement vector : $\rightarrow r = (dx^1, dx^2, ...dx^d)$, in d dimensional space. Now, we have some other displacement vector : $\rightarrow r = (dy^1, dy^2, ...dy^d)$ in the y set of coordinates. How are these related? It should be clear that

$$dy^n = \frac{\partial y^n}{\partial x^m}dx^m$$

We recognize this as a contravariant vector. More generally,

$$A'^{ik} = \frac{\partial x^i}{\partial x^l}\frac{\partial x'^k}{\partial x^m}A^{lm}$$

It is important to note that the lhs is a vector, and the rhs is a matrix multiplying a column vector.

Now imagine we have 2 vectors, say A^m and B^n, and we multiply these two together.

MANIFOLDS

In mathematics (specifically in differential geometry and topology), a smooth manifold is a subset of Euclidean space which is locally the graph of a smooth (perhaps vector-valued) function. A more general topological manifold can be described as a topological space that on a small enough scale resembles the Euclidean space of a specific dimension, called the dimension of the manifold. Thus, a line and a circle are one-dimensional manifolds, a plane and sphere (the surface of a ball) are two-dimensional manifolds, and so on into high-dimensional space. More formally, every point of an n-dimensional manifold has a neighborhood homeomorphic to an open subset of the n-dimensional space R^n.

SCHWARZSCHILD METRIC AND THE GEODESIC EQUATIONS

The solutions of the einstein field equations are metrics of spacetime. The metric tensor (simply, the metric) is the fundamental object of study. The metric captures all the geometric and causal structure of spacetime. Mathematically, spacetime is represented by a 4-dimensional differentiable manifoldM and the metric is given as a covariant, second-rank, symmetric tensor on M, conventionally denoted by g. Moreover the metric is required to be nondegenerate with signature (-+++). A manifold M equipped with such a metric is called a Lorentzian manifold.

The Schwarzschild solution is a spherically symmetric, vacuum solution solution of the einstein field equations. Karl Schwarzschild found the solution while serving in WW1 on the russian front. He sent the solution to Einstein, who had only produced an approximate solution at this point. Einstein responded:

"I have read your paper with the utmost interest. I had not expected that one could formulate the exact solution of the problem in such a simple way. I liked very much your mathematical treatment of the subject. Next Thursday I shall present the work to the Academy with a few words of explanation. [2] "

The solution describes the spacetime [g field] around a spherical, uncharged, non-rotating mass.

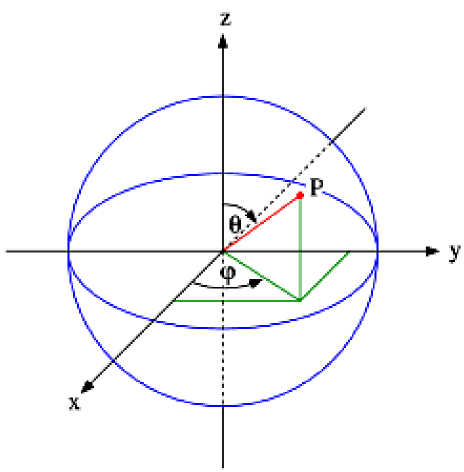

FIG. 5. Coordinate System

We will take our units such that c=1.

The timelike formulation of the Schwarzschild metric is just

$$d\tau^2 = w(r)dt^2 - v(r)dr^2 - r^2 d\theta^2 - r^2 sin^2(\theta) d\varphi^2$$

where $w(r) = 1 - \frac{2GM}{r}$ in units of c = 1 and $v(r) = 1/w(r)$.

