Supplementary Lecture Series 11:
General Relativity for Physics Students

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreward</td>
<td>2</td>
</tr>
<tr>
<td>1. The Equivalence Principle, Gravity, and Apparent Forces</td>
<td>2</td>
</tr>
<tr>
<td>2. Motion in a Rotating, Relativistic Reference Frame</td>
<td>11</td>
</tr>
<tr>
<td>3. Tidal Forces, non-Euclidean Geometry and “Local Inertial Reference Frames”</td>
<td>19</td>
</tr>
<tr>
<td>4. Gravitational Red Shift</td>
<td>23</td>
</tr>
<tr>
<td>5. The Twins Again</td>
<td>33</td>
</tr>
<tr>
<td>6. Similarities and Differences of Electromagnetism and Gravity</td>
<td>36</td>
</tr>
<tr>
<td>7. The Equation of Motion of Particles in Curved Space-Time</td>
<td>46</td>
</tr>
<tr>
<td>8. Covariant Derivatives and Covariant Vector Fields</td>
<td>53</td>
</tr>
<tr>
<td>9. The Equivalence Principle, Metric Compatibility and Christoffel Symbols</td>
<td>56</td>
</tr>
<tr>
<td>10. The Curvature of Space-Time</td>
<td>60</td>
</tr>
<tr>
<td>11. The Einstein Field Equation: “Curvature ~ Energy-Density”</td>
<td>64</td>
</tr>
<tr>
<td>a. Geodesic Equation of Motion as a Consequence of the Field Equations</td>
<td>68</td>
</tr>
<tr>
<td>b. Linearized Gravity</td>
<td>71</td>
</tr>
<tr>
<td>12. The Schwarzschild Metric and Black Hole</td>
<td>77</td>
</tr>
<tr>
<td>13. Circular Orbital Motion in the Schwarzschild metric</td>
<td>94</td>
</tr>
<tr>
<td>14. Relativistic Tidal Forces</td>
<td>100</td>
</tr>
<tr>
<td>15. The Discovery of Gravitational Waves</td>
<td>103</td>
</tr>
<tr>
<td>a. Freely Propagating Gravity Waves in Minkowski Space Time</td>
<td>107</td>
</tr>
<tr>
<td>b. How to Detect Gravity Waves</td>
<td>110</td>
</tr>
<tr>
<td>c. The Radiation and Detection of Gravity Waves from Binary Systems</td>
<td>112</td>
</tr>
<tr>
<td>17. Contrasting Special and General Relativity, The Cosmological Constant and Dark Energy</td>
<td>120</td>
</tr>
<tr>
<td>Special Topic A. The Isotropic Metric and Linearized General Relativity</td>
<td>126</td>
</tr>
<tr>
<td>Special Topic B. Slowly Rotating Stars, Frame-Dragging and Lense-Thirring Precession</td>
<td>128</td>
</tr>
<tr>
<td>Special Topic C. Gravito-Magnetism</td>
<td>133</td>
</tr>
<tr>
<td>Special Topic D. Rotating Stars, Black Holes and an Introduction to the Kerr Metric</td>
<td>136</td>
</tr>
<tr>
<td>Special Topic E. Newton in Orbit around a Schwarzschild Black Hole</td>
<td>144</td>
</tr>
</tbody>
</table>
Foreward

This lecture series supplements the textbook “Special Relativity, Electrodynamics and General Relativity: From Newton to Einstein” (ISBN: 978-0-12-813720-8) for undergraduate physics majors at the junior or senior level, by the author of this lecture series, John B. Kogut.

It accompanies a set of Video Lectures that are described on the web site https://relativitydoctor.com/. The Video Lectures closely follow the text of these lectures.

There are also two accompanying Video Lecture series on Special Relativity, and Classical Differential Geometry. Results from those lectures are referred to and used in this lecture series.

The three lecture series cover and supplement the material in the textbook “Special Relativity, Electrodynamics and General Relativity: From Newton to Einstein” (ISBN: 978-0-12-813720-8).

1. The Equivalence Principle, Gravity, and Apparent Forces

Everyone knows the story of Isaac Newton and the falling apple. There was a plague in Great Britain, so the students were sent out of the cities to reduce their chances of catching the contagion.

Supposedly Newton relaxed under an apple tree, contemplating the current ideas of mechanics. When an apple fell on his head, he invented the idea of the gravitational force—Earth's enormous mass exerted a force on the apple, breaking its stem and causing a collision with Newton's precious
head. This event led Newton, over the course of later months back at the university, to the force law of gravity,

\[ F_{12} = -G \frac{m_1 m_2}{r_{12}^2} \hat{r}_{12}, \]  

(1.1)

where \( G \) is Newton's constant (\( G \approx 6.67 \times 10^{-11} \text{ Nm}^2/\text{kg}^2 \)), which sets the scale for gravitational forces; \( m_1 \) and \( m_2 \) are the masses, which are a distance \( r_{12} \) apart; \( F_{12} \) is the force that \( m_1 \) exerts on \( m_2 \); and \( \hat{r}_{12} \) is the unit vector pointing from \( m_1 \) to \( m_2 \). Newton arrived at the inverse square character of the force law, Eq. (1.1), to explain the extensive planetary data accumulated by Kepler and others. Choosing \( m_1 \) to be the mass \( M_s \) of the Sun and \( m_2 \) to be the mass of Earth \( M_e \), the equation of motion of the Earth around the Sun is given by Newton's Second Law, force equals mass times acceleration (see also Figure 1.1).

\[ M_e \ddot{r}_e = -G \frac{M_s M_e}{|r_e - \hat{r}_{es}|} \hat{r}_{es}. \]  

(1.2)

Fig. 1.1 The Coordinate system for the earth moon system
A crucial element of Eq. (1.2) is the fact that the mass of the accelerating body, Earth in this case, cancels out of the equation of motion. We describe this by saying that the inertial mass, the mass on the left-hand side of Newton's Second Law (mass times acceleration equals force), equals the gravitational mass, the mass in the gravitational force expression. It is said that Galileo was the first physicist to investigate this point, before Newton codified classical dynamics, and establish the equality of these two masses experimentally. Galileo, through his assistants, dropped masses off the Leaning Tower of Pisa and observed that they accelerated identically in the gravitational field provided by Earth. Modern experiments pioneered by Lörand Eötvös in the early days of the 20th century and many others in modern times have established the equality of the gravitational and inertial masses to high precision, better than one part in $10^{13}$! The axiom that the two masses are strictly identical evolved into a central ingredient in the soon-to-be famous Equivalence Principle. Under Einstein, the Equivalence Principle developed into the statement that there is no way to distinguish the local effects of a uniform gravitational field from those in an accelerating reference frame in otherwise empty, force-free space. The Equivalence Principle allows us to understand accelerating reference frames in terms of gravity and gravity in terms of accelerating reference frames. This principle is explained and discussed in much greater detail as we journey forth.

The Equivalence Principle and the inverse square law of gravity are both under constant experimental scrutiny by high precision experiments. We accept both ideas as exact throughout our discussions. However, if one or both should fail ever so slightly, many topics of theoretical physics would need fundamental changes.
Classical physicists understood that the Principle of Equivalence made gravity a very special phenomenon. Other forces (such as electrostatics) or mechanical devices (such as springs) produce accelerations that are inversely proportional to the mass of the body. However, there are forces besides gravity that are familiar from day-to-day experience that produce accelerations that are independent of the mass of the body. These are called “apparent” forces and are strictly geometrical in origin. For example, when you drive a car and accelerate from a stop sign or decelerate at a red traffic light, you experience forces of this type. Another example consists of centripetal and Coriolis forces. These are the forces that occur when you measure acceleration in a rotating coordinate system, called a non-inertial frame of reference. Recall that an inertial frame is one in which an isolated body moves in a straight line at constant velocity. To change the body's velocity, a force must be applied. The velocity of the body changes according to Newton's Second Law, $f = ma$, where $f$ is the applied force and $a$ is the body's acceleration, $a = d^2r/dt^2$. A simple example of a rotating, non-inertial frame is afforded by a turntable spinning at a constant angular velocity $\omega$ in an otherwise inertial environment. From the perspective of a coordinate system fixed in the rotating turntable, a body moving at a constant velocity in the inertial frame is accelerating. Clearly this acceleration is independent of the body's mass and is purely a result of the coordinate transformation between the inertial and the rotating non-inertial frames.

Let’s begin with a few words about the centrifugal and Coriolis accelerations in the context of Newton's world. Later we will revisit rotating reference frames in the context of general relativity.

Consider Newton’s second law in a rotating frame, described by plane polar curvilinear coordinates $(r, \theta)$. The relation between the Cartesian coordinates in the inertial frame and the rotating plane polar coordinates is,
\[ x = r \cos(\theta + \omega t) \quad y = r \sin(\theta + \omega t) \]  \hspace{1cm} (1.3)

where \( \omega \) is the rate of rotation of the turntable. An exercise in kinematics produces the equations of motion for the radial and angular variables complete with centripetal forces, Coriolis forces and external forces per unit mass \((f_r, f_\theta)\),

\[
\frac{d^2 r}{dt^2} - r \left( \omega + \frac{d\theta}{dt} \right)^2 = f_r
\]

\[
\frac{d^2 \theta}{dt^2} + \frac{2}{r} \frac{dr}{dt} \left( \omega + \frac{d\theta}{dt} \right) = f_\theta
\]

It is instructive, but somewhat tedious, to begin with Newton’s second law written in Cartesian coordinates,

\[
\frac{d^2 x}{dt^2} = f_x \quad \frac{d^2 y}{dt^2} = f_y
\]

apply the transformation Eq. (1.3) and arrive at Eq. (1.4). We will develop systematic ways to do such calculations very easily in lecture 7 and will illustrate them with this particular example.

This exercise will show us that there are two sources of “apparent” forces in the equations of motion Eq. (1.4): a. the use of curvilinear coordinates, the case \( \omega = 0 \), and b. the non-inertial character of the rotating frames, \( \omega \neq 0 \). There are clearly two distinct geometrical effects and are usually discussed separately. We will not follow tradition here because they enter the equations of motion in the same fashion. When we develop the equations of general relativity we will see that sources of energy-momentum produce curved space-time and non-inertial frames, and that gravitational “forces” enter equations such as Eq. (1.4) on the left hand side as “apparent” forces and not on the right hand side where other forces, such as the Lorentz force of electromagnetism will appear!

Now we come to the good stuff. Einstein argued that gravity is also an apparent force.!
acceleration $-g$ independent of its mass. The Equivalence Principle states that this problem is indistinguishable from the motion of a body in a frame free of gravity but accelerating upward at the rate $g$. These facts were well appreciated by classical physicists in Newton's era, but this statement of the Equivalence Principle was popularized and pursued in the context of the relativistic theory of space-time by Einstein. Einstein posed the equivalence—the impossibility of distinguishing the physics in a constant gravitational field from that in an accelerating frame—by imagining a physicist doing experiments in an elevator in otherwise empty space time. The elevator is accelerating upward at a rate $g$, and, Einstein claimed, if the elevator has no windows so the physicist cannot see the tricks being played on him, there is no local experiment he can run that can distinguish this environment from one at rest on the surface of a planet where gravity generates the approximately uniform acceleration $g$. (On the surface of Earth, Newton's gravitational force law gives $g = GM_e/R_e^2 \approx 9.8 \text{ m/s}^2$, where $R_e$ is the radius of Earth, neglecting all other planetary masses.)

Is uniform gravity really an apparent force much like centripetal and Coriolis forces, familiar from our experiences with turntables? Can we transform to another coordinate system where the apparent force vanishes identically? In the case of a constant gravitational field, we can consider an observer in free fall in that environment. According to the Equivalence Principle, this free-falling frame of reference is force-free and is an inertial frame where isolated masses travel along straight lines.

The idea that gravity can be transformed away locally by passing to a free-falling, inertial frame of reference is very useful. Since we know that masses move along straight lines in inertial frames, we can use the Equivalence Principle to solve any mechanics problem in a small enough region of space time where a given gravitational field can be treated as uniform—just consider the
motion from the perspective of a freely falling frame where special relativity holds, solve the problem, and finally map it back to the coordinates an observer would use at rest in the gravitational field. (It is important here to check that there are no external non-gravitational forces, such as electrostatics, in the environment. These forces are not “apparent” and cannot be transformed away by a slick choice of reference frame.)

But, as emphasized by Einstein, one can also go beyond ordinary mechanics problems because the Equivalence Principle applies to any process. An interesting application concerns the deflection of light by a gravitational field. Because light moves along straight lines in inertial frames, the Equivalence Principle implies that it experiences an acceleration when it moves transverse to a gravitational field, as shown in Figure 1.2.

Consider an initially horizontal beam of light in the earth’s gravitational field, labeling $x$ as the horizontal position and $y$ the vertical position. Then the Equivalence Principle predicts the local bending of light,
\[
\frac{d^2 y}{dx^2} \approx \frac{d^2 y}{c^2 dt^2} \approx -\frac{g}{c^2}
\]

As emphasized by Rindler [1], this is a remarkable result which has far reaching implications. It states that since light travels at a finite speed, it has “weight”! No other assumptions about light need be invoked in this argument. Light, and all other physical phenomena, must travel on locally curved paths in gravity. This point strongly motivates the idea that gravity is an aspect of geometry and does not belong in the menagerie of forces! The argument above just gives the local bending of light in a uniform gravitational field. In order to compare to experiments we need to consider the path light traverses past a stellar object like the sun. The full machinery of general relativity is required here, as we illustrated in the textbook. Historically the deflection of light by the gravitational field of the Sun was first computed using Newtonian mechanics and a “corpuscular” model of light in the early days of the nineteenth century. Einstein repeated the calculation in the context of relativistic, curved space-time in 1919 and found that the Newtonian prediction is too small by a factor of two.

The experimental observation of the deflection of light in a gravitational field comes from a lensing effect, as illustrated in Figure 1.3.
Fig 1.3  Light rays traveling past a massive star and being “focused”. A lensing effect.

When light passes by a large astronomical object, it is “attracted”, as implied by the Equivalence Principle, and it is deflected as shown in the figure. The effect was observed originally by carefully measuring the background stars around the halo of the Sun during an eclipse and comparing those measurements to the position of the background stars when the Sun is in a different part of the sky. These measurements are difficult but even the earliest measurements by Sir Arthur Eddington in 1919 favored Einstein over Newton.

Finally, let’s reconsider the scattering of light in the presence of a large star somewhat more critically. Because the direction of the gravitational acceleration varies as we move around the star, the use of the Equivalence Principle must be stated more carefully. As we pass around the Sun, the Equivalence Principle can be applied only \textit{locally}; that is, only over a region of space-time where the gravitational field is essentially uniform can we find a freely falling inertial frame in which the field is essentially eliminated. The theory speaks of “local inertial frames” to accommodate spatially varying gravitational fields. We certainly cannot eliminate the effects of
gravity from large space-time regions! The spatial dependence of gravitational fields means that the mathematical details of the theory change from point to point. For example, a body falling in a spatially varying gravitational field executes straight-line motion in each local inertial frame approximating the varying gravitational field. The actual trajectory of the body is obtained by patching together its simple trajectories in contiguous inertial frames. This sounds awfully complicated. Mathematically, the language for this motion is neatly given by differential geometry.

2. **Motion in a Rotating, Relativistic Reference Frame.**

In special relativity we plot the world line of a particle's motion on a Minkowski diagram (Figure 2.1).

![Fig. 2.1 World line of massive particle traveling into its future light cone.](image)
In general relativity we will use notions such as the invariant interval and metric more than we did in our discussions of special relativity. So, let’s review these ideas. A problem familiar from elementary mechanics, such as the rotating reference frame, will help us move gradually into new, uncharted subjects.

Recall that the proper time $d\tau$ that passes on a clock attached to the moving particle can be calculated from the invariant interval $ds$,

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2.$$  \hspace{1cm} (2.1)

Consider the meaning of the symbols in this formula. In a frame $S$, we imagine two events, one at $(ct, x, y, z)$ and another at $(ct + cdt, x + dx, y + dy, z + dz)$. $ds$ is then the invariant interval between them. The events might be ticks on a clock, measurements of the ends of a rod, or whatever. The crucial point is that $ds$ is the same in all reference frames, so if we can calculate and understand it in one frame, we have its value in all frames. The proper time $d\tau$ that passes on a clock attached to the moving particle can be computed from $ds$ by boosting to a frame $S'$ where the particle is at rest. In $S'$, $dt' = d\tau$ (proper time), $dx' = dy' = dz' = 0$, so $ds^2 = c^2 d\tau^2$. So, if the particle is moving along the $x$ axis at velocity $v$, then $dx = v \cdot dt$, $dy = 0$, and $dz = 0$, and Eq. (2.1) reduces to

$$ds^2 = c^2 d\tau^2 = c^2 dt^2 - v^2 dt^2 = (c^2 - v^2)dt^2$$ \hspace{1cm} (2.2a)

or

$$d\tau = \sqrt{1 - v^2 / c^2} dt \equiv dt / \gamma,$$ \hspace{1cm} (2.2b)

which is just the expression of time dilation—the moving clock runs more slowly.

Minkowski diagrams, space-time pictures, are the natural arena for discussing dynamics because they show time and position information together. Because space and time mix under boosts, we must work in four-dimensional space-time. If a free particle's world line passes through $P_1 = (ct_1, x_1)$ and $P_2 = (ct_2, x_1)$, we know its velocity $v = (x_2 - x_1) / (t_2 - t_1)$ and its path.
As a first step toward developing relativistic particle motion in a gravitational field, consider relativistic force-free motion in a rotating reference frame. In rotating frames, there are centripetal and Coriolis “apparent” forces. Choose a cylindrical spatial coordinate system, 
\[ x = r \cos \varphi, \quad y = r \sin \varphi, \quad z = z \]
as shown in Figure 2.2.

![Cylindrical coordinate system](image)

Fig. 2.2 Cylindrical coordinate system.

Then the spatial distance becomes
\[ dx^2 + dy^2 + dz^2 = dr^2 + r^2 d\varphi^2 + dz^2 \]
and the space-time invariant interval reads
\[ ds^2 = c^2 dt^2 - (dr^2 + r^2 d\varphi^2 + dz^2) \]

To describe a turntable that is rotating about the z axis at angular velocity \( \omega \), as in Figure 2.3, we introduce a new azimuthal angle \( \varphi' \),
\[ \phi' = \phi - \omega t, \] (2.3)

so a point with fixed \( \phi' \) has its \( \phi \) increasing as \( \omega t \). This simple equation mixes the time coordinate with a spatial coordinate, so

\[ d\phi' = d\varphi - \omega dt \]

and the invariant interval written in terms of the rotating coordinates becomes

![Coordinates of rotating coordinate frame.](image)

\[ ds^2 = (c^2 - \omega^2 r^2)dt^2 - (dr^2 + r^2 d\phi'^2 + 2\omega r^2 d\phi' dt + dz^2) \] (2.4)

and does not neatly separate into a spatial and temporal part. Note that the coordinate-time mixing term, \( 2\omega r^2 d\phi' dt' \), is proportional to \( \omega \). We will see analogous terms when we study rotating Black Holes in later lectures.
Before dealing with $ds^2$ in its full glory, consider a clock at rest in the rotating reference frame and a distance $r$ from the $z$ axis. Two ticks of the clock occur at a given $r, \varphi'$, and $z$, so $dr = d\varphi' = dz = 0$, and Eq. (2.4) reduces to

$$ds^2 = c^2 d\tau^2 = (c^2 - \omega^2 r^2)dt^2.$$ 

So, the proper time kept by the clock is

$$d\tau = \sqrt{1 - \omega^2 r^2 / c^2} dt.$$  \hspace{1cm} \text{(2.5)}

This result is the usual time dilation formula because the velocity of the clock relative to the inertial $x, y, z$ frame is $v = \omega r$: as $r$ increases, the clock's velocity (transverse) increases proportionally and the time dilation effect is enhanced. We must restrict the possible values of $\omega$ and $r$ so that $\omega r < c$ in order that expressions such as Eq. (2.5) remain physical.

Now consider the cross term, $-2\omega r^2 d\varphi' dt$, which mixes time and space intervals. To appreciate what is going on, imagine setting up a cylindrically symmetrical gridwork of rods and synchronized clocks to make position and time measurements in the rotating reference frame. First, consider two clocks having the same $r$ and $\varphi'$ but different $z$ values, as shown in Figure 2.4.
Fig. 2.4 Synchronizing clocks at different $z$ coordinates in a rotating coordinate frame.

The clocks are rotating about the $z$ axis at an angular velocity $\omega$. We place a signal generator halfway between them, at rest in the rotating frame, to synchronize them. Because clocks 1 and 2 have velocities transverse to $z$, they receive the light pulses simultaneously in both the fixed $(x, y, z)$ inertial frame and the rotating frame. The synchronization procedure produces clocks that are synchronized in both frames.

However, clocks at the same $r$ and $z$ but different $\varphi'$ suffer a different fate, as illustrated in Figure 2.5, where the view is from above.
We have placed our signal generator halfway between the two clocks at their common $r$ and $z$ values. Note that the signal generator and the two clocks are each moving in separate directions and each is experiencing a different centrifugal acceleration. However, if we move clocks 1 and 2 nearby, so they are separated by just an infinitesimal distance, we have a more familiar situation. In this case, all three objects have essentially the same velocity and they can be viewed from a locally inertial frame moving at their common velocity. Because clocks 1 and 2 are separated in their direction of motion, we know from our discussion of the relativity of simultaneity that they are not synchronized in the inertial frame defined by the coordinates $(ct, x, y, z)$. The reason for this is clear—since light travels at the speed limit $c$ with respect to any inertial observer, an observer at rest in the $(ct, x, y, z)$ frame notes that clock 2 receives the light signal from the signal generator after clock 1 does because clock 2 is racing away from the source and clock 1 is
racing toward the source. So, events that are synchronous in the rotating frame (clocks 1 and 2, for example) are not synchronous in the fixed inertial frame \((ct, x, y, z)\) if they occur at different \(\phi\) values. In fact, we know how large this effect is from our discussion of the relativity of simultaneity—it is the product of the velocity of the moving clocks times the distance between them in their rest frame divided by \(c^2\). The relevant velocity is \(\omega r\). The relevant distance in the inertial frame \((ct, x, y, z)\) is \(rd\phi\), which corresponds to a larger distance \(\gamma(r)rd\phi\) in the clock's rest frame. Here \(\gamma(r) = (1 - \omega^2r^2/c^2)^{-1/2}\), so we have explicitly written \(\gamma\) as a function of \(r\). (The notion of a local inertial frame is certainly important here.) So, \((\omega r)[\gamma(r)rd\phi]/c^2\) is the time discrepancy that the frame \((ct, x, y, z)\) notes on the clocks. But this is a time interval in the rotating frame, and we want the time difference in the fixed frame. This is given by multiplying by another factor of \(\gamma(r)\) to account for time dilation, so the time difference is

\[
\frac{\gamma^2(r)\omega r^2d\phi'}{c^2} = \frac{\omega r^2d\phi'}{c^2 - \omega^2r^2},
\]

But if the two clocks are synchronized in the rotating frame, then the time that passes there, \(dt'\), must be related to \(dt\), the time that passes in the inertial frame, by

\[
dt' = dt - \frac{\omega r^3}{c^2 - \omega^2r^2}d\phi'.
\]

It must be that if we use this \(t'\) axis, the invariant interval will split into a spatial part and a temporal part. Substituting \(dt = dt' + \omega r^2d\phi'(c^2 - \omega^2r^2)\) into Eq. (2.3), we find

\[
ds^2 = (c^2 - \omega^2r^2)dt'^2 - \left(dr^2 + \frac{c^2r^2d\phi'^2}{c^2 - \omega^2r^2} + dz^2\right)
\]

and everything has worked out fine: the coordinates are orthogonal with a clear distinction between time-like and space-like intervals.
When we develop more mathematical machinery for Riemannian space-time, we will analyse metrics like Eq. 2.6 and solve problems in relativistic particle dynamics. Note, however, that Eq. 2.6 is just a re-parametrization of flat Minkowski space-time for $r < c/\omega$, so its space-time curvature vanishes, etc. The parametrization of Minkowski space-time with the primed coordinates has a serious flaw: it has created coordinate singularities in the metric. We will find analogous problems in some metrics of black holes later in the course and they will point to new and bizarre effects! We will refer back to Eq. 2.4 and Eq. 2.6 when we study the Kerr metric which describes rotating stars and black holes.

3. Tidal Forces, non-Euclidean Geometry and “Local Inertial Reference Frames”

There are aspects of the Equivalence Principle which are quite subtle. The claim is that freely falling reference frames are locally inertial frames where the rules of special relativity hold. The key word is “local”. What does that mean quantitatively? Let’s answer the question with a thought experiment.

Suppose I am in the freely falling reference frame with four small balls. I am outside a planet of huge mass $M$ with my four toys and I orient myself so that my feet are nearest the planet and I say that the planet is in the $-z$ direction. Newtonian mechanics will be adequate for the measurements and experiments we are about to describe. One of the balls is placed at my feet, one at the top of my head, another on my right side and the last one at my left side. They are all freely floating. If the reference frame were perfectly inertial, the four balls would retain their initial positions with respect to me at all times. What do I observe? The ball on my head accelerates away to larger $z$, the ball at my feet accelerates away to smaller $z$ and the balls at my right and left side
accelerate towards me at half the rate of the first two balls! This surprises me. In fact, I feel rather poorly...I feel stretched in the z direction and squeezed in the transverse directions.

What is happening? Apparently the size of my body exceeds that of a local inertial frame centered at my midsection and I feel the effects of gravity at my extremities. To describe the motions of the four balls, set up a Euclidean Cartesian coordinate system, \( x^i, i = 1,2,3 \), with its origin at the center of the planet. From Newton’s perspective the freely falling frame is accelerating toward the mass \( M \). Call the acceleration of gravity at my midsection \( \mathbf{g} \). It points in the \(-z\) direction.

Newton would describe the motions of my body at position \( \mathbf{r}(t) \) with the second law,

\[
\frac{d^2\mathbf{r}}{dt^2} = -\nabla \Phi(\mathbf{r})
\]  

(3.1)

where \( \Phi(\mathbf{r}) = -\frac{GM}{r} \) is Newton’s gravitational potential and \( r = \sqrt{x^2 + y^2 + z^2} \). Label the location of each ball with the vector \( \mathbf{r} + \mathbf{\epsilon} \), where \( |\mathbf{\epsilon}| \ll r \). The equation of motion of each ball then reads,

\[
\frac{d^2}{dt^2} (\mathbf{r} + \mathbf{\epsilon}) = -\nabla \Phi(\mathbf{r} + \mathbf{\epsilon})
\]  

(3.2)

Since \( \mathbf{\epsilon} \) is small we can expand the gravitational potential around the point \( \mathbf{r} \),

\[
\Phi(\mathbf{r} + \mathbf{\epsilon}) = \Phi(\mathbf{r}) + \mathbf{\epsilon} \cdot \nabla \Phi(\mathbf{r}) + \cdots
\]

Finally we can take the difference of the two equations of motion Eq. (3.1) and (3.2) to isolate the equation of motion for each ball to first order in \( \mathbf{\epsilon} \),

\[
\frac{d^2\mathbf{\epsilon}}{dt^2} = -\nabla (\mathbf{\epsilon} \cdot \nabla \Phi(\mathbf{r}))
\]  

(3.3)

We learn that the balls accelerate with respect to me by an amount that varies as the second derivatives of the potential. We can obtain an explicit formula by working out the derivatives,

\[
\nabla \Phi(\mathbf{r}) = \frac{GM}{r^3} \hat{r}
\]
\[ \nabla \left( \frac{GM}{r^3} \mathbf{e} \cdot \mathbf{r} \right) = -3 \frac{GM}{r^4} (\mathbf{e} \cdot \mathbf{r}) \hat{r} + \frac{GM}{r^3} \mathbf{e} \cdot \hat{r} \]

So, finally, in Cartesian components,

\[ \frac{d^2 \epsilon^i}{dt^2} = \frac{GM}{r^3} \sum_k \epsilon^k \left( 3r^k r^i - r^2 \delta^{ki} \right) \quad (3.4) \]

where \( r^i = (x, y, z) \). We see that the right hand side of this equation has the spatial distribution of a “Quadrupole moment”,

\[ Q^{(ij)} = (3r^i r^j - r^2 \delta^{ij}) \]

We shall see more Quadrupole moments in other gravitational phenomena below.