Solving for the Christoffel symbols we get the following system of geodesic equations

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0$$

$$t: \rightarrow \frac{d^2t}{d\lambda^2} + \frac{2MG}{r(r-2MG)}\frac{dt}{d\lambda}\frac{dr}{d\lambda} = 0$$

$$r: \rightarrow \frac{d^2r}{d\lambda^2} + \frac{MG}{r^2}(1 - \frac{2GM}{r})(\frac{dt}{d\lambda})^2 + \frac{MG}{r(2MG-r)}(\frac{dr}{d\lambda})^2 + (2MG -$$

$$\theta: \rightarrow \frac{d^2\theta}{d\lambda^2} + \frac{2}{r}\frac{d\theta}{d\lambda}\frac{dr}{d\lambda} - sin\theta cos\theta(\frac{d\varphi}{d\lambda})^2 = 0$$

$$\varphi: \rightarrow \frac{d^2\varphi}{d\lambda^2} + \frac{2}{r}\frac{d\varphi}{d\lambda}\frac{dr}{d\lambda} + 2cot\theta\frac{d\varphi}{d\lambda}\frac{d\theta}{d\lambda} = 0$$

Where λ is some affine parameter along the geodesic. Denoting $\dot{x^\mu} = dx^\mu/d\lambda$ and setting $\theta = \frac{\pi}{2}$ for simplicity. We get

$$t: \rightarrow \ddot{t} + \frac{2MG}{r(r-2MG)}\dot{t}\,\dot{r} = 0$$

$$r: \rightarrow \ddot{r} + \frac{MG}{r^3}(r - 2MG)\dot{t}^2 + \frac{MG}{r(2MG-r)}\dot{r}^2 + (2MG - r)\dot{\varphi}^2 = 0$$

$$\varphi: \rightarrow \ddot{\varphi} + \frac{2}{r}\dot{\varphi}\,\dot{r} = 0$$

Solving the t equation for $\dot{t}$ and the φ equation for $\dot{\varphi}$ we get the following conserved quantities

$$\dot{t} = \frac{E}{w(r)} \qquad \dot{\varphi} = \frac{L}{r^2}$$

Where E and L are constants corresponding in some form to the energy and angular momentum of the particle.

TIME-LIKE AND LIGHT-LIKE GEODESICS

For timelike geodesics, we can set $\lambda = \tau$ using the proper time as an affine parameter. From the metric we then can determine an expression for $\dot{r}$ as follows

$$d\tau^2 = w(r)dt^2 - v(r)dr^2 - r^2 d\varphi^2$$

$$1 = w(r)\dot{t}^2 - v(r)\dot{r}^2 - r^2\dot{\varphi}^2$$

$$w(r) = E^2 - \dot{r}^2 - w(r)\frac{L^2}{r^2}$$

$$\dot{r}^2 = E^2 - w(r)\left(1 + \frac{L^2}{r^2}\right)$$

So for timelike Schwarzschild geodesics we have

$$\boxed{(\frac{dr}{d\tau})^2 = E^2 - w(r)\left(1 + \frac{L^2}{r^2}\right)} \qquad \boxed{\frac{dt}{d\tau} = \frac{E}{w(r)}}$$

$$\boxed{\frac{d\varphi}{d\tau} = \frac{L}{r^2}}$$

For lightlike geodesic we set $d\tau = 0$ and leave λ as some affine parameter

$$d\tau^2 = w(r)dt^2 - v(r)dr^2 - r^2d\varphi^2$$

$$0 = w(r)\dot{t}^2 - v(r)\dot{r}^2 - r^2\dot{\varphi}^2$$

$$0 = E^2 - \dot{r}^2 - w(r)\frac{L^2}{r^2}$$

$$\dot{r}^2 = E^2 - w(r)\frac{L^2}{r^2}$$

So for lightlike Schwarzschild geodesics we have

$$\boxed{(\frac{dr}{d\lambda})^2 = E^2 - w(r)\frac{L^2}{r^2}} \qquad \boxed{\frac{dt}{d\lambda} = \frac{E}{w(r)}} \qquad \boxed{\frac{d\varphi}{d\lambda} = \frac{L}{r^2}}$$

MINKOWSKI SPACE

In order to determine the constants E and L in timelike and lightlike geodesics, we note that for sufficiently large radius the Schwarzschild metric resembles the Minkowski metric η_{ij}. Therefore, if we consider only paths originating relatively far from the gravitating body we can approximate the local metric as η_{ij}.