Now we can get our answers! Choose \( x = 0, y = 0 \) and \( z = r \). Then, upon substitution, into the equation of motion Eq. (3.4), we have,

\[ \frac{d^2 \epsilon_1}{dt^2} = - \frac{GM}{r^3} \epsilon_1, \quad \frac{d^2 \epsilon_2}{dt^2} = - \frac{GM}{r^3} \epsilon_2, \quad \frac{d^2 \epsilon_3}{dt^2} = + \frac{2GM}{r^3} \epsilon_3 \]

And we read off the stretching and squeezing alluded to above: there is elongation in the z-direction and squeezing in the transverse directions.

This exercise illustrates several points.

First, we have rediscovered gravitational tidal forces. Recall from your mechanics course that the moon creates tidal forces on the earth and these forces are responsible for the daily high and low tides in a harbour such as London or New York. As the earth rotates on its axis under the moon, there are two high tides and two low tides each solar day, approximately. This is because of the elongation we discovered in \( \epsilon^{(3)} \) and squeezing we discovered in the transverse directions \( \epsilon^{(1)} \) and \( \epsilon^{(2)} \). We visualize this effect in Figure 3.1.
In part a. of the figure we show the forces exerted on the earth by the moon, and in part b. we *subtract* the force exerted on the *center* of the earth so we concentrate on the “differential” force across the earth. Here we see the pattern of forces derived in Eq. (3.4) that are described by the Quadrupole moment: there is elongation in the direction toward the moon and squeezing in the transverse directions. We also see that the rotational frequency of the quadrupole moment is *twice* that of the earth: there are two high tides and two low tides in a twenty-four hour period, approximately.

We will have more to say about this later when we develop the equations of general relativity in later lectures.
4. Gravitational Red Shift

The gravitational red-shift represents one of the simplest and decisive tests of the Equivalence Principle. We are interested in how light propagates in a gravitational field. Suppose that light of frequency $v_o$ is emitted upward from the surface of a static, non-rotating planet to be observed some distance $h$ away. What frequency $v_o$ does the observer measure? We know that light travels at the speed limit $c$ in any inertial frame and we know that a freely falling frame is inertial. So, consider first the propagation of the light signal from the surface of the planet from the perspective of an inertial frame falling freely in the approximately uniform gravitational field $g$ as shown in Figure 4.1.

Fig. 4.1 Pulse of light traveling from emitter E to observer O, as viewed from a falling inertial frame.
The light pulse travels a distance \( h \) in the time interval from \( t = 0 \) to \( t = h/c \), and the falling frame has a velocity \( v = gh/c \) downward relative to the observers at rest on the planet when the light pulse reaches its destination. The local falling frame is inertial and the light pulse has frequency \( \nu_e \) when it is emitted and when it is absorbed, as measured in the falling frame. But at \( t = h/c \), the inertial frame is moving downward, opposite to the direction of propagation of the light pulse, so an observer at rest on the planet detects a Doppler shift toward the red when she observes the light pulse at \( t = h/c \). The frequency shift is, using the longitudinal Doppler shift formula derived in the lecture series on Special Relativity,

\[
\nu_o = \sqrt{\frac{1 - v/c}{1 + v/c}} \nu_e \approx \sqrt{1 - 2v/c} \nu_e \approx \left(1 - \frac{gh}{c^2}\right) \nu_e,
\]

where we substituted the velocity \( v = gh/c \), and linearized the formula for our present application, where \( v/c \ll 1 \).

The fractional change of the frequency is

\[
\frac{\nu_o - \nu_e}{\nu_e} = \frac{\Delta \nu}{\nu} = -\frac{hg}{c^2}.
\]

(4.1)

So, the observer a height \( h \) above the surface of the planet observes a red-shifted wave.

Another way of presenting this result, which is more fundamental, is to say: \emph{Identically constructed clocks run slower in lower gravitational potentials}. Because frequency varies as the reciprocal of time, \( \nu = 1/t \), we can express Eq. (4.1) as a fractional time difference,

\[
\frac{\Delta t}{t} = \frac{hg}{c^2}.
\]

(4.2)

In other words, an observer E at height 0 could send light signals to observer O at height \( h \) once a second. Eq. (4.1) then states that observer O detects these signals more widely spaced in time, at a
diminished or red-shifted frequency. So, observer O concludes that the clock at height 0 is running more slowly than hers, according to Eq. (4.2).

Just as the observer at height 0 sent light signals to an observer at height $h$ to compare clocks, their roles could be interchanged. Then the same argument, modified to account for the fact that now the freely falling frame is accelerating in the same direction as the light ray, so the Doppler shift formula applies with $v$ replaced by $-v$, predicts that the lower observer detects a blue-shifted light ray. The lower observer concludes that the clock in the higher gravitational potential runs faster than his by a fractional change of $-hg/c^2$.

Another way to use the Equivalence Principle to analyze the gravitational red-shift is to replace the problem in a gravitational field with one in an accelerating reference frame (e.g., a spaceship), as shown in Figure 4.2.

![Diagram](image)

Fig. 4.2 A light ray traveling up from a planet is equivalent to a light ray traveling with respect to an accelerating rocket in an inertial frame.
The spaceship of height $h$ is in a background inertial frame, so light travels at velocity $c$ there regardless of its origins. Light leaves the base of the ship (E), travels up during a time $t = h/c$, and is received at the tip of the ship (O). But O is receding at velocity $v = at = gh/c$, so it detects the frequency of light

$$\nu' = \sqrt{\frac{1 - \frac{v}{c}}{1 + \frac{v}{c}}} \nu \approx \sqrt{1 - 2\frac{v}{c}v} \approx (1 - \frac{gh}{c^2})\nu$$

in agreement with our earlier argument.

It is interesting that the best experimental determination of gravitational red-shift comes from terrestrial, controlled experiments that rely on the Mössbauer effect to measure tiny frequency shifts rather than from observations of distant stars. Recall the Mössbauer effect discussed in the lecture series on Special Relativity. The experiment involved an emitting atom within a regular crystal array. The emitting atom shares its recoil energy and momentum with its entire crystal environment, so its recoil velocity is reduced to essentially zero and all the energy between the quantum energy levels in the atom shows up as the energy of the emitted light wave. If the light emitted from such a crystal is incident on another identical crystal, absorption is possible because the energy of the light exactly matches an energy difference in the spectrum of the receiving atom, which can absorb the light without recoiling. If either the absorbing crystal or emitting crystal has a small relative velocity, the attendant Doppler shift shifts the energy of the light enough to make resonant absorption impossible. (Recall that in quantum physics, the frequency of light is related to its energy by the Planck relation $E = hv$, where $h$ is Planck's constant, $h = 6.627 \cdot 10^{-34}$ J-s, so we can speak equally well in terms of frequency $\nu$ or energy when discussing emission and absorption of photons by atoms.) In fact, the energy levels of atoms are not infinitely precise—each energy level has a natural width that is a consequence of the Heisenberg uncertainty relation.
Mössbauer was then able to study the profiles of spectral lines, obtaining their widths and detailed shapes by using his crystals and the physics of the Doppler shift.

R. V. Pound, G. A. Rebka, and J. L. Snider used the Mössbauer effect to measure the redshift of light near the surface of Earth [2]. They placed a source of $^{57}$Co 22.6 m higher than a receiver. The Equivalence Principle predicts a fractional frequency shift of $\Delta \nu / \nu_e \approx 2.46 \cdot 10^{-15}$, which is very tiny: three orders of magnitude smaller than the intrinsic natural uncertainty in the frequency of the emitted light $\Delta \nu_{QM}, \Delta \nu_{QM} / \nu \approx 1.13 \cdot 10^{-12}$. In order to obtain an observable effect, they gave the source a sinusoidal velocity $u = u_o \cos \omega t$ to induce a controlled, mechanical Doppler shift, $\Delta \nu_D / \nu = - (u_o / c) \cos \omega t$. Then the resonance-absorption cross section depends on $\Delta \nu + \Delta \nu_D$, and if $\Delta \nu_D >> \Delta \nu$, an exercise in quantum mechanics shows that the cross section has a term linear in $\cos \omega t$ with a strength proportional to $\Delta \nu$, the gravitational red-shift. By isolating this characteristic external frequency $\omega$ in the cross-sectional measurements, the experimenters confirmed the prediction of general relativity to good precision, about 10%. It is interesting that this terrestrial experiment is much more decisive than astronomical data of light emitted from distant stars because of all the inherent uncertainties in such observations (the Doppler effects of the movement of the star and the movement of the emitting atom in the turbulent hot surface of the star must be accounted for somehow).

We end this discussion with an illustration of the gravitational red-shift that derives it from energy conservation. We show that the gravitational red-shift is necessary for the internal consistency of relativistic dynamics—without it we could construct a perpetual motion machine!

Consider a pulley on the surface of Earth supporting two observers at the ends of a string as shown in Figure 4.3 [3].
Let the difference of heights of the observers be \( h \), as usual. Let the lower observer shine a flashlight at the higher observer, who sees the light, so it is absorbed on his retina. (Notice the similarity of this argument to our original derivation of \( E = mc^2 \), except the apparatus is now vertical and in an ordinary gravitational field.) Let the light have energy \( E \), so from special relativity we know that it has a mass equivalent of \( E/c^2 \). So, if the light travels from the lower to the higher observer, then mass \( E/c^2 \) has been transferred between the observers; the upper one is now slightly heavier than the lower one, so it sinks and does work equal to the change in the potential energy, \( hg \cdot E/c^2 = E \cdot \Delta V/c^2 \). Our pulley system is now back to its original configuration, and we can repeat the process as many times as we wish and have an inexhaustible source of work.

We have done it! We have made a perpetual motion machine! Well, except for one thing—we forgot about the gravitational red-shift! The light detected by the higher observer in Figure 4.3 has a smaller frequency, given by our gravitational red-shift formula, \( \Delta v/v = -\Delta V/c^2 \). How can this get us out of our conundrum? It is easy to see (as we verify below) that the energy that light
carries transforms between frames exactly as its frequency. If that is the case, then the energy the light deposits on the retina of the higher observer is diminished by \( E \cdot \Delta V/c^2 \), which is exactly the work we were hoping to get out of our machine! So, in reality, when the light from the lower observer reaches the higher one, its energy is diminished by the gravitational red-shift by just the amount we wanted to generate. So, yet another perpetual motion machine design bites the dust!

Let us check that light's frequency and energy transform identically in special relativity. Recall that if we know the energy \( E \) and the \( x \) component of the relativistic momentum \( p_1 \) in frame \( S \), then the energy in frame \( S' \) is \( E' = \gamma (E - vp_1) \). But for light propagating in the \( x \) direction, \( p_1 = E/c \), so the transformation law becomes

\[
E' = \gamma \left( E - \frac{\nu E}{c} \right) = \gamma \left( 1 - \frac{\nu}{c} \right) E = \frac{1 - \nu/c}{1 + \nu/c} E,
\]

which we recognize as the Doppler shift formula for the light's frequency. And so the proposed perpetual motion machine fails miserably.

The fact that light's frequency \( \nu \) and energy \( E \) transform identically under boosts is important in the quantum theory of light we have touched on in our discussion of relativistic collisions and quantum energy levels. Planck's equation, \( E = h\nu \) (where \( h \) is Planck's constant), is consistent with special relativity because \( E \) and \( \nu \) share the same transformation law.

This observation suggests yet another way of viewing the gravitational red-shift—it is just a consequence of energy conservation! When a photon, a quantum of light energy, travels from the observer on the surface of a planet where its frequency is \( \nu_o \) and its height is 0, to an observer where its frequency is \( \nu_o \) and its height is \( h \), the total energy, accounting for the gravitational potential, must be conserved:

\[
E_c = E_o + hg \frac{E_o}{c^2}.
\]
The change in potential energy has been written as \( hg(E_0/c^2) \), which is the change in the potential \( hg \) in the uniform gravitational field times the mass equivalent of the energy \( E_0 \) there. Solving for \( E_0 \),

\[
\frac{E_0}{E_e} = \frac{1}{1 + hg/c^2} \approx 1 - \frac{hg}{c^2},
\]

which we can write as a fractional change in energy, which is also the fractional change in frequency,

\[
\frac{E_0 - E_e}{E_e} = \frac{\nu_o - \nu_e}{\nu_e} \approx -\frac{hg}{c^2},
\]

and we have derived Eq. (4.1) again from a different perspective!

This result gives us a nice alternative view of gravitational red-shift. Why does the frequency of light change as it propagates away from the surface of a celestial body? Because it propagates to a new location where its gravitational potential is larger, so its relativistic energy, and hence its frequency, must be diminished accordingly!

In these applications we proceeded approximately and discussed only weak gravitational fields. Once we develop General Relativity, we will obtain the redshift formula exactly.

Before we go on to other topics, let’s assimilate what we have learned about the gravitational redshift into the language of General Relativity. As discussed earlier, the language of General Relativity is that of classical differential geometry and Riemannian geometry. We speak about the invariant interval in space time and space time curvature.

The metric of Minkowski space time was developed in the lectures on Special Relativity,

\[
ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = \sum_{\sigma\rho} \eta_{\sigma\rho} dx^\sigma dx^\rho
\]


where $\eta_{\sigma\rho}$ is the Minkowski metric of Special Relativity, a second rank covariant tensor whose non-zero components are $\eta_{00} = 1, \eta_{11} = \eta_{22} = \eta_{33} = -1$ in Cartesian coordinates. We used this formalism to discuss time dilation and found, in Eq. 2.2,

$$ds^2 = c^2dt^2 = (1 - v^2/c^2)c^2dt^2$$

Now we want to write the metric in more general terms in order to confront problems in gravity. We write,

$$ds^2 = \sum_{\sigma\rho} g_{\sigma\rho}dx^\sigma dx^\rho$$

and among our tasks in General Relativity will be the determination of the metric tensor $g_{\sigma\rho}$.

Let’s incorporate what we have learned about gravitational redshift into this language. We saw that the frequency difference $\Delta \nu$ is related to the potential difference $\Delta V$ at the space time points where the frequencies are observed,

$$\frac{\Delta \nu}{\nu} \approx -\frac{\Delta V}{c^2}$$

Since frequencies vary as the reciprocal of time intervals, we also had,

$$\frac{\Delta t}{t} \approx \frac{\Delta V}{c^2}$$

So, time measurements depend on the gravitational potential. So, the invariant interval and the metric must depend on $V(r)$. We write,

$$ds^2 = \sum_{\sigma\rho} g_{\sigma\rho}dx^\sigma dx^\rho = g_{00}(r)c^2dt^2 + \cdots$$

To determine how $g_{00}$ depends on $V(r)$, consider two pulses of light emitted at $r_E$ and observed at $r_O$ in a static gravitational potential $V(r)$. The time interval between the pulses at $r_E$ is the same as the time interval between the reception of the pulses at $r_O$ because the time of transit
from $r_E$ to $r_O$ is time independent. Call the time interval $dt$. Then the proper time differences at $r_E$ and $r_O$ are,

$$(\Delta \tau_E)^2 = g_{00}(r_E) \Delta t^2$$  \hspace{1cm} \text{and} \hspace{1cm} (\Delta \tau_O)^2 = g_{00}(r_O) \Delta t^2$$

Therefore,

$$\frac{\Delta \tau_E}{\Delta \tau_O} = \sqrt{\frac{g_{00}(r_E)}{g_{00}(r_O)}}$$  \hspace{1cm} (4.3)$$

But our redshift example showed that,

$$\frac{\Delta \tau_E}{\Delta \tau_O} \approx \frac{1 + V(r_E)/c^2}{1 + V(r_O)/c^2}$$

So, to first order in $V(r)/c^2 \ll 1$,

$$g_{00}(r) \approx 1 + 2V(r)/c^2$$

where the factor of 2 comes from expanding the square root in Eq. 4.3. This will prove to be an important result.

We can also generalize this result to strong gravity. We had,

$$\frac{dv}{v} = -\frac{dV}{c^2}$$

Integrating,

$$\frac{v_B}{v_A} = \frac{e^{-V_B/c^2}}{e^{-V_A/c^2}} = e^{-(V_B-V_A)/c^2}$$

By the line of argument above, this implies,

$$ds^2 = e^{2V/c^2}c^2dt^2 - dl \cdot dl$$

where $dl \cdot dl$ indicates the spatial components of the metric. So,

$$g_{00}(r) = e^{2V(r)/c^2}$$

where $V(r)$ is the Newtonian gravitational potential.
This result will be important when we consider other applications, such as accelerating reference frames. Of course, the more we can infer about $dl \cdot dl$, the more physics we can do. More later.

5. The Twins Again

In our lectures on Special Relativity, we left our twins, Mary and Maria, when they were coping with the fact that Mary was four years older than Maria after Maria’s space trip. In that discussion, we found that from Maria's perspective Mary ages an unexpectedly large amount when Maria jumps from the outgoing to the incoming rocket at the midpoint of her trip. Recall from that discussion that the lines of constant time in the frame of the outgoing rocket are at widely different angles from the lines of constant time in the frame of the incoming rocket, as shown in the Minkowski diagram, Figure 5.1 (Figure 3.19, of the textbook).
Fig. 5.1 The world paths of Mary (vertical line) and Maria (the outgoing and incoming straight lines). According to Maria’s clock, Mary ages 6.4 years in the “split second” when Maria jumps from the outgoing to the incoming rocket.

In fact, we read off the figure that Maria measures that 6.4 years passes on Mary's clock during Maria's turnaround!

By using the device of two rockets, we have been able to analyze the twin paradox without explicitly considering space and time measurements in an accelerated reference frame. Now that we know that accelerated reference frames are equivalent to environments in a gravitational field, we can face the problem head on. Our only limitation is that our discussion of general relativity so far is good only for weak gravitational fields, or, equivalently, small values of $v/c$. We will do better when we present the full apparatus of General Relativity, Einstein's field equations, in later lectures. That work will justify this application where large gravitational effects occur.

View the trip from Maria's frame. From Maria's perspective, Mary goes out and back. By the Equivalence Principle, then, Maria's acceleration can be replaced by a gravitational field. We already know that the important portion of the trip is the period during which Maria detects Mary's reversal. Call the distance between the sisters $vT$ at this point, where $T$ is the time they have been traveling apart at velocity $v$, according to Mary. At this point, the turnaround, Maria actually turns on her rocket motors and experiences an acceleration $a$ for a time $t$. However, invoking the Equivalence Principle, we can replace the rocket environment by one in which there is a gravitational field in Maria's frame that produces an acceleration $a$, as shown in Figure 5.2.
The gravitational potential difference between the sisters is distance times acceleration, \( avT \), which causes Mary's clock (her heartbeat) to gain the total amount \( (avT/c^2)t \), according to our formulas developed for the gravitational red-shift. Finally, \( t \) must be long enough to accommodate the reversal from the outgoing speed \( v \) to the incoming speed \( -v \), so \( at = 2v \). Therefore, the total time gained by Mary's clock is \( 2Tv^2/c^2 \).

Does this result agree with our previous analysis of the Twin Paradox? Return to Figure 5.1 and compute the time that Maria states passes on Mary's clock during the turnaround. We pull out our trusty relativity of simultaneity formula, which reminds us that Maria measures clocks at rest in Mary's frame that are separated by a distance \( x \) to be out of synchronization by an amount \( xv/c^2 \). So, if we change from the outgoing to the incoming rocket, causing \( v \), to change by \( \Delta v \), the time that passes on Mary's clock is \( x\Delta v/c^2 \). But \( x \) is the distance between the sisters in Mary's frame, which is \( vT \) in our discussion here. And finally, because the velocity \( v \) reverses, \( \Delta v = 2v \). So, our
earlier relativity of simultaneity discussion predicted that Mary's clock, as measured by Maria, gains a time of $2Tv^2/c^2$, in complete agreement with the general relativity result obtained here!

Using the numbers of our previous discussion in, $T = 5$ and $v/c = 0.8$, so Mary's time gain is $2 \cdot 5 \cdot 0.8^2$, which is 6.4 years, as we found earlier.

6. Similarities and Differences of Electromagnetism and Gravity

It is interesting to consider the similarities between the two field theories considered in this book, electromagnetism and general relativity. We will also return to this subject after we have developed general relativity for strong gravity in a later lecture.

Our investigations in electromagnetism started with Coulomb’s Law, the force between two charges $q_1$ and $q_2$ a distance $r$ apart,

$$ F(r) = k \frac{q_1 q_2}{r^2} \hat{r} $$

and our investigations in gravity started with Newton’s Law of the gravitational force between two masses $m_1$ and $m_2$ a distance $r$ apart,

$$ F_g(r) = -G \frac{m_1 m_2}{r^2} \hat{r} $$

We rewrote Coulomb’s Law as a local differential equation on our way to developing the field theory of electromagnetism,

$$ \nabla \cdot E(r) = -\nabla^2 V(r) = 4\pi k \rho(r) $$

where $\rho$ is the charge per unit volume, charge density, at the position $r$ and $V(r)$ is the electrostatic potential, $E = -\nabla V(r)$. 

36
In the case of gravity we introduce the potential \( V_g(\mathbf{r}) \) due to the presence of particle \( m_2 \) at the origin and write \( F_g(\mathbf{r}) = -m_1 \nabla V_g(\mathbf{r}) \). Then the same steps that led from Eq. (6.1) to (6.3) produce the differential equation,

\[
\nabla^2 V_g(\mathbf{r}) = 4\pi G \rho_g(\mathbf{r}) \tag{6.4}
\]

where the source of the gravitational field is the mass density \( \rho_g(\mathbf{r}) \).

Both forces \( \mathbf{F}(\mathbf{r}) \) and \( F_g(\mathbf{r}) \) are long range, falling off as \( r^{-2} \). In the case of electromagnetism we found that the long-range character of the electrostatic force led to the fact that the \textit{dynamical} electromagnetic field satisfies the wave equation,

\[
\left( \nabla^2 - \frac{\partial^2}{c^2 \partial t^2} \right) \mathbf{E}(\mathbf{r}, t) = \left( \nabla^2 - \frac{\partial^2}{c^2 \partial t^2} \right) \mathbf{B}(\mathbf{r}, t) = 0
\]

in free space. This predicted that electromagnetic radiation exists, that electromagnetic waves travel at the speed limit and that the photon has a vanishing rest mass. This development suggests that the gravitational potential will be replaced by the gravitational field when we pass from Newton’s to general relativistic space-time. We have already seen that Newton’s gravitational potential enters general relativity through deviations in the space-time metric \( g_{\mu\nu} \) from the flat Minkowski metric of special relativity. So, we expect that fluctuations in the gravitational field will manifest themselves in general relativity as fluctuations in the metric and they will propagate at the speed limit and will satisfy the wave equation in Minkowski space-time if they are small. The associated quantum, the “graviton”, will be massless just like the photon. These educated “guesses” are, in fact, true, but their verification requires investigations into general relativity, the subject of later lectures.

The electrodynamic force law states that like charges repel and unlike charges attract. For example, electrons attract their anti-particle relative, the positron, but electrons repel each other.
This is a characteristic of the spin one nature of the photon, as you will learn in field theory. However, Newton’s Law of gravity states that masses attract one another and only positive masses are known to exist. The universal attraction of masses, actually energies in the relativistic version of Newton’s Law, the Einstein field equations, is an essential element of the Equivalence Principle. The equality of inertial and gravitational masses is the basis of interpreting gravity as an aspect of geometry. These ideas generalize to field theories because field theory predicts that the masses of particles and their associated anti-particles are identical: particle and anti-particles must have opposite charges but they must have identical positive inertias.

We have seen in our studies of magnetism that the screening characteristic of electrodynamics is essential to its phenomenology. For example, the discovery and relevance of magnetism in practical situations relies on the fact that in most situations matter is effectively neutral and static electric forces are absent, leaving us able to find magnetic effects even though they are typically much smaller effects, suppressed by several powers of $v/c$. In these cases the positive charges of nuclei are screened by the negative charges of mobile electrons. There is no such screening in gravity so its various velocity dependent forces are often overwhelmed by static gravitational attractions between slowly moving objects.

At the quantum level, the spin one nature of the photon explains why like charges repel and unlike charges attract. What does the universal nature of gravity’s attraction tell us at the quantum level? In electrodynamics the carrier of the force is the spin one photon and in gravitation the carrier of the force is the graviton, the quantum description of gravitational radiation that we will study later. The graviton is described by a traveling wave in the metric $g_{\mu\nu}$ analogous to the photon which is a traveling wave in the electric and magnetic fields. The graviton carries spin two and its spin two nature predicts the universal attraction of the theory. If the graviton is traveling
through empty space, it travels at the speed limit, as will be discussed further after we present linearized General Relativity. Since both the photon and the graviton travel at the speed limit, they have some additional features in common: they both have but two degrees of freedom! In the case of electromagnetic radiation, we have already seen that there are two polarization states which describe it: right-handed or left-handed polarized waves. Contrast this to the spin one state of a massive particle described by quantum mechanics: it has three spin states, which are described by its spin projection in some direction, say the z axis, of -1, 0, or +1. This gives three degrees of freedom. Why does radiation only have two such states, -1 and +1? The special feature of radiation is, of course, that it travels at the speed limit, which is a Lorentz invariant notion. So, its spin state must also be Lorentz invariant. This will be true if the spin points either in the same direction as the propagation of the wave $k$ or in the opposite direction, $-k$. We saw this in our discussion of electromagnetic waves in the lecture series on Special Relativity where we found that right and left handed circularly polarized electromagnetic waves were the only possibilities because of the transversality conditions. The same argument applies to the graviton: it will propagate with its spin 2 aligned with its direction of motion or anti-aligned. In both cases the radiated field has two degrees of freedom, even though they have different spin values! We will check this point when we discuss gravitational radiation and the LIGO experiment.

Now let’s discuss the differences between Eq. (6.1) and (6.2) and between Eq. (6.3) and (6.4). First, the charges $q_1$ and $q_2$ in Eq. (6.1) are Lorentz invariants, they are the same in all frames. The masses, $m_1$ and $m_2$ in Eq. (6.2), are the rest masses of the two particles at fixed positions in the static force law. We need to ask how they transform if we boost Eq. (6.2) or Eq. (6.4) to a frame with velocity $v$. In the case of electromagnetism, the charge density $\rho(r)$ was identified as the zeroth component of a four vector $J_\mu = \rho v_\mu = (\gamma\rho_0 c, \gamma\rho_0 v)$ where $\rho_0$ is the rest
frame charge density and \( \gamma = \gamma(\mathbf{v}) \). The fact that the source of the electric field \( E \) was the zeroth component of a four vector, the charge current, was critical in the derivation of Maxwell’s equations from Gauss’ Law and special relativity.