In planar coordinates, the Minkowski metric is represented as

$$\eta_{ij} \to d\tau^2 = dt^2 - dr^2 - r^2d\varphi^2$$

We note that $dr^2 + r^2d\varphi^2 = d\vec{r} \cdot d\vec{r}$ and the linear velocity $\vec{v} = \frac{d\vec{r}}{dt}$. Therefore we see that $|\vec{v}|^2dt^2 = v^2dt^2 = d\vec{r} \cdot d\vec{r}$ and hence

$$d\tau^2 = dt^2 - v^2dt^2 = (1 - v^2)dt^2$$

For reference, note that we have scaled the space coordinates such that $c = 1$ hence v will be a proportion of the speed of light. Let us define $\gamma^2 = \frac{1}{1-v^2}$ then we see

$$\frac{dt}{d\tau} = \gamma$$

Let us consider now $\frac{d\varphi}{d\tau}$

$$\frac{d\varphi}{d\tau} = \frac{d\varphi}{dt}\frac{dt}{d\tau} = \gamma\frac{d\varphi}{dt}$$

Hence we arrive at the some useful properties of Minkowski space

$$\boxed{\frac{dt}{d\tau} = \gamma} \qquad \boxed{\frac{d\varphi}{d\tau} = \gamma\frac{d\varphi}{dt}} \qquad \boxed{\gamma^2 = \frac{1}{1 - v^2}}$$

FULL FORM OF THE TIMELIKE GEODESIC EQUATIONS

Let us consider first the timelike geodesics. Recall we found that the equations of motion are

$$\boxed{(\frac{dr}{d\tau})^2 = E^2 - w(r)\left(1 + \frac{L^2}{r^2}\right)} \qquad \boxed{\frac{dt}{d\tau} = \frac{E}{w(r)}}$$

$$\boxed{\frac{d\varphi}{d\tau} = \frac{L}{r^2}}$$

Assuming that the path originates far from the gravitating body, we can extract our constants from the Minkowski space properties. We have

$$\dot{t_0} = \frac{E}{w(r_0)} \to E = w(r_0)\gamma_0 = \frac{w(r_0)}{\sqrt{1 - v_0^2}}$$

$$\dot{\varphi}_0 = \frac{L}{r_0^2} \rightarrow L = r_0^2 \gamma_0 \frac{d\varphi}{dt}$$

We can find $\frac{d\varphi}{dt}$ by considering the angle ϕ between the gravitating mass and the initial path direction as seen by the source. The φ component of the velocity is found as

$$r\dot{\varphi} = vsin\phi \rightarrow \dot{\varphi}_0 = \frac{v_0 sin(\phi_0)}{r_0}$$

Hence the constant L is just

$$L = r_0^2 \gamma_0 \frac{v_0 sin(\phi_0)}{r_0} = \frac{v_0 r_0 sin(\phi_0)}{\sqrt{1-v_0^2}}$$

Therefore the full expression of timelike geodesics in the Schwarzschild solution is satisfied by the constants

$$\boxed{E \approx \frac{w(r_0)}{\sqrt{1-v_0^2}}} \qquad \boxed{L \approx \frac{v_0 r_0 sin(\phi_0)}{\sqrt{1-v_0^2}}}$$

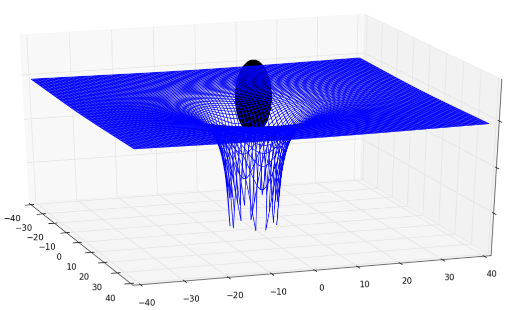

FIG. 6. A plot of the Spacetime surrounding a massive spherical object given by the Schwarzchild Solution

FULL FORM OF THE LIGHTLIKE GEODESIC EQUATIONS

Let us now consider the lightlike geodesics. We established the equations of motion to be

$$\boxed{(\frac{dr}{d\lambda})^2 = E^2 - w(r)\frac{L^2}{r^2}} \qquad \boxed{\frac{dt}{d\lambda} = \frac{E}{w(r)}} \qquad \boxed{\frac{d\varphi}{d\lambda} = \frac{L}{r^2}}$$