In the case of gravity, Eq. (6.2) and (6.4), we have the energy density acting as the source of the gravitational field. What are the properties of energy-per-volume under boosts? It is here that gravity and electromagnetism diverge. Energy is the zeroth component of the energy-momentum four vector \( p_\mu = (E/c, \mathbf{p}) \). Under a boost, volume\(^{-1}1/V\), transforms as \( \gamma \) which matches the behavior of the 0th component of a four vector such as the four velocity \( v_\mu = (\gamma c, \gamma \mathbf{v}) \).

So, the right hand side of the equation for the gravitational field, Eq. (6.4), is proportional to the product of the zeroth components of two four vectors. In a general inertial frame the right hand side will be proportional to the product \( p_\mu v_\nu \). If we have an extended source of gravity, perhaps many particles \( i \), then the right hand side of Eq. (6.4) would be proportional to \( \sum_i \mu_0^i c^2 v_l^i v_l^i \) where \( \mu_0^i \) is the rest mass of particle \( i \). Mathematical objects which are labelled by several four vector indices and transform under boosts accordingly are called tensors as we discussed in the lecture series on Special Relativity. They will play central roles in differential geometry and general relativity. We are familiar with the electromagnetic field tensor, \( F_{\sigma \rho} \), for example. The tensor uncovered here is \( T_{\mu \nu} \), the energy-momentum tensor. The energy-mass relation of special relativity implies that all forms of energy-momentum should contribute to \( T_{\mu \nu} \) in addition to the masses that contribute to Newton’s law of gravity. For example, the energy-momentum carried by the electromagnetic field contributes to \( T_{\mu \nu} \). We will pursue these ideas for gravity.

The fact that gravity couples to energy-momentum while electromagnetism couples to charge, a conserved quantity which is only carried by massive matter fields, produces more differences between these theories. In the case of electromagnetism, this property gave rise to the
Linear Superposition Principle which helped us solve electrodynamics problems. Gravity does not enjoy the Linear Superposition Principle and this fact makes it much more challenging. For example, the gravity field itself carries energy-momentum, so it couples to itself! The theory is intrinsically non-linear! The Newtonian limit of the theory is, in fact, linear, but this is an exception to the rule. We will see this explicitly when we solve for the gravitational field outside a spherically symmetric, static mass and discover a black hole. A dramatic implication of the non-linearity of gravity is the following: let’s say you are inside a black hole and try to escape by flailing about or turning on a rocket motor. Do you succeed in evading being sucked in? On the contrary, the extra energy you expend increases your attraction to the black hole and as a consequence you are sucked in more forcefully and your proper lifetime is diminished! In the language of the fundamentals of general relativity, the gravitons which generate the attraction between bodies, attract any object that carries energy-momentum and thus attract one another and generate an apparently very singular theory. This avalanche of interactions is one reason why physicists have yet to understand the quantized version of the theory!

Let’s return to our task of comparing and contrasting classical electromagnetism and classical gravity. We begin with static problems and consider the electrostatic potential far from a collection of charges and the analogous problem in gravity, the gravitational potential far from a collection of masses. Suppose there are charges $e_a$ at points $r_a$ near the origin of a convenient coordinate system. Then the electrostatic potential is given by a linear superposition of Coulomb potentials,

$$V(R) = k \sum_a \frac{e_a}{|R-r_a|}$$

(6.5)
We use the upper case notation \( R \) to indicate the point where the potential is measured and to emphasize that \( R = |R| \) is much larger than the largest \( |r_a| \). In this case it is useful to Taylor expand \( \frac{1}{|R-r_a|} \) around the origin,

\[
\frac{1}{|R-r|} = \frac{1}{R} - \sum_j x_j \frac{\partial}{\partial R_j} \frac{1}{R} + \frac{1}{2} \sum_{jk} x_j x_k \frac{\partial^2}{\partial R_j \partial R_k} \frac{1}{R} + \cdots \tag{6.6}
\]

where \( r = (x_1, x_2, x_3) \). Taking the derivatives and doing some straightforward algebra,

\[
\frac{1}{|R-r|} = \frac{1}{R} + \left( \sum_j x_j n_j \right) \frac{1}{R^2} + \left( \frac{1}{2} \sum_{jk} \left( 3x_j x_k - r^2 \delta_{jk} \right) n_j n_k \right) \frac{1}{R^3} + \cdots \tag{6.7}
\]

where \( n \) is the unit vector pointing in the \( R \) direction. If we apply this expansion to Eq. (6.5) we find,

\[
V(R) = k \left( \frac{1}{R} + D \frac{1}{R^2} + Q^{(2)} \frac{1}{R^3} + \cdots \right) \tag{6.8}
\]

where \( Q = \sum_a e_a \) is the total charge of the source, \( D = \sum_i d_i n_i \), \( d_i = \sum_a e_a x_i^{(a)} \) is the dipole moment of the source, and \( Q^{(2)} = \frac{1}{2} \sum_{jk} D_{jk} n_j n_k \) and \( D_{jk} = \sum_a e_a \left( 3x_j^{(a)} x_k^{(a)} - r^{(a)2} \delta_{jk} \right) \) is the quadrupole moment of the charge distribution.

Eq. (6.8) is particularly useful because it organizes the potential as increasing moments of the charge distribution and correlates these moments to inverse powers of \( R \), indicating how significant they are far from the source. For example, if the source is net neutral, \( Q = \sum_a e_a = 0 \), then the “monopole” term in Eq. (6.8) is absent and the next term, the dipole moment of the distribution dominates the field’s behavior at large \( R \). Since \( D \) is a relatively simple aspect of the charge distribution, it is easy to estimate \( V(R) \) for large \( R \) without knowing every detail of the source. The dipole and the quadrupole moments have particular properties which lead to particular sensitivities in \( V(R) \) to the orientation of the source. For example \( D = d \cdot n \), so \( V(R) \) varies as the cosine of the angle between the source’s dipole moment \( d \) and the orientation \( n \) of the observation position \( R \). Textbooks in electromagnetism develop these points in detail. If the source is neutral,
then it is easy to show that \( d \) is an intrinsic property of the charge distribution, i.e. it is independent of the origin of the coordinate system in which it is defined. If the charge distribution consists of two equal and opposite charges, then \( d \) is the product of the positive charge times the vector from the negative to the positive charge.

How does this compare with the same exercise in Newtonian gravity \( V_g(R) \)? Clearly the arithmetic details are the same with the replacement \( e_a \to m_a \). This replacement leads to two important, and perhaps surprising, differences. First, the strength of the \( R^{-1} \) term in the expansion of \( V_g(R) \) is the total mass of the distribution, \( M = \sum a m_a \). Unlike electrostatics where bulk matter is often neutral \( Q = \sum a e_a = 0 \) and the \( R^{-1} \) term is absent, in gravity the \( R^{-1} \) term is always present. Since each \( m_a \) is positive, there is no “screening” in gravity. Furthermore, we can always choose the origin of the coordinate system at the center of mass of the distribution, so the dipole term vanishes identically, \( r_{cm} = \sum a m_a r^{(a)} = 0 \). If there were no external forces on the collection of masses, then if the center of mass were initially at rest at \( r_{cm} = \sum a m_a r^{(a)} = 0 \), it stays at rest at \( r_{cm} = 0 \) forever. This means that the dipole term vanishes and the first correction to the \( \frac{M}{R} \) term in the potential is the quadrupole term \( \frac{Q^{(2)}}{R^3} \). It is the quadrupole moments of mass distributions that determine tidal forces in planetary systems and the gravitational radiation emitted from binary neutron star systems, as we shall soon see.

The differences between electromagnetism and gravity do not end here. Now let’s think about how one might observe a traveling electromagnetic wave and contrast that to how one could detect a traveling gravity wave. In the electromagnetic case we have the Lorentz force law \( F = dp/dt = e(E + u \times B) \). If the charged particle is at rest in frame \( S \), one could observe a traveling
\( E \) and \( B \) wave by observing the resulting force, a change in the particle’s momentum. No subtlety here. Traveling electromagnetic waves induce electric currents in circuits, etc.

Now consider the same situation in the case of gravity in a curved space. How could we use the kinematic state of an isolated point particle to detect a passing gravity wave? A gravity wave cannot be detected locally! Such a possibility would violate the Equivalence Principle! To understand this, recall that in a local region all the effects of gravity can be transformed away—we can jump into a local inertial reference frame and the effects of gravity disappear. Therefore, the gravity wave cannot be detected by observing what is happening in a strictly local region, in particular by observing one isolated particle. This means that a single isolated particle does not change its position relative to the free falling coordinate labels \((t, x, y, z)\). Subtle! We shall see this principle expressed in the equations of motion of point masses of General Relativity, especially in Appendix G, Sec. 3 of the textbook and lecture 14 below.

So, we can only detect gravity waves by recording their influence on two or more separated test particles! We have already seen above that tidal forces are caused by the spatial variations in gravitational potentials. Later we will see that oscillating quadrupole moments of mass distributions produce gravity waves. The gravity waves influence matter through tidal forces – the characteristic stretching and squeezing illustrated earlier. A gravity wave detector must consist of freely falling masses over extended regions of space-time. This is the concept behind the LIGO detector we will study in lecture 15 below.

Later in this lecture series we will learn that there are two characteristic quadrupole moments of a mass distribution whose oscillations can produce propagating gravitational waves. One is the + polarization which is a “breathing” mode in which “stretching” in one direction is
coordinated with “compression” in the perpendicular direction as shown in Figure 6.1.

Fig. 6.1 A ring of free test particles responds to a passing gravity wave in a + polarization state. Two “time shots” are shown. The wave is traveling into the page.

The second polarization, the × polarization, has the same characteristics but is rotated by 45 degrees. We have seen these patterns in our discussions of tidal forces and will see them again when we study gravitational waves.

Compare this to the dominant radiation patterns in electromagnetism. Oscillating dipole moments of charge distributions produce the strongest radiation: linearly polarized traveling electric and magnetic waves. The two independent linear polarization states are oriented at 90 degrees. Such waves carry considerable energy which we have calculated in problem sets using the Lamor formula. Although a point like local energy density of gravity is inconsistent with the Equivalence Principle, gravity waves carry energy and momentum that can be expressed using the space-time derivatives of the traveling gravitational wave. In fact when black holes collide and merge a great deal of mass is converted to energy and that energy is radiated away. The first such
astronomical event observed in 2015 by the gravity wave detector “Advanced LIGO” converted three solar masses into radiant energy. In the vicinity of the earth, the gravity wave, which originated more than one billion light years away, had a tiny amplitude which could be treated as a slight perturbation on the local Minkowski metric.

These remarks remind us of another important difference between electromagnetism and gravity. How are the detectors of electromagnetic and gravitational radiation different? Typical detectors of light respond to light’s intensity which is proportional to the square of its amplitude. The intensity of light typically decreases as the square of the distance from the source of the radiation. However, the situation in gravity is different. Detectors of gravity waves measure the amplitude of the wave which typically decreases as the inverse of the distance to the source. We will see this explicitly in the sections of this lecture series which focus on LIGO. This fact is critical in the observational science of gravitational waves since events which produce copious gravitational waves are rare and gravitational effects are so intrinsically weak.

7. The Equation of Motion of Particles in Curved Space-Time.

In our lectures on Special Relativity and Classical Differential Geometry, we have formulated and developed the theory of the equations of motion of point particles. Here we will begin a fresh discussion in the realm of curved space time, Riemannian manifolds, with an emphasis on geometric, coordinate independent, formulations of the problem.

Suppose that there is a vector \( \mathbf{V} \) within the curved space time and we can write it in terms of the basis \( \mathbf{b}_\beta, \beta = 1, \ldots, M \), which varies from point to point,

\[
\mathbf{V} = \sum_\beta V^\beta \mathbf{b}_\beta \quad (7.1)
\]
Now suppose that $\mathbf{V}$ sweeps out a curve parametrized by $\tau$. The rate of change of $\mathbf{V}$ along the curve is,

$$\frac{d\mathbf{V}}{d\tau} = \sum_\beta \frac{dV^\beta}{d\tau} \mathbf{b}_\beta + \sum_\beta V^\beta \frac{db_\beta}{d\tau}$$  \hspace{1cm} (7.2)$$

where the second term shows that the curvilinear coordinates have $\tau$ dependence, i.e. the basis vectors vary as the point of interest moves along the curve. Recall that when one solves geometry or physics problems in three dimensions using curvilinear coordinates like spherical or cylindrical coordinates, terms of this sort contribute. They are in fact the extra ingredients in expressions for the gradient, divergence or Laplacian in such coordinate systems.

Next we can use the chain rule to write $\frac{db_\beta}{d\tau}$ in terms of the orientation of the basis in curved space and its movement along the curve,

$$\frac{db_\beta}{d\tau} = \sum_\alpha \frac{\partial b_\beta}{\partial x^\alpha} \frac{dx^\alpha}{d\tau}$$  \hspace{1cm} (7.3)$$

The changing orientation of the basis can be expanded in the basis $(\mathbf{b}_\beta, \beta = 1, \ldots, M)$,

$$\frac{\partial b_\beta}{\partial x^\alpha} = \sum_\gamma \Gamma^\gamma_\beta_\alpha \mathbf{b}_\gamma$$  \hspace{1cm} (7.4)$$

where we have introduced the Christoffel symbols. We see that the Christoffel symbols record the fact that the basis vectors of a curvilinear coordinate system change their orientation as a point moves along a curve and this effects the rate of change of the components of the vector of interest.

Now we can substitute into Eq. (7.2) and relabel the summation indices to our advantage,

$$\frac{d\mathbf{V}}{d\tau} = \sum_\beta \frac{dV^\beta}{d\tau} \mathbf{b}_\beta + \sum_\alpha V^\beta \Gamma^\gamma_\beta_\alpha \frac{dx^\alpha}{d\tau} \mathbf{b}_\gamma$$

$$\frac{d\mathbf{V}}{d\tau} = \sum_\beta \left( \frac{dV^\beta}{d\tau} + \sum_\gamma V^\beta \Gamma^\gamma_\beta_\alpha \frac{dx^\alpha}{d\tau} \right) \mathbf{b}_\beta$$  \hspace{1cm} (7.5)$$

The quantity in parenthesis is the covariant derivative.
\[ \frac{dV^\beta}{d\tau} = \frac{dV^\beta}{d\tau} + \sum_{\alpha\gamma} \Gamma_{\gamma\alpha}^\beta \frac{dx^\alpha}{d\tau} V^\gamma \]  
\hspace{2cm} (7.6)

Since \[ \frac{dV^\beta}{d\tau} = \sum_{\alpha} \frac{\partial V^\beta}{\partial x^\alpha} \frac{dx^\alpha}{d\tau} = \sum_{\alpha} \frac{\partial V^\beta}{\partial x^\alpha} u^\alpha, \] we can write Eq. (7.6),

\[ \frac{dV^\beta}{d\tau} = \sum_{\alpha} u^\alpha D_\alpha V^\beta \]  
\hspace{2cm} (7.7)

and identify the covariant derivative introduced in Eq. (7.5),

\[ D_\alpha V^\beta = \partial_\alpha V^\beta + \sum_{\gamma} \Gamma_{\gamma\alpha}^\beta V^\gamma \]  
\hspace{2cm} (7.8)

If we apply Eq. (7.6) to the four velocity of the particle \( V = u = dx/d\tau \) then the vanishing of its acceleration \( d^2x/d\tau^2 = 0 \) becomes,

\[ \frac{dV^\beta}{d\tau} = 0 \]  
\hspace{2cm} (7.9)

which we recognize as the equation for a geodesic derived in our lecture series on classical differential geometry.

If a vector \( V^\beta \) satisfies \( D_\alpha V^\beta = \partial_\alpha V^\beta + \sum_{\gamma} \Gamma_{\gamma\alpha}^\beta V^\gamma = 0 \), setting Eq. 7.8 to zero, then we say that it is “parallel transported” along the curve traced out by the particle, \( x^\alpha(\tau) \). Note that the tangent vector to the curve, \( dx^\alpha(\tau)/d\tau \), is parallel transported along a geodesic. This serves as the definition of a geodesic, as discussed in the lecture series on classical differential geometry. It is easy to check that the inner product of \( V^\beta \) and the tangent \( u^\beta \) is conserved along the geodesic if \( V^\beta \) is parallel transported. This observation motivates the term “parallel transport”. This and other properties of the geometry of parallel transport, as well as many illustrations in the theory of two dimensional surfaces, were discussed in the differential geometry lecture series and notes, Supplementary Lecture 9. In the context of these lectures, the reader could provide the arguments for these statements in the language of Riemannian geometry after reading about Metric Compatibility in lecture 9 below.
In our applications to relativity we begin with the absence of acceleration in a local inertial frame, \( \frac{dv}{d\tau} = 0 \). This maps onto the equation of motion Eq. (7.9) in the curvilinear coordinate system that describes the same physics in a non-inertial frame of reference where the gravitational effects are experienced as the curvature of space-time. Eq. (7.9) describes the motion of massive particles in a gravitational field. We will see how all this works out as we develop the physics of General Relativity.

Note that only the symmetric piece of \( \Gamma^{\beta}_{\gamma\alpha} \) contributes to Eq. (7.9), so we treat the Christoffel symbols in our applications to relativity as explicitly symmetric in the lower indices, \( \Gamma^{\beta}_{\gamma\alpha} = \Gamma^{\beta}_{\alpha\gamma} \).

Let’s illustrate some of the ideas here by considering a familiar example: Newtonian mechanics in two dimensional polar coordinates. We want to calculate the Christoffel symbols and use them to write Newton’s second law in plane polar coordinates. The problem set also explores these examples in greater detail and from several different perspectives. In this Newtonian problem the force of gravity is inserted explicitly into the equation of motion. We will contrast this later with similar problems in general relativity where the “force” emerges from the curvature of space time through the Christoffel symbols.

Consider a non-relativistic particle in a gravitational potential generated by a mass \( M \) at the origin,

\[
\frac{d^2 r}{dt^2} = -\frac{GM}{r^2} \hat{r}
\]

This is a problem with a central potential so the motion of the particle in the potential will be planar. The particle’s velocity in plane polar coordinates is \( u^i = \left( \frac{dr}{dt}, \frac{d\theta}{dt} \right) \) and the equation of motion reads,

\[
\frac{du^i}{dt} + \sum_{kl} \Gamma^i_{kl} u^k u^l = -\frac{GM}{r^2} \delta^i_1
\]  

(7.10)
We need the Christoffel symbols. Let’s calculate them for plane polar coordinates using the geometric approach,

$$\partial_\beta e_\alpha = \sum_\gamma \Gamma^\gamma_{\alpha\beta} e_\gamma$$  \hspace{1cm} (7.11)

In Cartesian coordinates \( r = r \cos \theta \, i + r \sin \theta \, j \) and the coordinate vectors of plane polar coordinates are,

$$\frac{\partial r}{\partial r} = e_r = \cos \theta \, i + \sin \theta \, j \quad \frac{\partial r}{\partial \theta} = e_\theta = -r \sin \theta \, i + r \cos \theta \, j$$

Then one can calculate,

$$\frac{\partial e_r}{\partial r} = 0, \quad \frac{\partial e_r}{\partial \theta} = \frac{1}{r} e_\theta, \quad \frac{\partial e_\theta}{\partial r} = \frac{1}{r} e_\theta, \quad \frac{\partial e_\theta}{\partial \theta} = -r e_r$$

We can read off the Christoffel symbols from these results using Eq. (7.11),

$$\Gamma^r_{\theta\theta} = \Gamma^\theta_{r\theta} = \frac{1}{r}, \quad \Gamma^r_{\theta\theta} = -r$$

with the other Christoffel symbols vanishing. The equations of motion Eq. (7.10) become,

$$\frac{d^2 r}{dt^2} - r \left( \frac{d \theta}{dt} \right)^2 = -\frac{GM}{r^2}$$

$$\frac{d^2 \theta}{dt^2} + \frac{2 \, dr \, d\theta}{r \, dt \, dt} = 0$$

These are the results we wanted. The radial equation shows how the Christoffel symbols for the curvilinear coordinate system, plane polar coordinates, brings in the expected Centripetal “apparent” force. Similarly the Christoffel symbols bring in the Coriolis “apparent” force into the equation of motion for the angular variable. The reader should recognize this equation as the conservation law for the orbital angular momentum per mass, \( r^2 \frac{d\theta}{dt} = L \).

Now it is an easy and interesting exercise to write Newton’s second law and the Lorentz force law of electrodynamics in curvilinear coordinates, suitable for applications in general relativity.

Newton’s second law was written in four vector form in the lecture series on Special Relativity. In
Minkowski space-time we had,

\[ f^\mu = m \frac{d^2 x^\mu}{d\tau^2} = \frac{d}{d\tau} p^\mu \]  

(7.12)

where \( \tau \) is the proper time, \( p^\mu \) is the energy-momentum four vector, \( p^\mu = mu^\mu = m \frac{dx^\mu}{d\tau} \) and \( f^\mu \) is the four vector force. To write Eq. (7.12) in curvilinear coordinates, the ordinary differentials become covariant differentials which we write out in terms of Christoffel symbols,

\[ \frac{1}{m} f^\mu = \frac{du^\mu}{d\tau} = \frac{du^\mu}{d\tau} + \sum_{\nu,\sigma} \Gamma_{\nu\sigma}^{\mu} u^\nu u^\sigma = \frac{d^2 x^\mu}{d\tau^2} + \sum_{\nu,\sigma} \Gamma_{\nu\sigma}^{\mu} u^\nu u^\sigma \]  

(7.13a)

The Lorentz force law is a special case of this expression where \( f^\mu = q \sum_{\lambda} u_\lambda F_\lambda^\mu \), as we learned when we discussed Special Relativity. There we wrote \( f = q(E + u \times B) = dp/dt \) in Minkowski covariant form,

\[ \frac{q}{m} \sum_{\lambda} u_\lambda F_\lambda^\mu = \frac{d^2 x^\mu}{d\tau^2} + \sum_{\nu,\sigma} \Gamma_{\nu\sigma}^{\mu} u^\nu u^\sigma \]  

(7.13b)

Once we learn to calculate the Christoffel symbols more generally, we shall be able to write the Lorentz force law in curved space-time, outside a rotating black hole, for example.

In the textbook, and especially in its Appendices, such as C.2, we consider tensor transformation laws. These laws are very important because if we write a physics equation in tensor notation, then we learn that if it is true in one frame, then it is true in all frames. In the context of Special Relativity, this motivated us to write Maxwell’s equations and the Lorentz force law in this language. These ideas are even more important in General Relativity where the transformation laws become space time dependent. As we saw in the developments above, the space time dependence of the coordinates adds new considerations to the construction of tensors. We saw that the familiar partial differentiation with respect to coordinates is not generally a tensor operation: we needed to construct the covariant derivative to accomplish that.
Let’s consider a few more examples of space time dependent transformation laws and then state the general results. If we have a four dimensional displacement vector, we can expand it in the basis vectors of any frame,

\[ ds = \sum_a dx^a e_a = \sum_a dx'^a e'_a \]  

where we imagined a transformation \( x \rightarrow x'(x) \). We want to know how \( \{e'_a\} \) is related to \( \{e_a\} \).

But the chain rule tells us how the coordinates transform, \( dx^a = \sum_b \frac{\partial x^a}{\partial x'^b} dx'^b \), so

\[ ds = \sum_a dx^a e_a = \sum_{ab} \frac{\partial x^a}{\partial x'^b} dx'^b e_a = \sum_b dx'^b \left( \sum_a \frac{\partial x^a}{\partial x'^b} e_a \right) \]

Identify the transformation law for the covariant basis vectors,

\[ e'_a = \sum_b \frac{\partial x^b}{\partial x'^a} e_b \]  

7.16a

The transformation law for the contravariant basis vectors follow similarly,

\[ e'^a = \sum_b \frac{\partial x'^a}{\partial x^b} e^b \]  

7.16b

Covariant and contravariant basis vectors were introduced in the Appendix C.2 of the textbook and their geometric meaning was reviewed. These transformation laws were found there as well.

We can check the transformation law Eq. 7.16b by considering the orthonormality relations,

\[ e'^a \cdot e'_b = e^a \cdot e_b = \delta^a_b \]

Substituting in Eq. 7.16 a and b,

\[ \left( \sum_c \frac{\partial x'^a}{\partial x^c} e^c \right) \cdot \left( \sum_d \frac{\partial x^d}{\partial x'^b} e_d \right) = \sum_c \frac{\partial x'^a}{\partial x^c} \frac{\partial x^c}{\partial x'^b} = \frac{\partial x'^a}{\partial x'^b} = \delta^a_b \]

and everything works out.

Now take a vector and find its transformation laws,

\[ V = \sum_a V^a e_a = \sum_a V'^a e'_a \]

We can isolate the transformed components by taking projections onto the basis vectors,
\[ V'^a = e'^a \cdot V = \sum_b \frac{\partial x'^a}{\partial x^b} e^b \cdot V = \sum_b \frac{\partial x'^a}{\partial x^b} V^b \quad \text{(7.17a)} \]

and,

\[ V_a' = e'_a \cdot V = \sum_b \frac{\partial x'^a}{\partial x^b} e_b \cdot V = \sum_b \frac{\partial x'^a}{\partial x^b} V^b \quad \text{(7.17b)} \]

The generalization of these transformation laws to general higher rank tensors is clear: each covariant and contravariant index will transform by either Eq. 7.17a or b. These results are discussed more thoroughly in the textbook. The metric itself provides an example,

\[ g^c_d = \sum_{ab} \frac{\partial x^a}{\partial x^c} \frac{\partial x^b}{\partial x^d} g_{ab} \quad \text{(7.18)} \]

where the arguments of the metric label the same physical point, \( x' \) on the left and \( x(x') \) on the right. Let’s check this transformation law. The metric is defined by,

\[ ds^2 = \sum_{ab} g_{ab} dx^a dx^b = \left( \sum_{ab} \frac{\partial x^a}{\partial x^c} \frac{\partial x^b}{\partial x^d} g_{ab} \right) dx^c dx^d = \sum_{cd} g^c_d dx^c dx^d \]

where we used \( dx^a = \sum_b \frac{\partial x^a}{\partial x^b} dx'^b \) twice and we see that Eq. 7.18 follows.

Other examples of tensors include the field strength tensor \( F_\sigma \) of electrodynamics, the energy-momentum tensor \( T_\sigma \) that is the source of gravity in General Relativity, its covariant derivative \( D_\mu T_\sigma \), etc.

### 8. Covariant Derivatives and Covariant Vector Fields.

In the text above we considered covariant derivatives acting on contravariant vectors. We also need to differentiate covariant vectors. The general form of the relation must be,

\[ D_\mu W_\sigma = \partial_\mu W_\sigma + \sum_\gamma \bar{\Gamma}_\mu^\gamma_{\nu \sigma} W_\gamma \]

where we have borrowed the general structure of the equation from the discussion above
for contravariant vectors, but the coefficients in the parallel transport of covariant vectors, $\tilde{\Gamma}$, must be determined.