What are the constants E and L? Our previous approach of utilizing the Lorentz transforms from the Minkowski metric will not work as we will come across a singularity as $\gamma = \frac{1}{\sqrt{1-v^2}} = \frac{1}{0} = \infty$. We will instead utilize our choice of affine parameter λ to simplify the initial conditions. Let λ be defined such that $\frac{dt}{d\lambda} = 1$ initially. We then observe the following.

$$E = w(r_0)\dot{t_0} = w(r_0)$$

$$L = r^2 \frac{d\varphi}{d\lambda} = r^2 \frac{d\varphi}{dt}\frac{dt}{d\lambda}$$

In the previous section we found that $\frac{d\varphi}{dt}$ has an initial value of $\frac{v_0 sin(\phi_0)}{r_0}$. Here we have $v_0 = c = 1$ so the constant L is just

$$L = r_0^2 \frac{d\varphi}{dt}\frac{dt}{d\lambda} = \frac{r_0^2 sin(\phi_0)}{r_0} = r_0 sin(\phi_0)$$

So we see that the full full expression of timelike geodesics in the Schwarzschild solution is satisfied by the constants

$$\boxed{E \approx w(r_0)} \qquad \boxed{L \approx r_0 sin(\phi_0)}$$

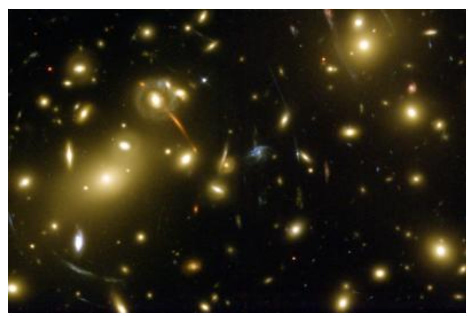

FIG. 7. Galaxy Cluster Abell 2218's Gravitational Lens. Hubble Space Telescope image

Thanks Dr. Sigurdson and Kevin, for being awesome.

[1] A. Einstein. How I Constructed the Theory of Relativity, Translated by Masahiro Morikawa from the text recorded in Japanese by Jun Ishiwara, Association of Asia Pacific Physical Societies (AAPPS) Bulletin, Vol. 15, No. 2, pp. 17-19 (April 2005). Einstein recalls events of 1907 in talk in Japan on 14 December 1922.

[2] Eisenstaedt, The Early Interpretation of the Schwarzschild Solution, in D. Howard and J. Stachel (eds), Einstein and the History of General Relativity: Einstein Studies, Vol. 1, pp. 213-234. Boston: Birkhauser, 1989.

[3] Albert Einstein (1905) "Zur Elektrodynamik bewegter Krper", Annalen der Physik 17: 891; English translation On the Electrodynamics of Moving Bodies by George Barker Jeffery and Wilfrid Perrett (1923); Another English translation On the Electrodynamics of Moving Bodies by Megh Nad Saha (1920).
[4] Kolecki, Glenn Research Center, NASA, 2002 http://www.grc.nasa.gov/WWW/k-12/Numbers/Math/documents/Tensors_TM2002211716.pdf
[5] Geroch, R. 'General Relativity: from A to B'. University of Chiacago Press, 1978.
[6] Heath, T. L. "Euclid's Elements" green lion press, 2007
[7] Dyson F. W., Eddington A. S., and Davidson C. R., Mem. Roy. Astron. Soc. 62, (1920)
[8] Chwolson, O. Astr. Nachricken. 221 (1924)
[9] Zwicky, F. Phys. Rev. 51 (1937)
[10] Walsh D., Carswell R. F., and Weymann R.J. Nature, 279 (1979)
[11] Mollerach, S. and Roulet E. "Gravitational Lensing and Microlensing" World Scientific Publishing, 2002.
[12] http://en.wikisource.org/wiki/On_the_Electrodynamics_of_Moving_Bodies
[13] http://hubblesite.org/gallery/album/exotic/pr2011025a/
[14] http://einstein.stanford.edu/SPACETIME/spacetime5.html
[15] Taylor, classical Mechanics