To relate $\tilde{\Gamma}$ to $\Gamma$, let’s exploit the fact that $\sum_\gamma V_\gamma W^\gamma$ is a scalar function if $V_\gamma$ and $W^\gamma$ are four vectors. For scalar functions $D_\mu$ reduces to ordinary $\partial_\mu$,

$$D_\mu (\sum_\gamma V_\gamma W^\gamma) = \partial_\mu (\sum_\gamma V_\gamma W^\gamma) = \sum_\gamma (\partial_\mu V_\gamma) W^\gamma + \sum_\gamma V_\gamma (\partial_\mu W^\gamma)$$  \hspace{1cm} (8.1)

But the covariant derivative satisfies,

$$D_\mu \left( \sum_\gamma V_\gamma W^\gamma \right) = \sum_\gamma (D_\mu V_\gamma) W^\gamma + \sum_\gamma V_\gamma (D_\mu W^\gamma)$$

just like ordinary differentiation, so,

$$D_\mu (\sum_\gamma V_\gamma W^\gamma) = \sum_\gamma (\partial_\mu V_\gamma + \sum_\alpha \tilde{\Gamma}_{\mu\gamma}^\alpha V_\alpha) W^\gamma + \sum_\gamma V_\gamma (\partial_\mu W^\gamma + \sum_\alpha \Gamma_{\mu\gamma}^\alpha W^\alpha)$$  \hspace{1cm} (8.2)

Comparing Eq. (8.1) and (8.2), we learn that,

$$0 = \sum_{\alpha\gamma} \left( \tilde{\Gamma}_{\mu\gamma}^\alpha V_\alpha W^\gamma + \Gamma_{\mu\gamma}^\alpha V_\gamma W^\alpha \right)$$

Since this expression must be true for all $V_\alpha$ and $W^\beta$, we learn that the coefficients must satisfy,

$$\tilde{\Gamma}_{\mu\gamma}^\alpha = -\Gamma_{\mu\gamma}^\alpha$$

In summary, for each upper index,

$$D_\mu V^\sigma = \partial_\mu V^\sigma + \sum_\gamma \Gamma_{\mu\gamma}^\sigma V^\gamma$$

and for each lower index we have now,

$$D_\mu W_\sigma = \partial_\mu W_\sigma - \sum_\gamma \Gamma_{\mu\sigma}^\gamma W_\gamma$$

Further manipulations show that these rules generalize to tensors with upper and lower indices.

For example, one can show that,
\[
D_\mu T^\rho_\sigma = \partial_\mu T^\rho_\sigma - \sum_\gamma \Gamma^\gamma_\mu_\sigma T^\rho_\gamma + \sum_\gamma \Gamma^\rho_\mu_\gamma T^\gamma_\sigma
\]

There are other properties of the covariant derivative we need throughout the text and some we have already anticipated,

1. When the covariant derivative acts on a scalar function, it reduces to the ordinary gradient,
\[
D_\rho f(x) = \partial_\rho f(x)
\]

2. The covariant derivative satisfies the “Leibniz rule” of ordinary differentiation. For example, the ordinary gradient satisfies,
\[
\partial_\mu (T_{\alpha\beta} V^\gamma) = (\partial_\mu T_{\alpha\beta}) V^\gamma + T_{\alpha\beta} (\partial_\mu V^\gamma)
\]
and the covariant derivative does also,
\[
D_\mu (T_{\alpha\beta} V^\gamma) = (D_\mu T_{\alpha\beta}) V^\gamma + T_{\alpha\beta} (D_\mu V^\gamma) \quad (8.3)
\]

The proof of Eq. (8.3) notes that it is true in the tangent space where the Christoffel symbols vanish in Cartesian coordinates and the covariant derivatives reduce to ordinary derivatives. But Eq. (8.3) is an expression involving four vectors when it is written in terms of \( D_\mu \), so if it is true in one coordinate system it is true in all. Note that the four vector character of the covariant derivative is essential in this slick line of argument.

Let’s consider a more fundamental and geometric derivation of the covariant derivative of a covariant vector field. Begin with an abstract four vector,
\[
\mathbf{V} = \sum_a V^a \mathbf{e}_a = \sum_a V_a \mathbf{e}_a
\]
So,
\[
\partial_b \mathbf{V} = \sum_a (\partial_b V_a) \mathbf{e}_a + \sum_a V_a (\partial_b \mathbf{e}_a) \quad 8.4
\]
We need \( \partial_b \mathbf{e}_a \) to proceed. We had \( \partial_b \mathbf{e}_a = \sum_d \Gamma^d_{ac} \mathbf{e}_d \), which defined the Christoffel symbols. We can isolate \( \Gamma^b_{ac} \) by projecting against a basis vector,
\[ e^d \cdot \partial_e e_a = \sum_b \Gamma_{ac}^b e^d \cdot e_b = \Gamma_{ac}^d \]

where we recalled the definition of the dual basis vectors,
\[ e^a \cdot e_b = \delta^a_b \]

In order to find \( \partial_b e^a \), consider,
\[ \partial_c (e^a \cdot e_b) = (\partial_c e^a) \cdot e_b + e^a \cdot (\partial_c e_b) = 0 \]

Substituting in \( \partial_c e_b = \sum_d \Gamma_{bc}^d e_d \), we learn that,
\[ \partial_c e^a = -\sum_b \Gamma_{bc}^a e^b \]

Substituting into Eq. 8.4,
\[ \partial_b V = \sum_a (\partial_b V_a) e^a - \sum_a \Gamma_{ab}^d e^d V_a = \sum_d \left( \partial_b V_d - \sum_a \Gamma_{ab}^a V_a \right) e^d \]

we identify,
\[ D_b V_d = \partial_b V_d - \sum_a \Gamma_{ab}^a V_a \]

as before.

9. The Equivalence Principle, Metric Compatibility and Christoffel Symbols

We introduced tangent spaces when we reviewed classical differential geometry. The tangent space is the best linear approximation to the curved space at and near a point \( P \).

For a surface embedded in Euclidean space a local coordinate mesh can be chosen so that its metric satisfies,
\[ g_{jk} = \delta_{jk} + O \left( \frac{|\Delta x|^2}{R^2} \right) \]

where \(|\Delta x|^2\) is the distance from \(P\) on the tangent plane and \(R\) is a measure of the curvature of the surface at \(P\), such as a principle radius of curvature there. One says that the surface is “locally flat”.

In the case of relativity we have,

\[ g_{\alpha\beta}(P) = \eta_{\alpha\beta}, \quad \partial_\gamma g_{\alpha\beta}(P) = 0, \quad \partial_\gamma \partial_\mu g_{\alpha\beta}(P) \neq 0 \] (9.1)

where \(\eta_{\alpha\beta}\) is the Minkowski metric. The equation \(g_{\alpha\beta}(P) = \eta_{\alpha\beta}\) is the first part of an expression of the Equivalence Principle: the freely falling frame at \(P\) is described locally by an inertial reference frame of special relativity. The second part of the statement is \(\partial_\gamma g_{\alpha\beta}(P) = 0\) in the freely falling frame. This means that there are no net forces in this frame at \(P\): freely falling frames are locally inertial. This is the essence of the Equivalence Principle and it allows gravity to be interpreted as an aspect of geometry. The last expression, \(\partial_\gamma \partial_\mu g_{\alpha\beta}(P) \neq 0\), means that the curvature of space-time is a frame independent physical effect which cannot be transformed away by a choice of coordinates. More about this later when we develop the Riemann tensor.

Before continuing we should carefully check that it is always possible to find coordinates \(x'^\mu\) near the point \(P\) so that,

\[ g_{\alpha\beta}(P) = \eta_{\alpha\beta}, \quad \partial_\gamma g_{\alpha\beta}(P) = 0 \]

To begin, recall the transformation law,

\[ g'_{cd} = \sum_{ab} \frac{\partial x^a}{\partial x'^c} \frac{\partial x^b}{\partial x'^d} g_{ab} \]

In the vicinity of the point \(P\) we can expand the transformation between coordinate systems,
\[ x^a(x') = x^a(P) + \sum_b \left( \frac{\partial x^a}{\partial x'^b} \right)_P (x'^b - x'^b_P) + \frac{1}{2} \sum_{bc} \left( \frac{\partial^2 x^a}{\partial x'^b \partial x'^d} \right)_P (x'^b - x'^b_P)(x'^d - x'^d_P) + \cdots \]

Now count the number of parameters in the transformation:

1. \( \left( \frac{\partial x^a}{\partial x'^b} \right)_P \) has \( 4 \times 4 = 16 \)

2. \( \left( \frac{\partial^2 x^a}{\partial x'^b \partial x'^d} \right)_P \) has \( 4 \times 10 = 40 \)

3. \( \left( \frac{\partial^3 x^a}{\partial x'^b \partial x'^c \partial x'^d} \right)_P \) has \( 4 \times 20 = 80 \)

Finally count the number of parameters in the metric:

4. \( g_{\alpha\beta}(P) \) has 10

5. \( \partial_\gamma g_{\alpha\beta}(P) \) has \( 4 \times 10 = 40 \)

6. \( \partial_\gamma \partial_\sigma g_{\alpha\beta}(P) \) has \( 10 \times 10 = 100 \)

So, there are enough parameters to arrange that \( g_{\alpha\beta}(P) = \eta_{\alpha\beta} \) and \( \partial_\gamma g_{\alpha\beta}(P) = 0 \), but there are not enough to set the second derivatives of the metric to zero. So, indeed, we can choose to use the Minkowski metric and have no forces at the point P, but we cannot transform away all the components of space time’s curvature. This agrees with the intuitive, physical argument.

Now let’s return to Eq. (9.1). It leads to a very useful formula for the Christoffel symbols in terms of the metric and its derivatives. We learned in the previous lecture that statements involving ordinary derivatives of tensors in the freely falling frame become statements about the covariant derivative of those tensors in a general coordinate frame. So the second equality in Eq. (9.1) becomes,

\[ D_\gamma g_{\alpha\beta}(P) = 0 \]
This relation holds for any arrangement of the tensor indices. Using the results of the previous lecture on the covariant derivative, we can write this expression out explicitly,

\[
\partial_\rho g_{\alpha\beta} - \sum_\mu \Gamma^\mu_{\rho\alpha} g_{\mu\beta} - \sum_\mu \Gamma^\mu_{\beta\rho} g_{\alpha\mu} = 0
\]

\[
\partial_\alpha g_{\beta\rho} - \sum_\mu \Gamma^\mu_{\alpha\beta} g_{\mu\rho} - \sum_\mu \Gamma^\mu_{\rho\alpha} g_{\beta\mu} = 0
\]

\[
\partial_\beta g_{\rho\alpha} - \sum_\mu \Gamma^\mu_{\beta\rho} g_{\mu\alpha} - \sum_\mu \Gamma^\mu_{\alpha\beta} g_{\rho\mu} = 0
\]

We can isolate a Christoffel symbol by subtracting the second and third equations from the first one and use the fact that the Christoffel symbol and the metric tensor are symmetric in their two lower indices to find,

\[
2 \sum_\mu \Gamma^\mu_{\rho\beta} g_{\alpha\mu} + \partial_\alpha g_{\beta\rho} - \partial_\rho g_{\beta\alpha} - \partial_\beta g_{\alpha\rho} = 0
\]

Finally we multiply through by the inverse of \( g_{\alpha\mu} \), i.e. \( g^{\rho\mu} \), and obtain the result,

\[
\Gamma^\mu_{\rho\beta} = \frac{1}{2} \sum_\alpha g^{\mu\alpha} \left( \partial_\rho g_{\beta\alpha} + \partial_\beta g_{\alpha\rho} - \partial_\alpha g_{\rho\beta} \right)
\]

(9.2)

So, the Christoffel symbols can be computed in terms of the metric and its derivatives. The presence of \( g^{\mu\alpha} \), the inverse of the metric, in Eq. (9.2) indicates the formula is also non-linear in the metric itself.

The equation \( D_\gamma g_{\alpha\beta}(P) = 0 \), which is called “metric compatibility”, is particularly useful when manipulating equations. It means that raising and lowering indices pass through the covariant derivative. For example,

\[
\sum_\beta g_{\alpha\beta} D_\rho V^\beta = \sum_\beta D_\rho \left( g_{\alpha\beta} V^\beta \right) = D_\rho V_\alpha
\]

Later in this lecture series we will introduce the Einstein field equations which express the idea that energy-momentum distributions warp space-time and produce its curvature. This
curvature means that there is a metric tensor $g_{\mu\nu}$ which depends on $x$. From $g_{\mu\nu}$ one can calculate the Christoffel symbols and compute the motion of massive particles. As the massive particles move, the energy-momentum tensor varies and the gravitational field changes which alters the motion of the massive particles, etc. Einstein’s equation for the gravitational field and the equation of motion are a coupled system of differential equations which in principle solve the problem of the cosmos.

10. The Curvature of Space-Time.

Now let’s return to parallel transport and relate it to the curvature of space-time. In a flat Minkowski space time a particle’s four velocity is a constant of motion. This means that $du^\mu = 0$ in a force free environment. We learned that this implies that $Du^\mu = 0$ in curved space time. This means that $u^\mu = dx^\mu / d\tau$ parallel transports along its trajectory from $x$ to $x + dx$. We already observed that under parallel displacement of two vectors, the angle between them remains constant. Therefore, the angle between a general parallel transported vector and the tangent to the geodesic remains unchanged, i.e. the components of the vector along the geodesic remain unchanged all along the path.

It is also clear that parallel displacement of a vector along a curve between two points will depend on the path chosen [4]. In particular, if a vector is parallel displaced along a closed path, upon returning to its initial point it will not coincide with its initial orientation. It is easy to demonstrate this by visualizing parallel transportation of vectors around the surface of a sphere of radius $R$, as discussed and illustrated in the lecture series on differential geometry.

One can generalize this observation by considering an infinitesimal closed path along which we parallel transport a vector $A^\mu$ [5],
\[ \Delta A^\mu = \oint \Gamma^\mu_{\alpha\beta} A^\alpha d\chi^\beta \]

Since \(\Delta A^\mu\) is the difference of two vectors at one point, it is a contravariant vector. As demonstrated in the lecture series on differential geometry, \(\Delta A^\mu\) is non-zero only because space-time is curved.

Another way to view this construction is to focus on the path dependence of parallel translation. In general, if we parallel transport a vector \(A^\mu\) along two separate paths from point 1 to point 2, the results will be different. This observation produces the concept of holonomy for the intrinsic curvature. Let \(D_\mu\) represent covariant differentiation in the direction \(\mu\) and let \(D_\nu\) represent covariant differentiation in the direction \(\nu\). Then consider two points \(p\) and \(q\) which are separated by two differentials, \(dx^\mu\) in the \(\mu\) direction and another, \(dx^\nu\) in the \(\nu\) direction. One can travel from \(p\) to \(q\) in two ways: first, move infinitesimally in the \(\mu\) direction and then in the \(\nu\) direction, or, second, move infinitesimally in the \(\nu\) direction and then in the \(\mu\) direction. The difference in the results of doing the translations from \(p\) to \(q\) along the two paths is a measure of the local curvature. One can write,

\[ D_\mu D_\nu A^\rho - D_\nu D_\mu A^\rho = -\sum_{\sigma} R^\rho_{\sigma\mu\nu} A^\sigma \]

which can be written as a “commutator”,

\[ (D_\mu D_\nu - D_\nu D_\mu)A^\rho = -\sum_{\sigma} R^\rho_{\sigma\mu\nu} A^\sigma \quad (10.1) \]

This equation, which defines the Riemann curvature tensor, \(R^\rho_{\sigma\mu\nu}\), is visualized in Figure 10.1. This is the generalization of the result in differential geometry where the right hand side was proportional to the Gaussian curvature.
Fig. 10.1 The closed path of covariant differentials used in defining local curvature.

Since the left hand side of Eq. (10.1) is the difference of two four vectors, the relation is a valid tensor equation which holds in any curvilinear coordinate system. In addition, the fourth(!) rank tensor in Eq. (10.1) \( R^\sigma_{\mu\nu\rho} \), the Riemann curvature tensor, is independent of the vector \( A_\rho \) used in the construction.

We will carry through the calculation in Eq. (10.1) and obtain an explicit formula for the Riemann curvature tensor, in terms of the Christoffel symbols and their first derivatives. The result will be,

\[
R^\rho_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \sum_\alpha \Gamma^\rho_{\mu\alpha} \Gamma^\alpha_{\nu\sigma} - \sum_\alpha \Gamma^\rho_{\nu\alpha} \Gamma^\alpha_{\mu\sigma} \quad (10.2)
\]

To derive this result, write out the commutator using the definition of the covariant derivative,
\[
\begin{align*}
D_\alpha D_\beta V^\rho - D_\beta D_\alpha V^\rho &= \partial_\alpha (D_\beta V^\rho) - \sum_\gamma \Gamma^\gamma_\alpha_\beta D_\gamma V^\rho + \sum_\gamma \Gamma^\rho_\alpha_\gamma D_\beta V^\gamma - (\alpha \leftrightarrow \beta) \\
&= \partial_\alpha \partial_\beta V^\rho + \sum_\gamma \left( \partial_\alpha \Gamma^\rho_\beta_\gamma V^\gamma + \sum_\gamma \Gamma^\rho_\beta_\gamma \partial_\alpha V^\gamma - \sum_\gamma \Gamma^\gamma_\alpha_\beta \partial_\gamma V^\rho \\
&\quad - \sum_\gamma \Gamma^\gamma_\alpha_\beta \Gamma^\rho_\gamma_\sigma V^\sigma + \sum_\gamma \Gamma^\rho_\alpha_\gamma \partial_\beta V^\gamma + \sum_\gamma \Gamma^\rho_\alpha_\sigma \Gamma^\gamma_\beta_\gamma V^\gamma - (\alpha \leftrightarrow \beta) \\
&= \sum_\sigma \left( \partial_\alpha \Gamma^\rho_\beta_\sigma - \partial_\beta \Gamma^\rho_\alpha_\sigma + \sum_\gamma \Gamma^\rho_\alpha_\gamma \Gamma^\gamma_\beta_\sigma - \sum_\gamma \Gamma^\rho_\beta_\gamma \Gamma^\gamma_\alpha_\sigma \right) V^\sigma
\end{align*}
\]

where the symmetry of the Christoffel symbol, \( \Gamma^\rho_\alpha_\beta = \Gamma^\rho_\beta_\alpha \), and the \( \alpha \leftrightarrow \beta \) terms led to considerable cancellation.

The Riemann tensor is very imposing since it has \( 4 \times 4 \times 4 \times 4 = 256 (!) \) components. However, it is highly constrained by symmetries. The Riemann tensor symmetry properties can be derived from formula Eq. (10.2) (Some are clear by inspection but others require work. They are derived in the problem set in the textbook.). First, lower the index on the tensor,

\[
R_{\rho\sigma\alpha\beta} = \sum_\gamma g^\gamma_\rho R^\gamma_\sigma\alpha\beta
\]

(10.3)

Then the symmetry properties read,

1. \( R_{\rho\sigma\alpha\beta} = -R_{\sigma\rho\alpha\beta} \)
2. \( R_{\rho\sigma\alpha\beta} = -R_{\rho\sigma\beta\alpha} \)
3. \( R_{\rho\sigma\alpha\beta} = R_{\alpha\beta\rho\sigma} \)
4. \( R_{\rho\sigma\alpha\beta} + R_{\rho\alpha\beta\sigma} + R_{\rho\beta\sigma\alpha} = 0 \)
5. \( D_\gamma R_{\rho\sigma\alpha\beta} + D_\rho R_{\sigma\gamma\alpha\beta} + D_\sigma R_{\gamma\rho\alpha\beta} = 0 \) (Bianchi Identity)

Other important curvature tensors of lower rank can be constructed from the Riemann tensor. The second rank Ricci tensor is
\[ R_{\alpha\beta} = \sum_{\gamma} R^\gamma_{\alpha\gamma \beta} \equiv \sum_{\gamma\sigma} g^{\gamma\sigma} R_{\sigma\alpha\gamma\beta} \] (10.4)

Properties 1 and 2 imply that it is symmetric, \( R_{\alpha\beta} = R_{\beta\alpha} \). Note that the other possible contractions like \( \sum_{\gamma} R^\gamma_{\gamma\alpha\beta} = \sum_{\gamma\sigma} g^{\gamma\sigma} R_{\gamma\alpha\beta} \) or \( \sum_{\gamma} R^\gamma_{\alpha\beta\gamma} = \sum_{\gamma\sigma} g^{\gamma\sigma} R_{\alpha\beta\gamma\sigma} \) vanish identically because of properties 1 and 3. This observation implies that the Ricci tensor \( R_{\alpha\beta} \) is the \textit{only} second rank tensor that can be constructed from the Riemann tensor.

The Ricci scalar is the fully contracted version,

\[ R = \sum_{\gamma} R^\gamma_{\gamma} = \sum_{\alpha\beta} g^{\alpha\beta} R_{\alpha\beta} \] (10.5)

In the case of two dimensional surfaces in three dimensional Euclidean space, the Ricci scalar is just twice the Gaussian curvature \( K \). This point is demonstrated in the problem set of the textbook.

In four dimensions the \textit{full} Riemann tensor \( R_{\alpha\beta\gamma\delta} \) is generally required to specify the curvature of space-time.


Now that differential geometry and curvature have been introduced, we can turn back to physics and consider the differential equation for the gravitational potential in Newton’s world,

\[ \nabla^2 V(r) = 4\pi G \rho(r) \] (11.1)

where the source of the potential is the mass density \( \rho(r) \) and the force on a particle of mass \( m \) is \( F(r) = -m\nabla V(r) \). We want to recast this equation in the language of general relativity, if possible.

We have seen in several examples that the Newtonian gravitational potential modifies the time-time component of the metric: \( g_{00} \) becomes \( 1 + \frac{2V(r)}{c^2} \) in the case of weak gravitational
potentials, $2V(r)/c^2 \ll 1$. If there is a particle of mass $M$ at the origin, $V(r) = -\frac{GM}{r}$ is the Newtonian gravitational potential. Then the time-time component of the metric becomes,

$$g_{00} = \eta_{00} + \frac{2V(r)}{c^2} = 1 - \frac{2GM}{c^2r} \quad (11.2)$$

We can write Eq. 11.1 in the more suggestive form,

$$\nabla^2 g_{00} = \frac{8\pi G}{c^4} T_{00} \quad (11.3)$$

where we identified the $0-0$ component of the energy-momentum tensor, $T_{00} = c^2\rho$. We will introduce the full tensor $T_{\mu\nu}$ later.

Is Eq. 11.3 the non-relativistic limit of a general covariant field equation for gravity? What are the constructive principles to guide us toward the fundamental differential equations of general relativity? They read:

1. The equation should be invariant to general coordinate transformations: it should be true in all frames of reference.

2. It should express the idea that the energy-momentum density is the source of the curvature of space time.

3. It should be a second order differential equation for the space time metric $g_{\sigma\rho}$.

4. It should embody conservation of energy-momentum locally in space-time.

5. It should reduce to special relativity far from any source of energy-momentum.

6. It should reproduce Newton’s theory of gravity in the non-relativistic, weak field limit.

We have already discussed these principles, but have not emphasized item 4 earlier. We have illustrated the energy-momentum tensor for free massive particles and in a problem set on electrodynamics we constructed the energy-momentum tensor from the electrodynamic field $F_{\mu\nu}$.

We want $T_{\mu\nu}$ to be subject to local energy-momentum conservation through a local differential
equation. The conservation law should be a true tensor relation so that it will be true in all frames of reference and will be consistent with causality. Recall our discussion of charge conservation in electrodynamics in Minkowski space time that led us to the continuity equation $\sum_{\mu} \partial^{\mu} J_{\mu} = 0$ for the current density $J_{\mu}$. The equation had a simple physical and geometric interpretation: the local charge density $\rho$ in a box changes only when currents $\mathbf{J}$ flow into or out of that box. Now we express the conservation of energy-momentum in a similar fashion. In flat Minkowski space time we have,

$$\sum_{\mu} \partial^{\mu} T_{\mu\nu} = 0$$

and $T_{00}$ is the energy density and $T_{0i}$ is the $i^{th}$ component of momentum density. $T_{\mu\nu}$ is constructed to be symmetric, as was clear from our free particle example and electrodynamics. In a curved space time Eq. (11.4) becomes

$$\sum_{\mu} D^{\mu} T_{\mu\nu} = 0$$

Now we need the left hand side of the equation, $\mathcal{E} = \frac{8\pi}{c^4} G T_{\mu\nu}$. It should satisfy several criteria,

1. It should be a local, conserved second rank tensor.
2. It should be constructed out of $g_{\sigma\rho}$, $g^{\sigma\rho}$ and their first and second derivatives in space-time.
3. It should represent the local intrinsic curvature of space time.

The only candidate tensor with most of these properties is the Ricci tensor $R_{\sigma\rho}$. However, we have not checked if it is conserved. To begin we know something about the covariant derivatives of the Riemann tensor $R_{\sigma\rho\nu\mu}$: it satisfies the Bianchi identities of differential geometry, symmetry property 5 discussed in the previous section,

$$D_{\gamma} R_{\rho\sigma\alpha\beta} + D_{\rho} R_{\sigma\gamma\alpha\beta} + D_{\sigma} R_{\gamma\rho\alpha\beta} = 0$$
This allows us to calculate the covariant four-divergence of the Ricci tensor,

\[ 0 = \sum_{\sigma\gamma} g^{\sigma\gamma} g^{\alpha\beta} \left( D_\gamma R_{\rho\sigma\alpha\beta} + D_\rho R_{\sigma\gamma\alpha\beta} + D_\sigma R_{\gamma\rho\alpha\beta} \right) = \sum_{\gamma} D^\gamma R_{\rho\gamma} - D_\rho R + \sum_{\gamma} D^\gamma R_{\rho\gamma} \]

So,

\[ \sum_{\gamma} D^\gamma R_{\rho\gamma} = \frac{1}{2} D_\rho R \quad (11.5) \]

So, the Ricci tensor is not quite right. However, we can make a conserved, local, second rank tensor out of \( R_{\sigma\rho} \) and \( R \), the Ricci scalar, \( R = \sum_\mu R^\mu_{\mu} \). The result is the Einstein tensor,

\[ G_{\sigma\rho} = R_{\sigma\rho} - \frac{1}{2} R g_{\sigma\rho} \quad (11.6) \]

Eq. (11.5) and (11.6) imply the desired conservation law,

\[ \sum_{\gamma} D^\gamma G_{\rho\gamma} = 0 \]

Finally, the Einstein field equation reads,

\[ G_{\alpha\beta} = \frac{8\pi G}{c^4} T_{\alpha\beta} \quad (11.7) \]

This is what we have been seeking. It is the center-piece of the theory. We will investigate its content in the following sections of this chapter.

The Einstein field equation is often written in terms of \( R_{\alpha\beta} \) and \( R \). Take the Einstein field equation Eq. (11.7), write it in terms of the Ricci tensor and Ricci scalar and then fully contract the indices,

\[ \sum_{\alpha\beta} g^{\alpha\beta} R_{\alpha\beta} - \frac{1}{2} \sum_{\alpha\beta} R g^{\alpha\beta} g_{\alpha\beta} = \frac{8\pi G}{c^4} \sum_{\alpha\beta} g^{\alpha\beta} T_{\alpha\beta} \]

Define \( R = \sum_\gamma R^\gamma_\gamma, T = \sum_\gamma T^\gamma_\gamma \) and note that \( \delta^\alpha_\beta = \sum_\gamma g^{\alpha\gamma} g_{\gamma\beta} \) and \( 4 = \sum_\gamma \delta^\gamma_\gamma \). We learn that,

\[ R = -\frac{8\pi G}{c^4} T \]

so the Einstein field equation can be written,
\[ R_{\alpha\beta} = 8\pi G \left( T_{\alpha\beta} - \frac{1}{2} T g_{\alpha\beta} \right) \]  

(11.8)

We should reflect on Eq. (11.7) and (11.8). These equations are the heart of general relativity and represent one of the supreme achievements of science. The discussion here is very efficient, tidy, prim and proper. This is not at all how science works! The creation of general relativity took years of blood sweat and tears, false starts, dead ends and, yes, mistakes. The synthesis of the principles of physics with Riemann’s formulation of differential geometry was something utterly new. It represented the first step in the development of modern field theory. Feynman stated,

“Einstein arrived at his field equations without the help of a developed field theory, and I must admit that I have no idea how he ever guessed at the final result.....I feel as though he had done it while swimming underwater, blindfolded, and with his hands tied behind his back!” [6]

a. Geodesic Equation of Motion as a Consequence of the Field Equations

At this point it appears that General Relativity consists of two coupled systems of differential equations: 1. The Einstein Field Equations for gravity, and 2. The Geodesic Equations for matter. This coupled system indicates that matter generates gravity and gravity determines the paths of the matter that created gravity. It is very significant that actually the geodesic equations are consequences of the Einstein Field equations and need not be postulated independently! Let’s demonstrate this point when the source of gravity is a single particle described by the energy-momentum tensor,

\[ T^{\alpha\beta} = \rho u^\alpha u^\beta \]  

11.9
The Einstein field equation reads \( G_{\alpha\beta} = \frac{8\pi G}{c^4} T_{\alpha\beta} \) and the Einstein tensor is conserved, so

\[
\sum_{\gamma} D^\gamma G_{\rho\gamma} = 0
\]

which implies that,

\[
\sum_{\gamma} D^\gamma T_{\rho\gamma} = 0 \quad 11.10
\]

Substituting Eq. 11.9 into 11.10,

\[
\sum_{\alpha} D_{\alpha} (\rho u^\alpha u^\beta) = \sum_{\alpha} D_{\alpha} (\rho u^\alpha) u^\beta + \rho \sum_{\alpha} u^\alpha D_{\alpha} u^\beta = 0 \quad 11.11
\]

If we contract this equation with \( u_\beta \), we find

\[
c^2 \sum_{\alpha} D_{\alpha} (\rho u^\alpha) + \rho \sum_{\alpha} u^\alpha \sum_{\beta} u_\beta D_{\alpha} u^\beta = 0 \quad 11.12
\]

But the second term vanishes because

\[
D_\beta \left( \sum_{\alpha} u^\alpha u_\alpha \right) = \sum_{\alpha} (D_\beta u^\alpha) u_\alpha + \sum_{\alpha} u^\alpha (D_\beta u_\alpha) = 2 \sum_{\alpha} (D_\beta u^\alpha) u_\alpha = 0
\]

where we observed that \( \sum_{\alpha} u^\alpha u_\alpha = c^2 \) so its derivative vanishes. Substituting back into Eq. 11.12, we have

\[
\sum_{\alpha} D_{\alpha} (\rho u^\alpha) = 0
\]

which is the relativistic conservation law. Finally, substituting back into Eq. 11.11, we have

\[
\sum_{\alpha} u^\alpha D_{\alpha} u^\beta = 0 \quad 11.13
\]

which states that the covariant derivative of the tangent vector along the curve \( x^\alpha(\tau) \) vanishes. But this is just the geodesic equation! To make this point explicit, consider a curve \( x^\alpha(\tau) \). Then its tangent satisfies,
\[
\frac{D u^\beta}{D \tau} = \sum_\alpha u^\alpha D_\alpha u^\beta = 0
\]

where we used Eq. 11.13 in the final step, setting it to zero. Writing this out,

\[
\frac{D u^\beta}{D \tau} = \frac{du^\beta}{d\tau} + \sum_{\rho\sigma} \Gamma^\beta_{\rho\sigma} u^\rho u^\sigma = 0
\]

Or,

\[
\frac{d^2 x^\beta}{d\tau^2} + \sum_{\rho\sigma} \Gamma^\beta_{\rho\sigma} \frac{dx^\rho}{d\tau} \frac{dx^\sigma}{d\tau} = 0
\]

In summary, we have derived the equation of motion of matter, the source of gravity, from the field equation for gravity itself. We see that this was possible because of the non-linear nature of General Relativity. In a linear theory, such as electrodynamics, this is not possible: the Lorentz force law does not follow from Maxwell’s equations. Maxwell’s equations only imply current conservation.

It is interesting that other field theories of high energy physics are also non-linear like General Relativity. For example, Quantum Chromodynamics, the theory of colored quarks and gluons in flat Minkowski space time share this property. The theory is invariant to space time dependent local rotations in color space. The quarks carry color (they comprise a triplet in color), as do the gluons (they comprise an octet in color) which provide the forces between quarks. In General Relativity, all forms of energy-momentum attract one another, so the pure gravity field self-interacts. Similarly, in Quantum Chromodynamics all forms of color interact, so the gluons self-interact as well. An important distinction between General Relativity and Quantum Chromodynamics (among many!), however, is that Quantum Chromodynamics is a quantum field theory, not a classical theory. Unfortunately, no quantum formulation of gravity in four dimensions is known.
b. Linearized Gravity

In the textbook we studied a number of problems where the metric had the form,

\[ g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \]

where \(|h_{\mu\nu}| \ll 1\) in various coordinate systems. This allowed us to linearize General Relativity in those coordinate systems. Applications included the weak field, static non-relativistic limit of the theory as well as the propagation of weak gravity waves far from their sources. These are important applications, so let’s develop this topic from scratch. First a warning: this approach obtains results to leading order in the deviation from the Minkowski metric and it therefore drops the non-linearities that make General Relativity fully self-consistent. We will have to be careful to use this approximation wisely.

The weak field limit of General Relativity can be viewed in two ways: 1. The field \(h_{\mu\nu}\) can be viewed as a field propagating in an ordinary, flat Minkowski space time, or 2. As a weak source of space time curvature. We will consider both interpretations, depending on which one is most compelling.

Let’s begin with a simple but useful technical point: when we work to first order in \(h_{\mu\nu}\), we can raise and lower indices on vectors and tensors which are already first order in \(h_{\mu\nu}\) with just the Minkowski metric \(\eta_{\mu\nu}\). This is an example of the perspective 1 above. This observation simplifies our calculations considerably.

When we study weak fields, we will be interested in two classes of coordinate transformations: Lorentz transformations and infinitesimal coordinate transformations,

\[ g_{\mu\nu} \rightarrow g'_{\mu\nu} = \eta_{\mu\nu} + h'_{\mu\nu} \]
where $|h'_{\mu\nu}| \ll 1$.

Consider first a Lorentz transformation,

$$x' \mu = \sum_{\rho} \Lambda^\mu_{\rho} x^\rho$$

where the $\Lambda^\mu_{\rho}$ are the components of the boosts we studied in special relativity,

$$x'^0 = \gamma(x^0 - vx^1/c)$$
$$x'^1 = \gamma(x^1 - vx^0/c)$$
$$x'^2 = x^2, \quad x'^3 = x^3$$

For a boost in the $x^1 = x$ direction, the metric transforms as,

$$g'_{cd} = \sum_{ab} \frac{\partial x^a}{\partial x'^c} \frac{\partial x^b}{\partial x'^d} g_{ab} = \sum_{ab} \Lambda^a_c \Lambda^b_d (\eta_{ab} + h_{ab}) = \eta_{ab} + \sum_{ab} \Lambda^a_c \Lambda^b_d (h_{ab})$$

where we recalled from special relativity that the Minkowski metric is invariant under boosts. We read off,

$$h'_{cd} = \sum_{ab} \Lambda^a_c \Lambda^b_d (h_{ab})$$

So $h_{ab}$ transforms as a second rank tensor in Minkowski space time. $S'$ is an admissible frame as long as $|h'_{\mu\nu}| \ll 1$.

Now consider infinitesimal general coordinate transformations,

$$x'^\mu = x^\mu + \xi^\mu$$

where $|\xi^n| \sim |h'^{\mu\nu}| \ll 1$. Now let’s calculate how $h^{\mu\nu}$ transforms. First we need the derivatives,
\[
\frac{\partial x'\mu}{\partial x^\nu} = \delta^\mu_\nu + \partial_\nu \xi^\mu, \quad \frac{\partial x^\mu}{\partial x'\nu} = \delta^\mu_\nu - \partial_\nu \xi^\mu
\]

where the equality actually means “correct to first order”. The metric transforms as,

\[
g'_{cd} = \sum_{ab} \frac{\partial x^a}{\partial x'^c} \frac{\partial x^b}{\partial x'^d} g_{ab} = \sum_{ab} (\delta^a_c - \partial_c \xi^a)(\delta^b_d - \partial_d \xi^b)(\eta_{ab} + h_{ab})
\]

\[
= \eta_{cd} + h_{cd} - \partial_c \xi_d - \partial_d \xi_c
\]

where we have raised and lowered indices using \(\eta_{ab}, \xi_c = \sum_d \eta_{cd} \xi^d\), which is correct to first order. We read off the transformation law for \(h_{cd}\),

\[
h'_{cd} = h_{cd} - \partial_c \xi_d - \partial_d \xi_c
\]

Following the terminology of electrodynamics, we refer to this result as a “gauge transformation”: we are still working in the same set of coordinates, but have defined a “new” tensor.

A final technical point: since \(g_{ab} = \eta_{ab} + h_{ab}\), then \(g^{ab} = \eta^{ab} - h^{ab}\). This result guarantees that \(g^{ab}\) is the inverse of \(g_{ab}\): to first order accuracy, \(\sum_a g^{ca} g_{ab} = \delta^c_b\).

Now we come to the heart of the matter: What is the field equation for \(h_{ab}\)? Our task is to linearize the Einstein field equation,

\[
R_{\sigma\rho} - \frac{1}{2} R g_{\sigma\rho} = \frac{8\pi G}{c^4} T_{\alpha\beta}
\]

We need to insert \(g_{\sigma\rho} = \eta_{\sigma\rho} + h_{\sigma\rho}\) into the left hand side of this equation and extract all the first order effects. First the Christoffel symbols become to first order in \(h_{\sigma\rho}\),

\[
\Gamma^\mu_{\rho\beta} = \frac{1}{2} \sum_{\alpha} \eta^{\mu\alpha} \left( \partial_\rho h_{\alpha\alpha} + \partial_\alpha h_{\rho\beta} - \partial_\alpha h_{\rho\beta} \right) = \frac{1}{2} \left( \partial_\rho h^\mu_\beta + \partial_\rho h^\mu_\beta - \partial^\mu h_{\rho\beta} \right)
\]

Next we need the Riemann tensor,
\[ R^\sigma_{\mu\nu\rho} = \partial_\nu \Gamma^\sigma_{\mu\rho} - \partial_\rho \Gamma^\sigma_{\mu\nu} + \cdots \]

where the neglected terms are all second order in \( h_{\sigma\rho} \). Substituting in the linear approximations to the Christoffel symbols produces,

\[ R^\sigma_{\mu\nu\rho} = \frac{1}{2} \left( \partial_\nu \partial_\mu h^\sigma_{\rho} + \partial_\rho \partial_\sigma h_{\mu\nu} - \partial_\nu \partial_\sigma h_{\mu\rho} - \partial_\rho \partial_\mu h^\sigma_{\nu} \right) \]

Then the Ricci tensor becomes,

\[ R_{\mu\nu} = \sum_\sigma R^\sigma_{\mu\nu\sigma} = \frac{1}{2} \left( \partial_\mu \partial_\nu h - \sum_\rho \partial_\nu \partial_\rho h^\rho_{\mu} - \sum_\rho \partial_\rho \partial_\mu h^\rho_{\nu} \right) \]

where \( h = \sum_\rho h^\rho_{\mu} \), the trace of \( h_{\mu\nu} \), and \( \Box = \sum_\rho \partial_\rho \partial^\rho \), is the wave operator, d’Alembertian.

Finally, the Ricci scalar is,

\[ R = \sum_\rho R^\rho_{\mu\nu\rho} = \Box h - \sum_\rho \partial_\rho \partial_\mu h^\rho_{\mu} \]

Collecting everything, the field equation becomes,

\[ \partial_\mu \partial_\nu h + \Box h_{\mu\nu} - \sum_\rho \partial_\nu \partial_\rho h^\rho_{\mu} - \sum_\rho \partial_\rho \partial_\mu h^\rho_{\nu} - \eta_{\mu\nu} \left( \Box h - \sum_\rho \partial_\rho \partial_\mu h^\rho_{\mu} \right) = -\frac{16\pi G}{c^4} T_{\mu\nu} \]

A slight simplification occurs if we define,

\[ \bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} h \quad 11.14 \]

It is useful to invert this expression, Eq. 11.14. First take the trace of Eq. 11.14,
\[ \sum_{\sigma} \tilde{h}_{\sigma} = \sum_{\sigma} h_{\sigma} - \frac{1}{2} \sum_{\sigma} \eta_{\sigma} h \]

which reads, if we define the trace of \( \tilde{h}_{\mu\nu} \) to be \( \bar{h} = \sum_{\rho} \bar{h}_{\rho} \),

\[ \bar{h} = h - \frac{1}{2} \cdot 4 \cdot h = -h \]

Substituting into Eq. 11.14, we have,

\[ h_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} \bar{h} \]

Substituting into the field equation, we find,

\[ \Box \bar{h}_{\mu\nu} + \eta_{\mu\nu} \sum_{\rho} \partial_\rho \partial_\mu \bar{h}^{\rho\mu} - \sum_{\rho} \partial_\nu \partial_\rho \bar{h}_\rho^\mu - \sum_{\rho} \partial_\rho \partial_\mu \bar{h}_\rho^\nu = - \frac{16\pi G}{c^4} T_{\mu\nu} \]

which is better but is still quite messy. But it reminds us of electrodynamics where the wave equation for the vector potential \( A^\mu \) is only simple if we choose the Lorenz gauge, \( \sum_{\mu} \partial_\mu A^\mu = 0 \).

Then the wave equation reads,

\[ \Box A^\mu = \frac{4\pi k}{c^2} J^\mu \]

Can we make a similar choice here,

\[ \sum_{\mu} \partial_\mu h^{\mu\nu} = 0 \quad (?) \]

Recall the gauge transformation,

\[ h'_{\cd} = h_{\cd} - \partial_\cd \xi_\ad - \partial_\ad \xi_\cd \]
The gauge transformation for $\tilde{h}_{\mu\nu}$ follows by substitution into Eq. 11.14,

$$\tilde{h}'_{cd} = h_{cd} - \partial_c \xi_d - \partial_d \xi_c + \eta_{cd} \sum_a \partial_a \xi^a$$

The divergence of this equation becomes,

$$\sum \partial_\rho \tilde{h}'^{\rho\mu} = \sum \partial_\rho h^{\rho\mu} - \square \xi^\mu$$

So, if we choose $\xi^\mu$ to satisfy $\sum \partial_\rho h^{\rho\mu} = \square \xi^\mu$ then $\sum \partial_\rho \tilde{h}'^{\rho\mu} = 0$, and we have a simple wave equation! Dropping the unnecessary prime, it reads,

$$\square \tilde{h}_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu} \tag{11.15}$$

This is the equation we will study to understand gravitational waves generated by a source $T_{\mu\nu}$. It is good as long as $\tilde{h}_{\mu\nu}$ is small. Close to a source we would probably need to use the full Einstein field equation and then match onto the solution to this equation far enough away from the source. We must supplement this wave equation with the definition Eq. 11.14 to obtain $h_{\mu\nu}$ and then find the metric $g_{\mu\nu}$.

This wave equation suggests that the Einstein field equation can be interpreted as a generalization of the Minkowski space time wave equation to curved space time. Since the gravitational field both propagates through space time and curves it, the general field equations are quite challenging. Eq. 11.15 also shows us one of the limitations of the linear approximation we mentioned before: we see that the Lorenz gauge for $\tilde{h}_{\mu\nu}$ implies that $\sum_{\nu} \partial^\nu T_{\mu\nu} = 0$ and the approximation fails to reproduce the geodesic equations for the matter source, as discussed above.

In the model application of linear gravity to gravitational radiation, we will simply model the source and see what waves $h_{\mu\nu}$ result. This is analogous to approaches to radiation problems in
electrodynamics. Of course, the full self-consistent problem can be attached using computer
simulations without linearizing the field equations. Those results reduce to linear gravity far from
the source but they predict the amplitudes of the radiated field quantitatively.

Finally we should check that Eq. 11.15 reduces to Newton’s description of the gravitational
field in the static, weak field, non-relativistic case. In this case \( T_{00} \to c^2 \rho \) and the wave equation
reduces to the Poisson equation,

\[
-\nabla^2 \tilde{h}_{00} = -\frac{16\pi G}{c^2} \rho
\]

So, if we identify \( \tilde{h}_{00} = 4\Phi/c^2 \) we retrieve Newton’s equation for gravitation,

\[
\nabla^2 \Phi = 4\pi G \rho
\]

12. The Schwarzschild Metric and Black Hole

To begin to understand general relativity we need to understand its simplest problem: the
space-time outside a spherical, static mass \( M \). This is the first problem one solves in Newton’s
theory of gravity and everyone knows the solution for the gravitational potential,

\[
\Phi(r) = -\frac{GM}{r}
\]  \hspace{1cm} (12.1)

as we discussed in our introductory remarks. The force felt by a particle of mass \( m \) in the
potential Eq. (12.1) is \( F = -m \nabla \Phi(r) = -\frac{GmM}{r^2} \hat{r} \): it is a static, radially symmetric force which
obeys the inverse square law.

Can we find the metric in general relativity that supplants this result? The metric should be
static and spherically symmetric,
\[ ds^2 = U(r)c^2dt^2 - V(r)dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) \]  \hspace{1cm} (12.2)

in spherical coordinates. The Einstein field equations should determine the functions \( U \) and \( V \) which only depend on the radial coordinate \( r \). This metric should apply outside the mass \( M \). If \( M \) consists of a mass distribution, its spherical symmetry and static nature in the coordinate system chosen in Eq. (12.2) is required. Finally, the metric must solve the Einstein field equations outside \( M \),

\[ R_{\alpha\beta} = 8\pi G \left( T_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} T \right) = 0 \]  \hspace{1cm} (12.3)

and in the non-relativistic limit we should retrieve Eq. (12.1) and Newton’s equation for the motion of a mass \( m \) in this environment, \( \mathbf{F} = -m \nabla \Phi(r) = m \frac{d^2\mathbf{r}}{dt^2} = -\frac{GM}{r^2} \mathbf{r} \).

Eq. (12.3) is deceptively simple. It hides a set of coupled, second order, non-linear differential equations for the components of the metric \( U(r) \) and \( V(r) \). Why second order? The Einstein field equations were constructed with this attribute in mind because we know from experience with Newton’s equations of motion and the equations of electrodynamics that second order differential equations have physical solutions if they are supplemented with sensible initial conditions. Higher order differential equations are full of pathologies, including run-away solutions, non-causal effects, etc. Why non-linear? General relativity is inherently non-linear. No linear superposition principle here! This fact makes analytic problem solving in general relativity much harder than in electrodynamics. The differential equation will be non-linear because of the universal character of gravity: any bit of energy-momentum attracts any other, so any patch of curvature attracts any other patch of curvature. In field theory one says that the theory has “intrinsic self-interactions”. We discuss these effects further in later sections of this chapter, but we will see them in play immediately.
Let’s begin. No tricks, just hard labor. We will illustrate some of the algebra and manipulations involved, but will not display all (!) of it. The student should commit themselves to work along here and the problem set will encourage that.

The metric reads,

\[ g_{00} = U, \quad g_{11} = -V, \quad g_{22} = -r^2, \quad g_{33} = -r^2 \sin^2 \theta \]

We will need its inverse,

\[ g^{00} = \frac{1}{U}, \quad g^{11} = -\frac{1}{V}, \quad g^{22} = -\frac{1}{r^2}, \quad g^{33} = -\frac{1}{r^2 \sin^2 \theta} \]

From these we can calculate the Christoffel symbols,

\[ \Gamma_{\nu \sigma}^{\mu} = \frac{1}{2} \sum_{\lambda} g^{\mu \lambda} \left( \partial_{\sigma} g_{\lambda \nu} + \partial_{\nu} g_{\lambda \sigma} - \partial_{\lambda} g_{\nu \sigma} \right) \]

Many components of \( \Gamma_{\nu \sigma}^{\mu} \) are zero because of the static, spherical character of this problem: time derivatives \( \partial_0 \) vanish identically and derivatives \( \partial_{\phi} \) do also and \( \partial_{\theta} \) is non-zero only if it applies to \( g_{33} \) or \( g^{33} \).

We calculate for \( i, j = 1,2,3 \),

\[ \Gamma_{00}^{0} = \frac{1}{2} g^{00} \left( \partial_0 g_{00} + \partial_0 g_{00} - \partial_0 g_{00} \right) = 0 \]

\[ \Gamma_{01}^{0} = \frac{1}{2} g^{00} \left( \partial_1 g_{00} + \partial_0 g_{01} - \partial_0 g_{01} \right) = \frac{1}{2} g^{00} \partial_1 g_{00} = \frac{\partial_{r} U}{2U} \]

\[ \Gamma_{ij}^{0} = \frac{1}{2} g^{00} \left( \partial_j g_{0i} + \partial_i g_{0j} - \partial_0 g_{ij} \right) = 0 \]
The full set of non-vanishing Christoffel symbols reads, using a convenient short-hand that “primes” indicate differentiation with respect to \( r \), \( U' \equiv \partial_r U \), \( V' \equiv \partial_r V \),

\[
\Gamma^0_{01} = \frac{U'}{2U}, \quad \Gamma^1_{00} = \frac{U'}{2V}, \quad \Gamma^1_{11} = \frac{V'}{2V}, \quad \Gamma^1_{22} = -\frac{r}{V}, \quad \Gamma^1_{33} = -\frac{r \sin^2 \theta}{V}
\]

\[
\Gamma^2_{12} = \frac{1}{r}, \quad \Gamma^2_{33} = -\sin \theta \cos \theta
\]

\[
\Gamma^3_{13} = \frac{1}{r}, \quad \Gamma^3_{23} = \cot \theta
\]

The “real” work begins with the Ricci tensor,

\[
R_{\mu \nu} = \sum_\beta R^\beta_{\mu \nu} = \sum_\beta \partial_\nu \Gamma^\beta_{\mu \beta} - \sum_\beta \partial_\beta \Gamma^\beta_{\mu \nu} + \sum_{\alpha \beta} \Gamma^\alpha_{\mu \beta} \Gamma^\beta_{\nu \alpha} - \sum_{\alpha \beta} \Gamma^\alpha_{\mu \nu} \Gamma^\beta_{\nu \beta}
\]

Writing this expression out reveals that \( R_{\mu \nu} = 0 \) identically if \( \mu \neq \nu \). This preliminary saves us a lot of extra labor, although its verification, while straight-forward, is tedious.

The non-trivial contributions to \( R_{00} \) are,

\[
R_{00} = -\partial_1 \Gamma^1_{00} + \Gamma^0_{01} \Gamma^1_{00} - \Gamma^1_{00} \Gamma^1_{11} - \Gamma^1_{00} \Gamma^2_{12} - \Gamma^1_{00} \Gamma^3_{13} = -\frac{U''}{2V} + \frac{U'}{4V} \left( \frac{U'}{U} + \frac{V'}{V} \right) - \frac{U'}{rV}
\]

Similarly,

\[
R_{11} = \frac{U''}{2U} - \frac{U'}{4U} \left( \frac{U'}{U} + \frac{V'}{V} \right) - \frac{V'}{rV}
\]

\[
R_{22} = r \frac{U'}{2UV} + \frac{1}{V} - \frac{r V'}{2V^2} - 1
\]
\[ R_{33} = \sin^2 \theta R_{22} \]

These are coupled, non-linear, second order differential equations. However, notice that if we multiply \( R_{00} \) by \( V/U \) and add the result to \( R_{11} \), most terms cancel leaving,

\[ \frac{U'}{rU} + \frac{V'}{rV} = 0 \]

Or more simply,

\[ U'V + V'U = 0 \]

which implies that the product \( UV \) is a constant! To determine the constant, imagine that \( r \) becomes large. Then the metric should reduce to the Minkowski metric, so the constant is unity,

\[ U(r)V(r) = 1 \]

Finally, substitute this relation into \( R_{22} = 0 \) and find,

\[ U + rU' = 1 \]

which can be written as,

\[ d(rU)/dr = 1 \]

which integrates to,

\[ rU = r + k \quad \text{or} \quad U = 1 + k/r \]

where \( k \) is a constant. Since \( UV = 1 \), we learn that \( V = (1 + k/r)^{-1} \).

To determine the constant \( k \), imagine the limit of large \( r \). We should obtain the weak gravitational field case that we already analyzed using the Equivalence Principle,
\[ g_{00} \to 1 - \frac{2GM}{c^2r} \]

So, we read off \( k = -\frac{2GM}{c^2} \) and we have the Schwarzschild metric,

\[ ds^2 = \left( 1 - \frac{2GM}{c^2r} \right) c^2 dt^2 - \left( 1 - \frac{2GM}{c^2r} \right)^{-1} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \quad (12.4) \]

The physics in this formula will be the subject of several sections of this chapter. What makes it so exciting? It contains a Black Hole!

There is an ominous singularity in Eq. 12.4: for \( 2GM/rc^2 = 1 \), the proper time,

\[ d\tau = \sqrt{1 - 2GM/rc^2} \, dt, \]

vanishes, and the proper distance, \( dl = dr/\sqrt{1 - 2GM/rc^2} \), diverges. The critical distance,

\[ r_{\text{sch}} = \frac{2GM}{c^2}, \]

is called the Schwarzschild radius. For a typical celestial body such as the Sun, \( r_{\text{sch}} \) is much smaller than the body's actual radius and these singularities are not physical—the expression for \( ds^2 \), Eq. (12.4), is only true outside the massive body, just like Newton's gravitational potential, \( V(r) = -GM/r \), is only true outside the mass \( M \) (\( r_{\text{sch}} \) is about 2,900 meters for the sun whose physical radius is about \( 7 \times 10^5 \) kilometers. For the earth \( r_{\text{sch}} \) is about 0.88 centimeters!). However, there are astrophysical bodies that have collapsed under their own enormous weight and whose radii are less than their Schwarzschild radii. They are called Black Holes, and they provide a physical realization of the peculiar, extreme conditions described by the Schwarzschild metric for strong fields.

The Schwarzschild metric has singularities at \( r = 0 \) and \( r = 2GM/c^2 \). Are either of these singularities properties of the coordinate mesh or are they intrinsic physical properties of the black hole? The Schwarzschild radius is certainly physically significant, but we will learn that it is not a
true singular point. One way to see this is to calculate some components of the Riemann tensor, 
\(R_{\alpha\beta\gamma\delta}\), and some curvature invariants and see that they are free of singularities at the Schwarzschild radius. For example, one can calculate the fully contracted “square” of the Riemann tensor,

\[
\sum_{\alpha\beta\gamma\delta} R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta} = \frac{48 G^2 M^2}{c^4 r^6}
\]

Only the origin is a true singularity. We will find other results below that clarify this finding.

We want to understand the light cones in this space time to see how massive and massless particles propagate in this environment. Recall that in Minkowski space time we can consider left and right moving light rays at each point in the \((ct, x)\) plane. Massive particles at \((ct, x)\) propagate into the future of the light cone at that point. (We leave the \(y\) and \(z\) coordinates implicit and treat the problem as two dimensional) We want to make the analogous constructions for the Schwarzschild metric to understand the physics inside and outside the Schwarzschild radius \(r_s = \frac{2GM}{c^2}\). Consider the radial motion of light rays. A light ray follows the null interval,

\[
d s^2 = 0 = \left(1 - \frac{2GM}{c^2 r}\right) c^2 dt^2 - \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2
\]

So,

\[
\frac{dt}{dr} = \pm \left(1 - \frac{2GM}{c^2 r}\right)^{-1}
\]

To visualize this equation, plot \(t\) against \(r\) in a space-time diagram as shown in Figure 12.1. As we consider \(r\) values closer to \(r_s\), the light cones become narrower.
Fig. 12.1 The light cones in Schwarzschild coordinates as $r$ approaches the Schwarzschild radius.

At large $r$ the slope of the light cone $dt/dr$ approaches $\pm 1$ as in flat Minkowski space-time. However, as the light ray approaches the Schwarzschild radius, a distant observer concludes that the light ray takes longer and longer to make fixed incremental progress toward smaller $r$.

Eq. 12.5 can be integrated to find the paths of outgoing and incoming photons,

\[ ct - r = \frac{2GM}{c^2} \ln \left| \frac{c^2 r}{2GM} - 1 \right| + \text{const}. \quad \text{(outgoing)} \quad 12.6a \]

\[ ct + r = -\frac{2GM}{c^2} \ln \left| \frac{c^2 r}{2GM} - 1 \right| + \text{const.} \quad \text{(incoming)} \quad 12.6b \]

It is instructive to plot these geodesics and see how the light cones are effected by the gravitational field. We are especially interested in the region near the Schwarzschild radius $r_s$ and the singularity at the origin. Fig. 12.2 shows the plot which we want to describe and discuss.
Fig. 12.2  Incoming and outgoing light rays in Schwarzschild coordinates. Future light cones abruptly tilt over 90° at the Schwarzschild radius.

For large $r \gg r_s$, the outgoing and incoming paths which describe the edge of the local light cone propagate at ±45°. These are analogues of left and right moving light rays discussed earlier in the context of special relativity. Massive particles propagate within the future light cone and massless particles (photons and gravitons) propagate on the edges of the future light cone as shown in Fig. 12.2. We see that incoming photons approach $r_s$ but never reach it, as is clear from Eq. 12.5, which shows that $\frac{dt}{dr}$ approaches plus infinity for incoming photons. Similarly, outgoing photons starting at $r$ as $r \to r_s$ become progressively more vertical.
The Schwarzschild radius $r_s$ is a boundary, an abrupt change in behavior of the light cones. As discussed earlier and is clear for Eq. 12.4 when $r$ falls below $r_s$, $t$ becomes a spacelike variable and $r$ becomes a timelike coordinate. In Fig. 12.2 we see the incoming photons propagating in the $-t$ direction and ending on the singularity at $r = 0$. So, although the incoming photon is traveling in the $-t$ direction, it is simply diminishing its spacelike deviation from $r = 0$. Similarly outgoing photons starting at $r < r_s$ propagate in the $+t$ direction and also bend to the left and end on the singularity at $r = 0$. We see from the intersections of the incoming and outgoing photons inside $r_s$ that the light cones have tilted by $90^\circ$ to the left compared to the light cones outside the Schwarzschild radius. The tilted light cones thus direct massive and massless particles inevitably to the singularity at $r = 0$ where they are presumably destroyed by the divergent curvature there. The plot emphasizes that this fate is inescapable because it follows from space time geometry: no matter how hard one tries to avoid the $r = 0$ singularity, if one ever passes through the event horizon from the outside, then one’s path to $r = 0$ is inevitable. Even if the particle is rocket propelled, its future light cone bends toward $r = 0$. It cannot cross the edge of the tilted light cones and it powers into the $r = 0$ singularity where it is destroyed.

The variables $(ct, r, \theta, \phi)$ of the Schwarzschild metric Eq.12.4 are not the best for exploring the region of space time having $r < r_s$. They have a coordinate singularity at $r_s$. In this coordinate system, if we drop a massive particle from rest at large $r$, then it takes an infinite interval of time $t$ to reach $r_s$ and it “never” passes through the horizon. This conclusion is true from the perspective of a stationary observer at large $r$. Her proper time is $t$. By contrast, we will calculate below that the falling observer herself measures a finite proper time $\tau$ when she passes $r_s$. In addition, she doesn’t measure any special “catastrophic” forces as she passes through the horizon. Imagine her surprise when she finds that she can never retrace her steps back to $r > r_s$!
To make these points precise consider a spaceship heading toward $r_{Sch}$. The captain of the ship signals a distant observer by sending light rays every second according to his watch. The distant observer, however, detects these signals more and more slowly. Since $dt/dr$ becomes unbounded at $r_{Sch}$, the distant observer never detects the spaceship passing through $r_{Sch}$. This is a frame dependent result. What happens from the perspective of the captain of the spaceship? To answer this question, let’s study the path of the spaceship. If the spaceship is without power it follows a geodesic,

$$\frac{du_\sigma}{d\tau} - \sum_{\mu \rho} \Gamma^\rho_{\mu \sigma} u^\mu u_\rho = 0$$

Here we have written the geodesic equation for the covariant vector $u_\sigma$ and the Christofel symbols appear here with the “wrong” sign as discussed in lecture 8. The second term can be simplified,

$$\sum_{\mu \rho} \Gamma^\rho_{\mu \sigma} u^\mu u_\rho = \sum_{\mu \rho \alpha} \frac{1}{2} g^{\rho \alpha} (\partial_\mu g_{\alpha \sigma} + \partial_\sigma g_{\mu \alpha} - \partial_\alpha g_{\mu \sigma}) u^\mu u_\rho$$

$$= \sum_{\mu \rho \alpha} \frac{1}{2} (\partial_\mu g_{\alpha \sigma} + \partial_\sigma g_{\mu \alpha} - \partial_\alpha g_{\mu \sigma}) u^\alpha u^\mu = \frac{1}{2} \sum_{\mu \alpha} \partial_\sigma g_{\mu \alpha} u^\mu u^\alpha$$

where we observed cancellation between the first and third terms in the parentheses when the contraction with the symmetric combination $u^\alpha u^\mu$ was taken.

This form of the geodesic is convenient for finding conserved quantities which help to solve for the trajectory taken by the spaceship. If there is a direction for which $\partial_\sigma g_{\mu \alpha} = 0$, then $u_\sigma$ is conserved. Note that if $\sigma = 0$, then $\partial_0 g_{\mu \alpha} = 0$ because the Schwarzschild metric is time independent. In addition, $\partial_\varphi g_{\mu \alpha} = 0$ because the metric has no $\varphi$ dependence, it is unchanged by rotations around the $z$ axis. Call the first constant $E = u_0$ and the second one, which we will identify as the angular momentum per mass about the $z$ axis, $L = -u_\varphi$. What about motion in the $\theta$ direction? If we orient the coordinate system so that $\theta = \pi/2$, then $u_\theta = d\theta/d\tau = 0$ because
\[ du_\theta / d\tau - \frac{1}{2} \partial_\theta g_{\phi \phi} u^\phi u^\phi = 0, \quad g_{\phi \phi} = -r^2 \sin^2 \theta \quad \text{and} \quad \partial_\theta g_{\phi \phi} = -2r^2 \sin \theta \cos \theta \]

which vanishes at \( \theta = \pi / 2 \). So, if we orient the coordinate system so that initially \( \theta = \pi / 2 \), then \( \theta \) remains there. So, as expected from the symmetry of the problem, orbits in the Schwarzschild metric are planar.

We will need the components of the four velocity with upper indices,

\[ u^0 = g^{00} u_0 = \left( 1 - \frac{2GM}{c^2 r} \right)^{-1} E, \quad u^r = \frac{dr}{cd\tau}, \quad u^\theta = 0, \quad u^\phi = g^{\phi \phi} u_\phi = \frac{1}{r^2} L \]

Now we can get the “first integral” of the equations of motion by considering the invariant length of \( u^\mu = dx^\mu / d\tau \). This length is conserved along the geodesic, \( u \cdot u = \sum_\mu u^\mu u_\mu = c^2 \) for the spaceship. Why is \( u \cdot u = c^2 \)? This is a scalar quantity which can be evaluated in any convenient frame. Choose a free falling frame where the rules of special relativity apply. There we can choose the rest frame of the space ship and the result is immediate. Writing this expression out in terms of the components we have computed,

\[ E^2 \left( 1 - \frac{2GM}{c^2 r} \right)^{-1} - \left( 1 - \frac{2GM}{c^2 r} \right)^{-1} \left( \frac{dr}{cd\tau} \right)^2 - \frac{L^2}{r^2} = c^2 \]

which we can manipulate into a more standardized form,

\[ \frac{1}{2} \left( \frac{dr}{cd\tau} \right)^2 + \frac{1}{2} \left( 1 - \frac{2GM}{c^2 r} \right) \left( \frac{L^2}{r^2} + c^2 \right) = \frac{1}{2} E^2 \quad (12.6) \]

This expression is useful to solve for orbits in the vicinity of the mass \( M \). Recognize that this is the general relativity version of the Kepler problem of Newtonian mechanics. We will study Eq. (12.6) further in the next section of this chapter but here we want to specialize to purely radial motion, \( L = 0 \) with the spaceship starting at rest at some intial radius, \( r = R \). Applying these initial conditions to Eq. (12.6), we find \( E \) in terms of \( R \), \( E^2 / c^2 = \left( 1 - \frac{2GM}{c^2 R} \right) \). Now Eq. (12.6) becomes,

\[ \frac{1}{2} \left( \frac{dr}{cd\tau} \right)^2 + \frac{1}{2} \left( \frac{2GM}{c^2 R} - \frac{2GM}{c^2 r} \right) = 0 \]

which predicts the proper time it takes the spaceship to fall into the origin,
\[
\tau = \int_0^R \frac{dr}{2GM - \frac{2GM}{R}} = \frac{\pi}{2} R \sqrt{\frac{R}{2GM}} \tag{12.7}
\]

where we used integral tables in the last step.

We learn two things from these considerations. First, according to the captain of the spaceship a finite amount of time passes on her wrist watch before she reaches the origin. The Schwarzschild radius does not play any special role in the operation of her wristwatch. This is in marked contrast with the measurements made by a distant observer who watches the spaceship begin its free fall at \( R > r_{\text{Sch}} \) and must wait forever for the spaceship to reach the Schwarzschild radius. And second, the result Eq. (12.7) and the energy equation Eq. (12.6) bear an uncanny resemblance to the same problem in Newtonian mechanics. In Newton’s world energy conservation reads,

\[
E = \frac{1}{2} m v^2 - \frac{GMm}{r} \tag{12.8}
\]

where \( m \) is the mass of the spaceship. If the spaceship begins at rest at \( r = R \), then \( E = -\frac{GMm}{R} \), the motion is purely radial, \( v = \frac{dr}{dt} \), and the energy conservation equation becomes,

\[
\frac{dr}{dt} = \sqrt{2GM} \left( \frac{1}{r} - \frac{1}{R} \right)
\]

Which can be flipped over and integrated,

\[
t = \int_0^R \frac{dr}{\sqrt{2GM} \left( \frac{1}{r} - \frac{1}{R} \right)} = \frac{\pi}{2} R \sqrt{\frac{R}{2GM}}
\]

So we have the same answer as general relativity with the understanding that Newton’s time \( t \) is universal and in relativity the time that was calculated was the proper time of the captain of the spaceship.
Now we face a puzzle! In Newton’s world, Eq. (12.8), the velocity $v$ is unrestricted by the speed of light: for fixed $E$, $v$ becomes as large as you like as $r$ approaches zero. Looking back at our general relativity discussion, we see that $dr/d\tau$ follows the same systematics! Does this mean that the spaceship can travel faster than the speed of light inside a black hole? But notice something special about $dr/d\tau$: it is a “mixed” quantity, $r$ is a coordinate marker and $\tau$ is a proper time. It is not the velocity of an object as measured by an observer in an inertial frame. So we don’t have a “manifest” contradiction here, but the matter deserves more thought. That can be found in reference [7].

Now let’s return to the Schwarzschild metric and try to understand space-time better inside the Schwarzschild radius $r_s$. We have seen that the Schwarzschild coordinates $t$ and $r$ are not suitable for this: there is a coordinate singularity at $r_s$ and there is no singularity in other frames. Let’s look for coordinates where the singularity at $r_s$ is removed. This should be possible because the curvature is well-behaved there and an infalling space ship does not notice any dire dynamics when it passes through the Schwarzschild radius. First, reconsider the path of a radially infalling photon described in Schwarzschild coordinates,

$$c t = -r - r_s \ln|r/r_s - 1| + \text{const.}$$

Consider the null coordinate [8],

$$c t_+ = c t + r + r_s \ln|r/r_s - 1|$$

This variable is constant along the path of an infalling photon and should be better behaved. We see that

$$c d t_+ = c\, dt + \frac{r}{r - r_s}\, dr$$

and we can write the metric in terms of $(c t_+, r, \theta, \phi)$,

$$ds^2 = (1 - r_s/r)c^2 dt_+^2 - 2c dt_+ dr - r^2(d\theta^2 + \sin^2 \theta\, d\phi^2)$$
In these variable the metric is regular at $r_s$ and we can consider all values of $r$, $0 < r < \infty$. The radial null geodesics satisfy,

$$(1 - r_s/r)\left(\frac{cdt_+}{dr}\right)^2 - 2\left(\frac{cdt_+}{dr}\right) = 0$$

which has the incoming solution,

$$\frac{cdt_+}{dr} = 0 \quad \text{or} \quad ct_+ = \text{const}$$

And outgoing solutions,

$$\frac{cdt_+}{dr} = 2(1 - r_s/r)^{-1} \quad \text{or} \quad ct_+ = 2r + 2r_s \ln|r/r_s - 1| + \text{const}$$

It is also convenient to use a timelike variable $t'$ instead of $t_+$

$$ct' = ct_+ - r = ct + r_s \ln|r/r_s - 1|$$

Then the metric reads [8],

$$ds^2 = (1 - r_s/r)c^2 dt'^2 - \frac{2cr_s}{r} dt'dr - (1 + r_s/r)dr^2 - r^2(d\theta^2 + \sin^2 \theta d\varphi^2)$$

Now the incoming photon paths are,

$$ct' = -r + \text{const}$$

And the outgoing are,

$$ct' = r + 2r_s \ln|r/r_s - 1| + \text{const}$$

These are plotted in Fig. 12.3.
Fig. 12.3   Incoming and outgoing light rays plotted in (advanced) Eddington-Finkelstein coordinates. There is no singularity at $r_s$. Incoming particles pass through it smoothly. Light cones tilt over gradually toward the origin as $r$ ranges from $r_s$ to the origin. No particles can propagate from inside $r_s$ to the outside.

where we see that the future light cones bend over at the Schwarzschild radius and lead all infalling particles, massless and massive, to $r = 0$ where the singular curvature destroys them. Once a particle has reached $r \leq r_s$, then its future includes the $r = 0$ singularity. No particle starting at $r < r_s$ can escape to the outer region, $r > r_s$. The Schwarzschild radius is both an “event horizon” and a surface of “infinite redshift”.

92
The reader may want to consult references [4] and [9] for additional discussion and graphics. We learn that the surface given by the Schwarzschild radius is a one way membrane – once a light ray or massive particle passes through it from the outside, it can never return. Such a surface is called an “event horizon”. A distant observer can not perceive anything within $r_{Sch}$. Note that an event horizon is a global structure, not local: the spaceship passed from $r > r_{Sch}$ to $r < r_{Sch}$ without trauma, but she can never retrace her steps. Since this behavior is a matter of space-time geometry, it has no compelling analog in Newton’s world. Once the spaceship passes through the event horizon, then $c dt'/dr < 0$ and the spaceship relentlessly falls to the origin even if it is under its own rocket power. In fact, for a “typical” black hole, the spaceship and the astronauts inside it will be ripped apart by gravitational tidal “forces” before it reaches the surface of the black hole. Those forces will stretch and squeeze it mercilessly. We know how to estimate these forces from our discussion of tidal forces in lecture 3. Recall that if an astronaut is falling into a black hole feet first, that unfortunate person feels a tidal acceleration across his length of,

$$a = \frac{2GMd}{R^3}$$

where $R$ is his distance to the center of the mass $M$, $d$ is his height and $G$ is Newton’s constant, $6.67 \times 10^{-8} cm^3 gm^{-1} sec^{-2}$. Let’s suppose that the black hole has a mass of the sun ($1.9 \times 10^{33}$ grams), the astronaut is 2 meters tall and $R$ is 100,000 meters. Then the acceleration difference across his length is,

$$2 \times (6.67 \times 10^{-8}) \times (1.9 \times 10^{33}) \times (200)/(1.0 \times 10^7)^3 \approx 51 \times 10^6 cm/sec^2.$$ 

Compare this number with the acceleration of gravity at earth’s surface, 980 cm/sec$^2$. Clearly the tidal forces would stretch and squeeze the astronaut out of existence long before reaching the black hole’s surface.
There are more than one hundred confirmed black holes among the celestial bodies presently under observation. Astrophysicists predict that they are created by the gravitational collapse of very massive stars that run out of fuel late in their life cycles. Binary black holes have also been observed. In 2015 gravitational waves were first observed and their source proved to be a binary pair of black holes in the process of colliding and merging, as we will discuss in lecture 15 below.

The interested student can learn more about black holes in a book [7] written at the same level as this lecture series. The book also has a full bibliography and looks at recent astrophysical observations and data.

The modern study of black holes, their relation to quantum physics, and unification, a field pioneered by Stephen Hawking, is very stimulating because it places all of our physical laws in an extreme environment and challenges their self-consistency and our understanding of them. When you learn the elements of quantum mechanics and reconsider the properties of black holes, you will learn about Hawking radiation, the fact that black holes are, in fact, hot and radiate electromagnetic waves. They are not really black at all! The unification of gravity and quantum mechanics is an active field of research to which the researchers of the high energy physics community contribute through their investigations in super-strings, a framework for a theory of everything—electricity and magnetism, radioactivity (weak interactions), nuclear physics (strong interactions), and gravity.

13. Circular Orbital Motion in the Schwarzschild metric.

There are some striking differences between the Newtonian Kepler problem and orbital motion in the Schwarzschild metric that we should appreciate [7]. We derived in the previous lecture the crucial orbital equation,

\[
\frac{1}{2} \left( \frac{dr}{c dt} \right)^2 + \frac{1}{2} \left( 1 - \frac{2GM}{c^2 r} \right) \left( \frac{L^2}{r^2} + c^2 \right) = \frac{1}{2} E^2
\]

(13.1)
We can write this expression in the form of a one dimensional equation of motion in the radial coordinate,

\[
\frac{1}{2} \left( \frac{dr}{dt} \right)^2 + V_{\text{eff}}(r) = \frac{1}{2} E^2
\]  

(13.2a)

where,

\[
V_{\text{eff}}(r) = \frac{c^2}{2} - \frac{GM}{r} + \frac{L^2}{2r^2} - \frac{GML^2}{c^3r^3}
\]  

(13.2b)

Compare this equation to the corresponding equation in Newtonian mechanics,

\[
\frac{1}{2} \left( \frac{dr}{dt} \right)^2 + V_{\text{eff}}^N(r) = E
\]  

(13.3a)

where,

\[
V_{\text{eff}}^N(r) = -\frac{GM}{r} + \frac{L^2}{2r^2}
\]  

(13.3b)

Special effects in general relativity come from the last term in Eq. (13.2b) which is a relativistic effect, varying as $1/c^2$, that makes the origin more attractive than the centrifugal barrier, $L^2/2r^2$, makes it repulsive! In Figure 13.1 we compare $V_{\text{eff}}^N(r)$ for Newtonian mechanics with $V_{\text{eff}}(r)$ of general relativity.
Fig. 13.1 The potentials of General Relativity and Newtonian mechanics. The non-linearities of General Relativity create the $-r^{-3}$ attraction near the origin.

The striking difference is that for $L^2 \neq 0$, $V_{\text{eff}}(r)$ diverges as $-r^{-3}$ near the origin! This term makes it impossible to have stable circular orbits for small $r$. We can search for stable circular orbits in $V_{\text{eff}}(r)$ by looking for solutions to the two conditions,

$$\frac{dV_{\text{eff}}}{dr} = 0, \quad \frac{d^2V_{\text{eff}}}{dr^2} > 0$$

The first requirement gives,
\[
\frac{dV_{\text{eff}}}{dr} = 0 = \frac{GMr^2/c^2 - L^2r + 3GML^2/c^2}{r^4}
\]
which states how \(L^2\) depends on the radius of the orbit,

\[
L^2 = GMr \left(1 - \frac{3GM}{c^2r}\right)^{-1}
\]

(13.4)

So, we can have circular orbits only for \(r > 3GM/c^2\). In addition, since \(dr/d\tau = 0\) at a circular orbit, \(E^2 = 2V_{\text{eff}}(r)\) there and,

\[
E^2 = \left(1 - \frac{2GM}{c^2r}\right)^2 \left(1 - \frac{3GM}{c^2r}\right)^{-1}
\]

(13.5)

there as well. Finally, we can solve the expression Eq. (13.4) for \(r\),

\[
r_{\pm} = \frac{L^2 \pm \sqrt{L^4 - 12G^2L^2M^2/c^4}}{2GM/c^2}
\]

(13.6)

We learn that if \(L^2 < 12G^2M^2/c^4\), then there are no circular orbits. In addition, one can check that \(\frac{d^2V_{\text{eff}}}{dr^2} > 0\) for \(r_+\) and \(\frac{d^2V_{\text{eff}}}{dr^2} < 0\) for \(r_-\). But since \(L^2\) must be greater than \(12G^2M^2/c^4\) for stable orbits to exist, it follows from Eq. (13.6) that \(r_+ > 6GM/c^2\).

We must understand the term \(- \frac{GML^2}{c^2r^3}\) in \(V_{\text{eff}}\). It changes the Newtonian effective potential qualitatively near the origin and the extra attraction overwhelms the centrifugal barrier term which is so important in many physics problems ranging from planetary motion to quantum mechanics. The modifications to the Newton-Kepler predictions of planetary motion are dramatic even for our solar system where there are no exotic objects like black holes and the new term is very, very small compared to the other terms in \(V_{\text{eff}}\). One can check that for planetary problems the radial coordinate is always much, much larger than \(r_{\text{Sch}}\) so we are always working in the “weak” gravity regime where we expect Newtonian predictions to be very accurate. In fact the new term is a tiny perturbation but it makes distinct and measureable modifications to the predictions of Newtonian mechanics. The reason is interesting. Recall that Newtonian mechanics predicts that the orbits of
masses in a $1/r$ potential are perfect ellipses, closed periodic orbits. This is a special feature, a “dynamical symmetry”, of the $1/r$ potential of Newton’s law of gravity. The only other potential which produces closed, periodic orbits is the harmonic oscillator, a potential which grows as $r^2$ and isn’t relevant here. The point is that if there are any perturbations to the $1/r$ potential of Newton’s law of gravity, then the planetary equations lose their special symmetry and would not produce closed, periodic ellipses. Instead, the orbits would precess slightly year to year. Since the very slow precession of an orbit accumulates over time, the effect proves to be measurable using data obtained from observations employing conventional telescopes! The new term in $V_{eff}$ did, in fact, address a long standing discrepancy in planetary science: the other planets which orbit the sun perturb the orbit of Mercury and also cause it to precess. However, when these Newtonian effects were calculated, only 93% of the observed precession of Mercury was accounted for. The last 7%, 43 seconds of arc per century, remained a mystery for over 200 hundred years. In fact, the general relativity prediction for the precession rate of Mercury is almost exactly 43 seconds of arc per century! This was a great success for general relativity in its early days!

But why does general relativity predict the new term, $-\frac{GML^2}{c^2r^3}$? We can motivate it by writing it in a more suggestive form,

$$\left( - \frac{GM}{r} \right) \left( \frac{L^2/r^2}{c^2} \right)$$

(13.7)

Here $L^2/r^2$ is the energy due to the angular motion of the object and dividing by $c^2$ gives its mass equivalent. (A word about units here: L was defined as angular momentum per mass. So $L^2/r^2$ behaves as transverse velocity squared. Multiplying Eq. (13.2) by the particle’s mass restores energy units.) The first factor in Eq. (13.7) is the gravitational potential generated by the mass $M$, so the equation records the fact that all energies attract one another in a universal fashion. Newtonian mechanics misses this term because in Newton’s world only masses attract one another,
not other forms of energy. In Einstein’s world energy and mass and momentum are all unified into a single four vector, so the principle of covariance insists that they all contribute equally to gravitational attraction. Since the new term is a long range force, it falls off as the third power of the distance, it is important in planetary physics where gravitational effects are all weak. But it is also highly singular as \( r \to 0 \), so in the vicinity of black holes it can dominate over the more familiar centrifugal repulsion potential and produce new effects.

The new term illustrates that in general relativity all forms of energy attract one another. For example, the energy carried by the gravitational field itself attracts all other forms of energy. This makes the theory highly non-linear, as we have already seen in several applications, most explicitly in the Schwarzschild metric. This fact also shows up in attempts to make a quantum theory of gravity where gravitons, the analog to photons of electrodynamics, must interact with themselves. These interactions appear to generate new interactions which are very singular at short distances and lead to pathologies. Although classical general relativity is a theory which is full of triumphs, the search for a quantum version continues unrewarded.

Now let’s return to the problem at hand and consider the radial “in falling” of the particle with angular momentum. Consider Eq. (13.1). We want to calculate the proper time for the free radial motion by solving that equation for \( dt \), generalizing the discussion in Eq. (13.4),

\[
\int c\,dt = \int_{r_0}^{r_1} dr \left( E^2 - \left( 1 - \frac{2GM}{c^2 r} \right) \left( c^2 + \frac{L^2}{r^2} \right) \right)^{-1/2}
\]

The new term, \( \frac{GM L^2}{c^2 r^3} \), enters the denominator with a positive sign indicating that it tends to decrease the proper time needed to fall between \( r_1 \) and \( r_0 \). The centrifugal barrier, \( \frac{L^2}{r^2} \), enters with a minus sign and works in the other direction, as expected. However, if \( r < r_{Sch} \), then the new term
dominates and *increasing* the angular momentum *decreases* the proper time, the lifetime of the “in falling” object!

**14. Relativistic Tidal Forces**

It is interesting to generalize the discussion of tidal forces from Newtonian mechanics to general relativity. We will see that the Riemann tensor replaces the second derivatives of the gravitational potential in the equation of motion for the difference of the positions of two nearby free falling objects. This shows us that the curvature of space-time is the general relativistic version of tidal forces. Tidal forces are more fundamental than one might have thought! Let’s understand them in general relativity.

Consider the position of a freely falling object $x^\alpha(\tau)$. Its trajectory is given by the geodesic,

$$\frac{d^2 x^\alpha}{d\tau^2} + \sum_{\mu\nu} \Gamma^\alpha_{\mu\nu}(x) u^\mu u^\nu = 0$$  \hspace{1cm} (14.1a)

where $u^\mu = dx^\mu/d\tau$. A nearby freely falling object follows the trajectory $\tilde{x}^\alpha(\tau)$,

$$\frac{d^2 \tilde{x}^\alpha}{d\tau^2} + \sum_{\mu\nu} \Gamma^\alpha_{\mu\nu}(\tilde{x}) \tilde{u}^\mu \tilde{u}^\nu = 0$$  \hspace{1cm} (14.1b)

where we choose the same proper time variable to parametrize this nearby trajectory. Following the discussion of Newtonian tidal forces, the two trajectories are infinitesimally close together at $\tau = 0$ and are parallel as well. We are interested in the time evolution of the difference $e^\alpha(\tau) = \tilde{x}^\alpha(\tau) - x^\alpha(\tau)$. First consider Eq. (14.1b) and Taylor expand the Christoffel symbol about the position of the first particle $x^\alpha(\tau)$,
\[ \Gamma_\mu^\alpha(x^\alpha + \epsilon^\alpha) = \Gamma_\mu^\alpha(x^\alpha) + \sum_\alpha \epsilon^\alpha \partial_\alpha \Gamma_\mu^\alpha + O(\epsilon^2) \]

Substituting into Eq. (14.1b) and keeping terms of first order in \( \epsilon^\alpha \),

\[ \frac{d^2(x^\alpha + \epsilon^\alpha)}{d\tau^2} + \sum_{\mu\nu} \left( \Gamma_\mu^\alpha(x^\alpha) + \sum_\alpha \epsilon^\alpha \partial_\alpha \Gamma_\mu^\alpha + O(\epsilon^2) \right) \frac{d(x^\alpha + \epsilon^\alpha)}{d\tau} \frac{d(x^\alpha + \epsilon^\alpha)}{d\tau} = 0 \]

\[ \frac{d^2x^\alpha}{d\tau^2} + \frac{d^2\epsilon^\alpha}{d\tau^2} + \sum_{\mu\nu} \Gamma_\mu^\alpha u^\mu u^\nu + \sum_{\mu\nu} \epsilon^\alpha \partial_\alpha \Gamma_\mu^\alpha u^\mu u^\nu + 2 \sum_{\mu\nu} \Gamma_\mu^\alpha u^\mu \frac{d\epsilon^\nu}{d\tau} \approx 0 \]

Subtracting Eq. (14.1a) we have,

\[ \frac{d^2\epsilon^\alpha}{d\tau^2} + \sum_{\mu\nu\sigma} \epsilon^\sigma \partial_\sigma \Gamma_\mu^\alpha u^\mu u^\nu + 2 \sum_{\mu\nu} \Gamma_\mu^\alpha u^\mu \frac{d\epsilon^\nu}{d\tau} \approx 0 \] (14.2)

Next we need the evolution of the four vector \( \epsilon^\alpha(\tau) \) itself. From lecture 7, Eq. 7.6,

\[ \frac{d\epsilon^\alpha}{d\tau} = \frac{d\epsilon^\alpha}{d\tau} + \sum_{\mu\nu} \Gamma_\mu^\alpha u^\mu \epsilon^\nu \]

And,

\[ \frac{D^2\epsilon^\alpha}{D\tau^2} = \frac{d}{d\tau} \left( \frac{d\epsilon^\alpha}{d\tau} + \sum_{\beta\gamma} \Gamma_\beta^\alpha u^\beta \epsilon^\gamma \right) + \sum_{\delta\eta} \Gamma_\delta^\eta u^\delta \left( \frac{d\epsilon^\eta}{d\tau} + \sum_{\beta\gamma} \Gamma_\beta^\eta u^\beta \epsilon^\gamma \right) \]

\[ \frac{D^2\epsilon^\alpha}{D\tau^2} = \frac{d^2\epsilon^\alpha}{d\tau^2} + 2 \sum_{\beta\gamma} \Gamma_\beta^\gamma u^\beta \frac{d\epsilon^\gamma}{d\tau} + \sum_{\beta\gamma\delta} \partial_\delta \Gamma_\beta^\gamma u^\delta u^\beta \epsilon^\gamma + \sum_{\beta\gamma} \Gamma_\beta^\alpha \frac{du^\beta}{d\tau} \epsilon^\gamma + \sum_{\beta\gamma\delta\eta} \Gamma_\delta^\gamma \Gamma_\beta^\eta u^\delta \epsilon^\gamma \]

where we used,

\[ \frac{d\Gamma_\beta^\gamma}{d\tau} = \sum_{\delta} \frac{\partial \Gamma_\beta^\gamma}{\partial x^\delta} \frac{dx^\delta}{d\tau} = \sum_{\delta} \frac{\partial \Gamma_\beta^\gamma}{\partial x^\delta} u^\delta \]
Finally we eliminate $\frac{d^2 e^a}{d\tau^2}$ from the expression for $\frac{D^2 e^a}{D\tau^2}$ by using Eq. (14.2) and we use Eq. (14.1a) to eliminate $\frac{d\mu^\beta}{d\tau} = \frac{d^2 x^\beta}{d\tau^2}$. Now we have, noting that the terms involving $de^a/d\tau$ cancel,

$$\frac{D^2 e^a}{D\tau^2} = -\sum_{\beta\gamma\delta} \partial_\delta \Gamma^\alpha_{\beta\gamma} u^\beta u^\gamma \epsilon^\delta + \sum_{\beta\gamma\delta} \partial_\delta \Gamma^\alpha_{\beta\gamma} u^\beta u^\delta \epsilon^\gamma - \sum_{\beta\gamma\delta\eta} \Gamma^\alpha_{\beta\gamma} \Gamma^\beta_{\delta\eta} u^\delta u^\eta \epsilon^\gamma + \sum_{\beta\gamma\delta\eta} \Gamma^\alpha_{\delta\eta} \Gamma^\eta_{\beta\gamma} u^\beta u^\delta \epsilon^\gamma$$

Relabelling indices,

$$\frac{D^2 e^a}{D\tau^2} = -\left(\sum_{\beta\gamma\delta} \partial_\gamma \Gamma^\alpha_{\beta\delta} - \sum_{\beta\gamma\delta} \partial_\delta \Gamma^\alpha_{\beta\gamma} + \sum_{\beta\gamma\delta\eta} \Gamma^\alpha_{\gamma\eta} \Gamma^\beta_{\delta\eta} - \sum_{\beta\gamma\delta\eta} \Gamma^\alpha_{\delta\eta} \Gamma^\eta_{\beta\gamma}\right) u^\beta u^\delta \epsilon^\gamma$$

Recognize the Riemann tensor here,

$$\frac{D^2 e^a}{D\tau^2} = -\sum_{\beta\gamma\delta} R^\alpha_{\beta\gamma\delta} u^\beta u^\delta \epsilon^\gamma + O(\epsilon^2) \quad (14.3)$$

So, we learn that to first order in the small four vector difference $e^a(\tau)$, its evolution is determined by the curvature of spacetime. This result applies to strong gravity which tests the limits of general relativity. Excellent result, Eq. (14.3)! Compare it to the analogous result in classical differential geometry, the Jacobi equation, which shows that the intrinsic Gaussian curvature determines the evolution of the distance between two nearby geodesics.

In Appendix G of the textbook we use Eq. (14.3) to analyze the detection of gravitational waves.

Let’s compare this result, Eq. (14.3), to Newtonian tidal forces. We should have agreement between general relativity and Newton’s theory if we limit our attention to: 1. Weak gravitational fields, and 2. Non-relativistic motion so $u^\mu = (c\, dt/d\tau, dx/d\tau, dy/d\tau, dz/d\tau) \approx c(1,0,0,0)$. Then,
\[ \frac{d^2 \epsilon^i}{d\tau^2} \approx \frac{d^2 e^i}{c^2 dt^2} \approx -c^2 \sum_k R^i_{0k0} \epsilon^k \] (14.4)

where we used the fact that \( R^i_{000} \) vanishes identically and the sum over \( k \) is from 1 to 3. But the Newtonian prediction was,

\[ \frac{d^2 \epsilon}{dt^2} = -\nabla (\epsilon \cdot \nabla \Phi) \]

Comparing to Eq. (14.4) we have, \( c^2 R^i_{0k0} = \nabla_i \nabla^k \Phi \), so

\[ \nabla^2 \Phi = c^4 \sum_k R^k_{0k0} = c^4 R_{00} = c^4 \frac{8\pi G}{c^4} \left( T_{00} - \frac{1}{2} g_{00} T \right) \]

where we identified the Ricci tensor, \( R_{\alpha\beta} = \sum_{\gamma} R^\gamma_{\alpha\gamma\beta} \) and used the Einstein field equation,

\[ R_{\alpha\beta} = \frac{8\pi G}{c^4} \left( T_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} T \right) . \]

For a static mass density \( \rho \), we saw in Section 11.8 that \( T_{00} = c^2 \rho \).

Then \( T = \sum_{\gamma} T^\gamma_{0\gamma} = T^0_0 = +c^2 \rho \), so \( T_{00} - \frac{1}{2} g_{00} T = c^2 \rho - \frac{1}{2} (1)(+\rho) = \frac{1}{2} \rho \). Collecting everything,

\[ \nabla^2 \Phi = \frac{1}{2} \frac{8\pi G \rho}{c^4} = 4\pi G \rho \]

which is the differential form of Newton’s law of gravity. All is well.

**15. The Discovery of Gravitational Waves**

Gravitational radiation was predicted by Einstein in 1916. It is a crucial consequence of general relativity because its existence does for gravity what electromagnetic radiation did for electromagnetism: it establishes the dynamical field theoretic character of the gravitational force. As you are discovering in the problem sets, gravity is a much weaker force than the electromagnetic force. This caused the experimental discovery of gravitational waves to take a century of hard work!
However, on February 11, 2016, the LIGO collaboration announced the first observation of gravitational waves in a very exciting webcast. The critical observation occurred five months earlier on September 14, 2015 and several months of data analysis were required to understand the result in sufficient detail to justify publication. The data matched the predictions of general relativity if the source of the waves was the inward spiral and merger of a pair of black holes and subsequent “ringdown” of the resulting single black hole. The observation demonstrated the existence of binary stellar-mass black hole systems and constituted the first observation of a binary black hole merger. The data could be fit extraordinarily well by the hypothesis that two black holes of masses 29 and 36 solar masses merged into one black hole at a distance of about 1.3 billion light years from earth. During the final fraction of a second of the merger, it released more than 50 times the power of all the stars in the observable universe combined. The signal increased in frequency from 35 to 250 Hz (cycles per second) over 10 cycles (5 orbits of the pair of black holes) as it rose in strength for a period of 0.2 second. The mass of the new black hole was 62 solar masses, so the energy equivalent of three solar masses was emitted as gravitational waves in this catastrophic event. The signal was detected at both sites of the LIGO experiment, one in Hanford, WA and another in Livingston, LA, with a time difference of 7 milliseconds between them. It was inferred that the source, the merging binary black hole system, was 1.3 billion light years away in the general direction of the Magellanic Clouds in the Southern Celestial Hemisphere. What a historic moment!

This achievement was the result of decades of work by the LIGO collaboration which dated back to 1984, 31 years before fruition. LIGO stands for “Laser Interferometer Gravitational Wave Observatory” and the generation of the detectors that were successful in 2015 constitute the “advanced LIGO” version of the upgraded experiment. In fact, the upgraded
detectors (laser interferometers) had been operating for just two days before the discovery occurred. Each LIGO detector, which are 1,865 miles apart, are L-shaped detectors which have two arms at right angles to each other and are 2.5 miles in length from a central laboratory building. Lasers are beamed down each arm and are bounced back by free standing suspended mirrors many times to act as extraordinarily precise measuring devices. In fact the lasers allow the length of the arms to be measured to a precision of $10^{-4}$ the width of a single proton. This accuracy is required to measure the small scale of the effects imparted by a passing gravitational wave. The two LIGO detectors act as a check on one another. It is an amazing engineering feat to detect such tiny signals in an otherwise noisey world!

The basis of each LIGO detector are two perpendicular interferometers, as shown in Figure 15.1.

![Image of LIGO detector](image)

**Fig. 15.1** The arms of the LIGO detector.

These devices merge two light rays which creates an interference pattern. If the
peaks of the two light rays overlap, they combine to form a larger peak but when a valley of one light ray overlaps with the peak of the other ray, the two rays cancel out. The laser beams can be arranged to cancel each other out in ordinary operations. Then when a gravitational wave passes through the facility, it would stretch one arm and compress the perpendicular one (we will derive this crucial property of gravity waves below), the exact cancellation would be destroyed and some light would reach a sensitive photodetector. The responses of the photodetector then measure the character of the gravity wave, its amplitude and frequency. As mentioned above, the upgraded detector of advanced LIGO is sensitive to amplitudes as small as $10^{-4}$ the width of a single proton (this requires lasers with sufficient power and accuracy to make coherent beams of light which can travel more than one hundred times up and down an arm of the facility before combining with the laser beam traveling down the other arm) and to be sensitive to gravity waves with frequencies ranging from 75 to 500 Hz. We shall see that this frequency range matches the frequencies of the expected sources. It is amusing to note that these are common acoustic frequencies, the middle A tuning frequency provided by the oboe in symphony orchestras is 440 Hz, so the LIGO detectors act as super sensitive human ears!


Now let’s discuss gravitational waves [4] in more detail with an eye toward LIGO applications. We will need to model the sources of the gravity waves and the radiation itself. The sources will consist of two masses under the influence of their gravitational attraction. We will treat this problem with Newtonian mechanics because it is adequate up until the last fraction of a second before the merger of the two black holes. In addition we will treat the observed gravity wave as a tiny distortion in the Minkowski metric. These approximations will give us a semi-
quantitative grasp of the real phenomenon. Of course the experts do better! They treat the merging black holes with full, non-linear general relativity. Supercomputing is essential here. The subject of numerical general relativity where the field equations are simulated in regions of strong gravitational effects is a mature subject. It is interesting that the simple approximations and idealizations that we will sketch here match the realistic calculations rather well in most cases. Far from the source the experts also use linearized gravity because the deviation from the Minkowski metric in the vicinity of the LIGO detectors is extraordinarily small.

Before we turn to the LIGO experiment, we are well served by some preliminaries.

**a. Freely Propagating Gravity Waves in Minkowski Space Time**

We have seen that weak gravity waves propagate far from their sources according to the invariant wave equation of special relativity,

$$\Box \bar{h}_{\mu\nu} = 0$$

in the Lorenz gauge,

$$\sum_\rho \partial_\rho \bar{h}^{\rho\mu} = 0$$

These equations are solved by plane waves,

$$\bar{h}_{\mu\nu} = A_{\mu\nu} e^{i \sum_\sigma k_\sigma x^\sigma}$$

The wave equation implies that the wave vector $k_\sigma$ is lightlike, $\sum_\sigma k_\sigma k^\sigma = 0$, which means that gravitons are massless and travel at the speed limit just like light rays. We anticipated these facts earlier in the lecture series.

The quantity $A_{\mu\nu}$ is a symmetric $4 \times 4$ matrix which gives the polarization of the gravity wave. It obeys the Lorenz condition,

$$\sum_\sigma k_\sigma A^{\sigma\rho} = 0$$
We will use the Lorenz condition and the null character of \( k_\sigma \) to simplify \( A_{\mu\nu} \) below. We will see that gravity waves are transverse and have two polarization states. These results were inferred earlier from general principles.

Let’s review these properties for electromagnetic waves before we turn to gravitons. The electromagnetic vector potential \( A^\sigma \) satisfies the massless wave equation in free space time. Therefore, it is represented by a plane wave,

\[
A^\sigma = \epsilon^\sigma e^{i \sum k_\rho x^\rho}
\]

where \( \epsilon^\sigma \) gives the vector’s polarization. The Lorenz condition implies \( \sum k_\sigma \epsilon^\sigma = 0 \). If we consider light rays traveling in the \( z \)-direction so \( k^\sigma = (k, 0, 0, k) \) where \( kc = \omega \), then we learn that \( \epsilon^0 = \epsilon^3 \). Our final condition on \( \epsilon^\sigma \) follows from the fact that electrodynamics physics is unchanged by gauge transformations,

\[
A^\sigma \to A^\sigma + \partial^\sigma \Phi
\]

where \( \Phi \) is a gauge function. In addition, the Lorenz condition is maintained as long as \( \Phi \) satisfies the free wave equation, \( \Box \Phi = 0 \). A plane wave satisfies these conditions, \( \Phi = \chi e^{i \sum k_\rho x^\rho} \) where \( \chi \) is a constant amplitude. So, the gauge transformation for \( A^\sigma \) implies,

\[
A^\sigma \to A^\sigma + \partial^\sigma \Phi = \epsilon^\sigma e^{i \sum k_\rho x^\rho} + ik^\sigma \chi e^{i \sum k_\rho x^\rho}
\]

So if we choose \( \epsilon^0 + ik^0 \chi = 0 \), then the “new” \( \epsilon^\sigma \) has vanishing 0th and 3rd components: it is purely transverse,

\[
\epsilon^\sigma = (0, a, b, 0)
\]

This shows that physical photons have only two polarization states, two linear polarizations, for example. We write,

\[
A^\sigma = \epsilon^\sigma e^{i \sum k_\rho x^\rho} \quad \text{with} \quad \epsilon^\sigma = (0, a, b, 0) = a e_1^\sigma + b e_2^\sigma
\]

Finally, how do we detect traveling electromagnetic waves? Since the electric field is

\[
\vec{E} = -\frac{\partial}{\partial t} \vec{A} - \vec{\nabla} A^0
\]
and $A^0$ vanishes for the traveling wave, we have $\vec{E} = -i\omega \vec{e} e^{i\sum k_\rho x_\rho} = -i\omega \vec{e} e^{i\omega(t-z/c)}$. So, as the wave passes, charged particles will oscillate in a sinusoidal fashion transverse to the z-direction. The light ray is easily detected, as we all know.

Now let’s do the same exercises for traveling gravity waves. We want to understand the polarization $A_{\mu\nu}$ of the gravity wave. Let’s begin by simply counting degrees of freedom. $A_{\mu\nu}$ is a symmetric $4 \times 4$ matrix, so it has 10 components. However, the Lorenz gauge condition, $\sum_\sigma k_\sigma A_{\sigma\rho} = 0$, reduces the independent entries by 4, down to a total of 6. But then we can make infinitesimal coordinate transformation, $h'_{cd} = h_{cd} - \partial_c \xi_d - \partial_d \xi_c$, with the four vector $\xi_c$ which reduces the number of independent entries by another 4, down to a total of 2. These are the two states that we discovered when thinking about the Lorentz invariant spin states of a massless graviton.

Let’s provide the detail of the comments in the previous paragraph. When we write out the Lorenz condition, $\sum_\sigma k_\sigma A_{\sigma\rho} = 0$, for a wave traveling in the z-direction so $k^\sigma = (k, 0, 0, k)$ where $kc = \omega$, we learn that,

$$A_{\mu 3} = A_{\mu 0}$$

An infinitesimal coordinate transformation, $x'_\sigma = x_\sigma + \xi_\sigma$, induces the transformation $h'_{cd} = h_{cd} - \partial_c \xi_d - \partial_d \xi_c$, which induces the transformation $\bar{h}'_{cd} = \bar{h}_{cd} - \partial_c \xi_d - \partial_d \xi_c$ + $\eta_{cd} \sum_a \partial_a \xi^a$ in the variable $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu}$. If we choose $\xi_\sigma$ to satisfy the free wave equation, $\Box \xi_\sigma = 0$, then the Lorenz condition is maintained, $\sum_\sigma \partial_\sigma h_{\sigma\rho} = \sum_\sigma \partial_\sigma \bar{h}'_{\rho\sigma} = 0$. If we choose $\xi_\sigma$ to be a plane wave $\xi_\rho = \epsilon_\rho e^{i\sum k_\rho x_\rho}$, then the transformation for $A_{\mu\nu}$ reads,

$$A'_{\mu\nu} = A_{\mu\nu} - i\epsilon_\mu k_\nu - i\epsilon_\nu k_\mu + i \sum_\sigma \eta_{\mu\nu} \epsilon^\sigma k_\sigma$$

When we write out this relation for $k^\sigma = (k, 0, 0, k)$ we find,

$$A'^{00} = A^{00} - ik(e^0 + e^3) \quad A'^{11} = A^{11} - ik(e^0 - e^3)$$

$$A'^{01} = A^{01} - ik e_1 \quad A'^{12} = A^{12}$$

$$A'^{02} = A^{02} - ik e_2 \quad A'^{22} = A^{22} - ik(e^0 - e^3)$$
Now we choose \( \epsilon_\nu \) so that \( A^{'00} = A^{'01} = A^{'02} = 0 \) and \( A^{'11} = -A^{'22} \). Then,

\[
A^\mu_\nu = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & a & b & 0 \\
0 & b & -a & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]

which is a transverse and traceless matrix. We can write this result in terms of two basis tensors, following our notation for the traveling electromagnetic field,

\[
A^\mu_\nu = a e^\mu_1 + b e^\mu_2
\]

where \( e^\mu_1 \) is diagonal and traceless and \( e^\mu_2 \) is transverse, as we see in the expression for \( A^\mu_\nu \).

Note that since \( h^\mu_\nu \) is traceless, we have \( h^\mu_\nu = h_{\nu\nu} \), and we have the modification of the Minkowski metric directly from this exercise.

b. How to Detect Gravity Waves

Next we need to detect a traveling gravity wave. Consider first its effect on an isolated single particle. That particle travels on a geodesic,

\[
du^\sigma / d\tau + \sum_\mu_\nu \Gamma^\sigma_\mu_\nu u^\mu u^\nu = 0
\]

Suppose that the particle were initially at rest in this frame, \( u^\sigma = (c, 0, 0, 0) \). Then the geodesic equation gives,

\[
\frac{d}{d\tau} u^\sigma = -c^2 \Gamma^\sigma_\mu_\nu u^\mu u^\nu = -\frac{1}{2} c^2 \sum_\rho \eta^\sigma_\rho (\partial_\rho h_{\rho0} + \partial_0 h_{\rho0} - \partial_\rho h_{00})
\]

But in the transverse, traceless gauge, \( h^\mu_\nu = h_{\mu\nu} \) and only \( h_{ij} \) for \( i, j = 1, 2 \) are non-zero. So, \( \frac{d}{d\tau} u^\sigma = 0 \) and the particle doesn’t move relative to the coordinates! (We argued this point before, qualitatively.) In other words, because of the equivalence principle, the test particle moves in sync with the coordinate mesh. To get an observable we need to measure the physical distance between test particles as the gravity wave passes. The spatial position of a test particle, \( \xi^\sigma = (0, \xi^1, \xi^2, \xi^3) \), is constant in time, as we just derived. Let \( \xi^\sigma \) now be the coordinate difference between two test particles. Then the physical distance between them is given by,
\[ l^2 = \sum_{ij} g_{ij} \xi^i \xi^j = \sum_{ij} (\delta_{ij} - h_{ij}) \xi^i \xi^j \]

where the sums run over the spatial coordinates 1, 2, and 3. We see here that \( h_{ij} \) will have non-zero components and it oscillates in space time, so we do have a measurable effect here!

There is a simple analogy to help us understand this result. Suppose you are the harbor master at Well Fleet on the Cape. You label points in the bay with free floating buoys. Suppose there are two fishing boats, one near buoy A and the other near buoy B. As the tide changes or as a wave comes through the bay, the boats and the buoys move in lockstep but the physical distance between the two boats change: as more water races into the bay the boats (and buoys) are pushed apart. The boats’ positions relative to the buoys, however, doesn’t change.

Back to gravity. The expression for \( l^2 \) can be written in a slightly more conventional fashion. Let’s write it with a Euclidean metric \( \delta_{ij} \). This is easily done to first order in \( h_{ij} \). Define,

\[ \zeta^j = \xi^j + \frac{1}{2} \sum_k h^i_k \xi^k \]

So,

\[ l^2 = \sum_{ij} \delta_{ij} \zeta^i \zeta^j + O(h^2) \]

where we were careful with signs, \( h^i_k = \sum_l \eta^{il} h_{lk} = -h_{ik} \).

Suppose a gravity wave approaches with + polarization, \( A_{TT}^{\mu\nu} = ae_1^{\mu\nu} \). Then, taking real parts,

\[ h_{TT}^{\mu\nu} = ae_1^{\mu\nu} \cos k \cdot x = ae_1^{\mu\nu} \cos \left( \omega(t - z/c) \right) \]

So finally,

\[ \zeta^j = (\xi^1, \xi^2, 0) - \frac{1}{2} a \cos \left( \omega(t - z/c) \right) (\xi^1, -\xi^2, 0) \]
Let’s measure $\zeta^j$ with respect to a reference particle (labelled o) at the center of a ring of particles, as shown in Fig. 16.1 where we show its time evolution through one cycle of the gravity wave.

![Diagram of relative positions of test particles](image)

Note that the pattern in Fig. 16.1 of stretching and squeezing comprise a quadrupole moment, tidal forces, that we have seen in several applications already.

Finally if we take the other polarization $A_{TT}^{\mu\nu} = be_2^{\mu\nu}$ then the pattern of $\zeta^j$ is the same as Fig. 16.1, except rotated by 45°. This is the $\times$ polarization.

**c. The Radiation and Detection of Gravity Waves from Binary Systems**

Before we consider the LIGO experiment in more detail, let’s consider the different frames of reference used in our subsequent discussions and analyses. First, there is the frame fixed to the earth. Since the mechanical structures of LIGO’s arms are bound together by electromagnetic
forces which are many, many orders of magnitude greater than gravitational forces, the gravity waves detected by the apparatus do not have any effect here. In particular the speed of light relative to the earth frame, the massive arms of the structure, is the speed limit c because the gravitational fields in the vicinity of the LIGO apparatus are weak. In addition, the gravitational potential due to the earth’s mass is essentially constant over the apparatus and does not play a role in LIGO’s analysis. Gravity waves do effect proper lengths and times, such as the proper length between the freely suspended mirrors at the end of each arm of the detector. These are the lengths and times the analysis focuses on.

To begin suppose that the LIGO detector lies in the x-y plane and the source is far away in the z direction as shown in Figure 16.2.
Fig. 16.2  A binary star system radiating gravity waves to a detector at z.

Then we parametrize the traveling gravitational wave $\varepsilon(x^\mu)$ and write the metric in either of two forms,

$$ds^2 = dt^2 - (1 - \varepsilon)dx^2 - (1 + \varepsilon)dy^2 - dz^2 \quad (\varepsilon \ll 1, +\text{polarization}) \quad (16.1a)$$

or,

$$ds^2 = dt^2 - dx^2 - 2\varepsilon dx dy - dy^2 - dz^2 \quad (\varepsilon \ll 1, \times \text{polarization}) \quad (16.1b)$$

Let’s concentrate on the + polarization case expressed in the metric Eq. (16.1a). We are anticipating here that the source of the gravity wave will be an oscillating quadrupole moment of a mass distribution, so if there is a “stretch” in one direction transverse to the propagation direction...
of the wave, the z axis, then there is a “compression” in the perpendicular direction. The length of
an arm of the LIGO detector is about \(4 \times 10^3\) m. and the wavelength of the gravity wave is \(\lambda = c/\nu \approx 3 \times 10^8/10^2 = 3 \times 10^6\) m. Therefore, \(\epsilon(x^\mu)\) can be treated as a constant across the
apparatus. Since \(\epsilon(x^\mu)\) satisfies the free field wave equation in the vicinity of the detector and the
gravity wave from the merging black holes is traveling along the z axis, \(\epsilon\) is a function of \(t - z/c\).
When this wave passes and \(\epsilon\) is positive, then the physical length of the x-arm of the detector
increases and the physical length of the y-arm of the detector decreases according to Eq. (16.1),
\[
\Delta x_{arm} = \sqrt{1 + \epsilon} \Delta x \approx \left(1 + \frac{\epsilon}{2}\right) \Delta x, \quad \Delta y_{arm} = \sqrt{1 - \epsilon} \Delta y \approx \left(1 - \frac{\epsilon}{2}\right) \Delta y
\] (16.2)
Therefore, the time difference between light traveling to and fro along the two arms is,
\[
\Delta t_{arm} = \left(\frac{1}{2} \epsilon + \frac{1}{2} \epsilon\right) 2L/c = 2 \epsilon L/c
\] (16.3)
For \(N\) trips of each laser beam, down and back, along each arm,
\[
\Delta t_{arm} = 2N \epsilon L/c
\] (16.4)
The problems discuss these predictions and LIGO’s parameters in more detail.

Now let’s consider the source of the gravity waves, the right hand side of the field
equations. Let’s model the source as two masses \(M_1\) and \(M_2\) orbiting one another at a distance \(r\).
They move non-relativistically so Newtonian mechanics applies. An exercise in Newtonian
mechanics, that is reviewed in the problem set, reminds us that the binding energy of the pair is,
\[
E_B = -\frac{GM_1M_2}{2r}
\] (16.5)
Since the masses are in a constant accelerated state, the system can radiate gravity waves
and gradually lose energy. When this happens the distance \(r\) between them will monotonically
decrease and they will eventually collide and merge in the case of two black holes. This system
becomes a model for realistic systems that could radiate gravity waves and be detected by
advanced LIGO. The rate of energy loss to gravity waves requires a detailed analysis of the field
equations that will be discussed below. The result is that the rate of energy loss to gravity waves is,

\[ \frac{dE_B}{dt} = -\frac{32G^4}{5c^5r^5} (M_1M_2)^2(M_1 + M_2) \]  

(16.6)

By combining Eq. (16.5) and (16.6), one derives the rate at which \( r \) decreases due to the energy loss,

\[ \frac{dr}{dt} = -\frac{64G^3}{5c^5r^3} M_1M_2(M_1 + M_2) \]  

(16.7)

These expressions have excellent experimental support from the detection of binary neutron stars.

These results are pursued in more detail in the problem sets.

Finally we quote the result for the gravitational wave radiated from this system,

\[ \epsilon(t, z) = -\frac{4G^2M_1M_2}{rzc^4} \cos\left(2\pi\nu(t - z/c)\right) \]  

(16.8)

where \( \nu \) is twice the frequency of the orbiting binary, \( \nu_{\text{orbit}} \). We will show below that the quadrupole moment of the mass distribution is responsible for the radiation and the relation \( \nu = 2\nu_{\text{orbit}} \) follows just as it does for quadrupole radiation in electromagnetism. The Newtonian prediction for \( \nu_{\text{orbit}} \) follows from Kepler’s law,

\[ \nu_{\text{orbit}} = \frac{1}{2\pi} \left(\frac{G(M_1+M_2)}{r^3}\right)^{1/2} \]  

(16.9)

These relations are important because LIGO is sensitive to the range of frequencies 50-500 Hz and this matches binaries that are the candidates for LIGO sources. This will be discussed further below.

Now let’s return to the wave equation Eq. (11.15) and derive some of the features of gravity waves emitted by a binary system. We will see that \( \tilde{h}_{\mu\nu} \) is transverse and traceless in this application, so \( \tilde{h}_{\mu\nu} = h_{\mu\nu} \), and the wave equation gives us the modification of the Minkowski
metric directly. We will be satisfied with a semi-quantitative discussion. If the reader worked through Appendix F of the textbook, she knows the form of the solution to Eq. (11.15). It has the same form as the four vector potential radiated by a general charge distribution. In this case the gravity wave piece of the metric, the solution to Eq. (11.15), reads,

$$h_{\mu \nu}(t, x, y, z) = \frac{4G}{c^4} \iiint T_{\mu \nu}(x', y', z', t - R/c) \frac{dx'dy'dz'}{R}$$

(16.10)

where $R^2 = (x - x')^2 + (y - y')^2 + (z - z')^2 = c(t - t')$. The temporal offset, $t - t'$ accounts for the fact that the bits of the energy-momentum tensor that effect the gravity wave at time $t$ originated a time earlier that accounts for the time needed for the signal to a travel a distance $R$ at the speed limit $c$. This result was discussed further in Appendix F. In our application the observation point is very(!) far from the source, so we can pull the factor of $R^{-1}$ outside the integral,

$$h_{\mu \nu}(t, x, y, z) = \frac{4G}{c^4 R_0} \iiint T_{\mu \nu}(x', y', z', t - R/c) dx'dy'dz'$$

(16.11)

where $R_0$ is the distance from the observation point to the origin of the coordinate system which is chosen at the center of mass of the radiating system. We model the radiating system as two masses $M_1$ and $M_2$ orbiting one another in the x-y plane, as shown in the Figure 16.3.
We treat the two body system with Newtonian mechanics which proves to be adequate until the last fraction of a second when the two black holes merge. Furthermore, if we model the system with a circular orbit, then the two bodies simply rotate around their common center of mass with a frequency $\nu_{\text{orbit}} = \omega/2\pi$ which is given by Kepler’s law, as reviewed in the problem set. The orbits of the two masses are given by,

$$\begin{align*}
   x_1^{(1)}(t) &= r_1 \cos \omega t, \quad x_2^{(1)}(t) = r_1 \sin \omega t, \quad x_3^{(1)}(t) = 0 \quad (16.12a) \\
   x_1^{(2)}(t) &= r_2 \cos \omega t, \quad x_2^{(2)}(t) = r_2 \sin \omega t, \quad x_3^{(2)}(t) = 0 \quad (16.12b)
\end{align*}$$

The energy-momentum density Eq. (16.11) is proportional to the sum of two terms, one for each of the two particles, one proportional to $\delta \left(x' - x_1^{(1)}(t)\right) \delta \left(y' - x_2^{(1)}(t)\right) \delta(z')$ and one proportional to $\delta \left(x' - x_1^{(2)}(t)\right) \delta \left(y' - x_2^{(2)}(t)\right) \delta(z')$. As a result, the source of the gravity
waves is the energy-momentum tensor $T_{\mu\nu}$, discussed in Sec. 11.8, whose Cartesian coordinates behave as,

$$T_{ij} = \left( m_1 v_i^{(1)} v_j^{(1)} + m_2 v_i^{(2)} v_j^{(2)} \right)$$  \hspace{1cm} (16.13)

where $v_i^{(1)} = dx_i^{(1)}/dt$ for mass 1 and $v_j^{(2)} = dx_j^{(2)}/dt$ for mass 2. So, differentiating Eq. (16.12) to make the velocities $v_i^{(1)}$ and $v_j^{(2)}$, the non-zero spatial components of $T_{\mu\nu}$ are,

$$T_{11} = I \omega^2 \cos^2 \omega t = \frac{1}{2} I \omega^2 (1 + \cos 2\omega t)$$

$$T_{12} = T_{21} = I \omega^2 \cos \omega t \sin \omega t = \frac{1}{2} I \omega^2 \sin 2\omega t$$

$$T_{22} = I \omega^2 \sin^2 \omega t = \frac{1}{2} I \omega^2 (1 - \cos 2\omega t)$$

where $I = M_1 r_1^2 + M_2 r_2^2$ is the moment of inertia of the two body system. LIGO is only sensitive to the time dependent pieces of $T_{\mu\nu}$ which read,

$$T_{11} = +\frac{1}{2} I \omega^2 \cos 2\omega t$$

$$T_{12} = T_{21} = I \omega^2 \sin 2\omega t$$

$$T_{22} = -\frac{1}{2} I \omega^2 \cos 2\omega t$$  \hspace{1cm} (16.14)

Since $h_{ij}(t)$ is proportional to $T_{ij}(t - z/c)$ where $z$, as shown in Figure 16.2 above, is the distance to the source,

$$h_{ij}(t) = \frac{2G}{c^4 z} T_{ij}(t - z/c)$$  \hspace{1cm} (16.15)

We see that 1. $h_{11}(t) = -h_{22}(t)$, as anticipated in Eq. (16.1a), and 2. $\omega_{\text{gravity}} = 2\omega_{\text{orbit}}$, the frequency of the gravity wave is determined by that of the quadrupole moment which is twice that of the orbiting masses. Note also that $h_{ij}$ falls as the inverse of the distance to the source.
The reader is encouraged to follow the operation of advanced LIGO and see what additional progress it makes in observational astronomy. A space based version of LIGO called “LISA” is also in the planning and prototyping stage. It should be sensitive to much lower frequency waves and should observe the merging of super-heavy black holes that astronomers believe exist at the centers of galaxies.

The discussion here is only qualitative because we have applied the wave equation Eq (11.15) near the merging black holes where space time is actually strongly curved. The full Einstein field equations must be used to make quantitative predictions.

Appendix G, Sec. 3 of the textbook, contains a discussion of gravity wave detection using the geodesic deviation (“tidal forces”) equations applied to LIGO’s mirrors. This approach uses the Riemann curvature, a physical tensor, directly.


Let’s continue our discussion in lecture 6 where we considered some of the major differences between electromagnetism and relativity, and now contrast special relativity with general relativity.

Special relativity is characterized by having global symmetries. There are ten in all,

a. Translations in four space time dimensions,

b. Rotations about the three spatial axes, and

c. Boosts (Lorentz transformations) along any of the three spatial axes.

This collection of ten symmetries constitute the “Poincare Group”. General relativity takes these global symmetries and makes them local, i.e. space-time dependent. In special relativity the ten symmetry operations have the same action over all of space-time. In general relativity the degree
of symmetry is much higher. One requires that the theory be invariant to space-time dependent coordinate transformations. The theory must be invariant to coordinate transformations which might be appreciable only in the vicinity of some space-time event and negligible elsewhere. This was the guiding principle in the construction of Einstein’s field equations.

In special relativity we took great care to set up a coordinate system of meter sticks and clocks over all of space-time so that measurements could be compared between frames in a local fashion. We found that clocks which were synchronized in one frame were not synchronized in moving frames and we discovered time dilation, Lorentz contraction and the relativity of simultaneity. We never embarked on an analogous activity in general relativity because, as we found out when considering parallel translation, comparing measurements at non-zero distances in general relativity is not generally possible. The curvature of spacetime makes any attempt in this direction path dependent and hence non-universal. In the context of differential geometry this is a consequence of the fact that tangent planes at different points on a curved surface, or freely falling frames in general relativity, do not coincide and do not give the basis of global comparisons of time and spatial measurements.

We have seen that the implementation of the Equivalence Principle can be quite subtle. The Equivalence Principle means that in small enough regions of spacetime, the laws of physics reduce to those of special relativity. In a freely falling coordinate system it becomes impossible to detect the gravitational field by strictly local experiments. In a local inertial reference frame the metric at a point $P$ is given by the Minkowski metric $g_{\sigma\rho}^{(0)}(P)$ and the first derivatives of the metric must vanish, $\partial_\alpha g_{\sigma\rho}(P) = 0$ in the falling frame. This condition, $\partial_\alpha g_{\sigma\rho}(P) = 0$, means that there are no forces in the freely falling frame at $P$. This is the essence of the Equivalence Principle. One
must consider second derivatives of the metric to find non-zero elements in the freely falling frame. These quantities produce the curvature of spacetime.

The Equivalence Principle also makes general relativity intrinsically non-linear. The idea is that gravity couples universally to energy-momentum. But the gravitational field itself carries energy-momentum, so it must couple to itself. We saw the non-linearity in the Einstein field equations and the Schwarzschild metric. Even simpler, the expression for the Riemann tensor has terms which are linear as well as quadratic in the Christoffel symbols. The Christoffel symbols themselves involve the products of the inverse of the metric, the metric and its first and second derivatives. It is the intrinsic non-linearity of these equations which makes it so challenging to find exact solutions to general relativity. The theory does not enjoy the principle of linear superposition that plays such an important role in electrodynamics.

Modern quantum field theories of elementary particles are also based on local symmetry principles and universality. These theories are formulated in Minkowski spacetime, but they have “internal” symmetry operations that are local symmetries and these symmetries dictate interactions and conservation laws. The theory of strong interactions “Quantum Chromodynamics” is based on the premise that the three colors of quarks become a local symmetry principle. The theory contains quark fields as well as gluon fields which also carry color. The gluons are analogous to the electromagnetic field of electrodynamics but they carry the color quantum number themselves in contrast to the electromagnetic fields and its quantum, the photon, which are neutral. The fact that the gluons in Quantum Chromodynamics carry color implies that they interact among themselves which makes the theory intrinsically non-linear. This property of Quantum Chromodynamics is believed to underlie its attribute of Quark Confinement. Certainly the gauge theories of elementary particles and general relativity have more in common than appear at first. We saw in our
introduction of the Riemann tensor which underlies the Einstein curvature tensor which enters the
field equation for general relativity, that closed loops are the essential geometric objects used to
formulate the theory. Closed loops are also the basic geometric object underlying the construction
of gauge theories of elementary particle physics. In all cases we are formulating theories with local
symmetry groups and products of operators around closed loops capture the character of the
theories precisely. In general relativity the local symmetry group is that of Poincare, in Quantum
Chromodynamics it is rotations in local color space of quarks and gluons. Covariant derivatives
must be constructed and used in each theory to express coordinate-free physical differences. Of
course the language of quantum field theory is needed for elementary particle physics.

Unfortunately a quantum formulation of gravity eludes physics and it is not even known if the two
foundations of modern physics are mutually compatible. Perhaps string theory will shed light here.

Another particularly significant difference between field theories based on a flat Minkowski
metric and general relativity concerns the nature of the scale of energy. In all these theories except
general relativity, the absolute scale of energy is not significant – only differences of energy matter.
For example, in Newton’s world forces are given by the gradient of potentials, so the zero point of
the potential energy is irrelevant. All of this changes in general relativity because the source of the
Einstein tensor is the energy-momentum tensor, not its derivatives! The zero point of energy is
established by the energy density of the vacuum. In a quantum field theory, this is very perplexing.
In quantum field theory, quantum fluctuations in the vacuum contribute to the vacuum energy.
These contributions occur over all length scales from zero to infinity. It appears that to understand
and calculate the vacuum energy one must understand physics at arbitrarily high energies! Without
some unknown principle or symmetry, this knowledge is utterly beyond our reach. Given these
obstacles, the only way forward appears to be phenomenology.
This brings us to an introductory discussion of the “Cosmological Constant”. If the vacuum energy must be included in the Einstein field equations then it should have the form,

\[ T^{(\text{vac})}_{\mu\sigma} = \rho_{\text{vac}} g_{\mu\sigma} \]  \hspace{1cm} (17.1)

And Einstein’s equation becomes,

\[ G_{\mu\sigma} = \frac{8\pi G}{c^4} \left( T^{(M)}_{\mu\sigma} + \rho_{\text{vac}} g_{\mu\sigma} \right) \]

where \( T^{(M)}_{\mu\sigma} \) is the matter field contribution to the energy-momentum tensor that we have discussed and illustrated earlier. This equation is usually written,

\[ G_{\mu\sigma} - \Lambda g_{\mu\sigma} = \frac{8\pi G}{c^4} T^{(M)}_{\mu\sigma} \]  \hspace{1cm} (17.2)

where \( \Lambda \) is the “cosmological constant” and \( \Lambda = 8\pi G \rho_{\text{vac}} \). Eq. (17.1) has been coined “dark energy”.

We shall see that the cosmological constant produces an acceleration of the expansion of the universe. The observation of the spectra of light emitted by certain supernova indicates that,

\[ |\rho^{\text{obs}}_{\text{vac}}| \approx (10^{-12} \text{ GeV}) \sim 10^{-8} \text{ erg/cm}^3 \]  \hspace{1cm} (17.3)

This is a tiny energy density which is numerically negligible on the scale of galaxies, etc. but has large cosmological effects. It is interesting to ask if quantum field theory can provide a prediction for \( \Lambda \). In some future quantum field theory of gravity one might use dimensional analysis to make an “estimate”. The dimensional numbers in the theory are the Planck constant \( \hbar \), the speed of light \( c \), and Newton’s gravitational constant \( G \). These constants can be combined to produce a “Planck energy”,

\[ E_P = \left( \frac{\hbar c^5}{G} \right)^{1/2} \approx 1.22 \times 10^{19} \text{ GeV} \approx 1.95 \times 10^{16} \text{ erg} \]
If this “natural” energy scale sets the scale for the energy density of the vacuum, we would “predict”,
\[ \rho_{vac} \sim (10^{19} \text{ GeV})^4 \approx 10^{112} \text{ erg/cm}^3 \]
So, our “back of the envelop” estimate of \( \rho_{vac} \) is wrong by 120 orders of magnitude! A cosmological constant of this order would overwhelm the right hand side of Einstein’s equation and vitiate all of its great successes. Something is terribly(!) wrong with these estimates. Many suggestions have been made to circumvent this disaster but none of them are truly satisfactory and a resolution is an open challenge! Perhaps a full-fledged quantum field theory of gravity would provide a solution. Breakthroughs in string theory might be required as well. For the time being, we proceed phenomenologically and include the non-zero estimate Eq. (17.3) in cosmological calculations of the expansion of the universe.

Without reviewing modern developments in applications of general relativity to cosmology, the large scale structure of the universe, we can argue in elementary terms that \( \Lambda \) leads to an accelerating expansion of the universe. Consider the Newtonian limit of Eq. (17.2). Retracing our steps in lecture 11, we easily find that \( \Lambda \) modifies Newton’s law of gravity,
\[ \nabla^2 V = 4\pi G \rho - \Lambda c^2 \]  \hspace{1cm} (17.4)
We again see that \( \Lambda \) had better be very small on terrestrial scales to have escaped observation! Using Eq. (17.4) we calculate the gravitational acceleration a distance \( r \) outside a mass \( M \),
\[ g = -\nabla V(r) = -\frac{GM}{r^2} \hat{r} + \frac{1}{3} \frac{\Lambda c^2}{r} \hat{r} \]
We find a universal acceleration outward which grows(!) with distance.

The reader is encouraged to consult the literature on this timely subject and develop a more sophisticated understanding of the theoretical and experimental issues in this challenging field.
Special Topic A. The Isotropic Metric and Linearized General Relativity

Recall the discussion of linearized General relativity in Lecture 11. We showed that for weak, static gravity we could obtain Newton’s law of gravity, $\nabla^2 \Phi = 4\pi G \rho$ from the wave equation for $\bar{h}^{\mu\nu}$ if we identified $\bar{h}^{00} = 4\Phi/c^2$. In addition, we saw that all other components of $\bar{h}_{\alpha\beta}$ and $T_{\alpha\beta}$ are negligible in this non-relativistic, static, weak field approximation.

We need to find $h_{\mu\nu}$ in order to construct the metric $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. Begin with the definition

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} h$$  \hspace{1cm} A.1

Take the trace of this equation,

$$\bar{h} = \sum_{\mu} \bar{h}^\mu_\mu = \bar{h}^0_0 = \bar{h}_{00} = 4\Phi/c^2$$ which implies $h = -4\Phi/c^2$. Now we can get $h_{00}$. Write out Eq. A.1 for the 0-0 component, $\bar{h}_{00} = h_{00} - \frac{1}{2} \eta_{00} h$ which reads $4\Phi/c^2 = h_{00} + 2\Phi/c^2$, so $h_{00} = 2\Phi/c^2$. Next we can get $h_{11}$ from the 1-1 component of Eq. A.1,

$$\bar{h}_{11} = h_{11} - \frac{1}{2} \eta_{11} h = h_{11} - 2\Phi/c^2$$  \hspace{1cm} A.3

But $\bar{h}_{11}$ is negligible, so $h_{11} = 2\Phi/c^2$. Similarly, we find $h_{22} = h_{33} = 2\Phi/c^2$.

Now we have our prize, the metric in the static, weak field case,
\[ ds^2 = \left(1 + \frac{2\Phi}{c^2}\right)c^2 dt^2 - \left(1 - \frac{2\Phi}{c^2}\right)(dx^2 + dy^2 + dz^2) \quad \text{(A.4)} \]

This result is accurate to order \(1/c^4\). This coordinate system, which is discussed further in the problem sets in the textbook, is called “isotropic” because of its manifest rotational symmetry in the spatial coordinates. It is generally very handy in weak field applications.

Now let’s retrieve Newton’s second law in weak, linear gravity. It will follow from the geodesic equation which predicts how a mass \(m\) travels in a force-free space-time described by curvilinear coordinates,

\[
\frac{d^2\vec{x}_\beta}{d\tau^2} + \sum_{\alpha\gamma} \Gamma^\alpha_{\gamma\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\gamma}{d\tau} = 0 \quad \text{(A.5)}
\]

We need the Christoffel symbols for the metric Eq. (A.4). These are easily calculated using Eq. (9.2). The formulas in linearized, weak field gravity are only accurate to first order in \(2\Phi/c^2\), so we raise and lower indices using the Minkowski metric of the background spacetime. In addition, the gravitational field is time independent in the class of frames of reference used here. Then it is easy to verify (Problem 12.18),

\[
\Gamma^0_{00} = \frac{\partial \Phi}{c^2 \partial c_t}, \quad \Gamma^i_{00} = \frac{\partial \Phi}{c^2 \partial x^i}
\]

We simplify Eq. (A.5) for non-relativistic kinematics, \(p^0 \gg p^i\), and choose the \(\beta = 0\) component. The zeroth component of Eq. (A.5) then reduces to,

\[
\frac{d}{dt} (E + m\Phi) = 0 \quad \text{(A.6)}
\]

where we identified \(E = mc^2 \frac{dt}{d\tau}\) in the first term and we only kept leading terms for non-relativistic kinematics, \(\frac{dx^\gamma}{d\tau} \equiv (c, 0,0,0)\) in the second term. Eq. A.6 is just Newtonian energy conservation!

Similarly for \(\beta = i, (i = 1,2,3)\), Eq. (A.5) becomes Newton’s second law for a particle of mass \(m\) in a gravitational field,
\[
\frac{d}{dt} p = -m \nabla \Phi
\]
to leading order. These are the expected results, but obtained from the perspective of general relativity.

This example illustrates again that in Einstein’s world gravitation is curvature expressed through non-zero \( \Gamma^\beta_{\gamma \alpha} \), and in Newton’s world gravitation is a universal force, proportional to every body’s inertial mass. In the weak field, non-relativistic regime, these two approaches give identical predictions.

**Special Topic B. Slowly Rotating Stars, Frame-Dragging and Lense-Thirring Precession**

Now let’s apply this formalism to some simple but instructive situations [4,9,10]. What is the metric, to first order in \( h_{\mu \sigma} \), outside a rotating mass (a star or planet)? Although the source is rotating let’s suppose that there is no explicit time dependence. Examples might be a rigid sphere rotating at a constant angular velocity \( \omega \). Or perhaps the star only has cylindrical symmetry about a fixed axis of rotation. In either case, \( \partial_\alpha T^{\alpha \beta} = 0 \) and \( \partial_0 \bar{h}^{\alpha \beta} = 0 \). The wave equation Eq. 11.15 reduces to Poisson’s equation,

\[
\nabla^2 \bar{h}^{\alpha \beta}(\vec{x}) = \frac{-16\pi G}{c^4} T^{\alpha \beta}(\vec{x})
\]  

(B.1)

which was discussed and solved in the textbook in the context of electrodynamics,

\[
\bar{h}^{\alpha \beta}(\vec{x}) = \frac{4G}{c^4} \int \frac{T^{\alpha \beta}(\vec{y})}{|\vec{x} - \vec{y}|} d^3y
\]  

(B.2)

Furthermore, recall from lecture 6 that the simplest \( T^{\alpha \beta}(\vec{x}) \) reads,

\[
T^{\alpha \beta}(\vec{x}) = \rho(\vec{x}) u^{(\alpha}(\vec{x}) u^{\beta)}(\vec{x})
\]  

(B.3)
where $\rho(\vec{x})$ is the proper mass density of the source and $u^{\alpha}(\vec{x})$ is its four velocity. If we apply this source to non-relativistic stars and planets where $|u^{\ell}(\vec{x})/c| \ll 1$, then to first order $u^{\alpha} = dx^{\alpha}/d\tau \approx (c, \vec{u})$. And

$$
\begin{align*}
T^{00} &= \rho c^2 \\
T^{0i} &= \rho cu^i \\
T^{ij} &= \rho u^{\ell}u^{j} \approx 0
\end{align*}
$$  \hfill (B.4)

where we neglect $T^{ij}$ because it is a second order correction and the formalism here is accurate only to first order. The expressions for the metric follow,

$$
\begin{align*}
\bar{h}^{\alpha\beta}(\vec{x}) &= -\frac{4G}{c^2} \int \frac{\rho(\vec{y})}{|\vec{x} - \vec{y}|} \, d^3 y \\
\bar{h}^{0i}(\vec{x}) &= -\frac{4G}{c^2} \int \frac{\rho u^{\ell}(\vec{y})}{|\vec{x} - \vec{y}|} \, d^3 y
\end{align*}
$$  \hfill (B.5)

and $\bar{h}^{ij}(\vec{x}) \approx 0$. To bring out the similarities to problems of rotating charge distributions in electrodynamics, it is convenient to define,

$$
\begin{align*}
\Phi(\vec{x}) &= -G \int \frac{\rho(\vec{y})}{|\vec{x} - \vec{y}|} \, d^3 y \\
A^{i}(\vec{x}) &= -\frac{4G}{c^2} \int \frac{\rho u^{\ell}(\vec{y})}{|\vec{x} - \vec{y}|} \, d^3 y
\end{align*}
$$  \hfill (B.6)

So,

$$
\begin{align*}
\bar{h}^{00} &= \frac{4\Phi}{c^2} \\
\bar{h}^{0i}(\vec{x}) &= \frac{A^{i}}{c} \\
\bar{h}^{ij} &\approx 0
\end{align*}
$$  \hfill (B.7)

In order to retrieve the metric $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, we recall that $\bar{h}_{\alpha\beta} = h_{\alpha\beta} - \frac{1}{2} \eta_{\alpha\beta} h$, where $h = \sum_{\alpha} h^{\alpha}_{\alpha}$. Note that one can invert this expression by noting that $\bar{h} = \sum_{\alpha} \bar{h}^{\alpha}_{\alpha} = -h$, so $h_{\alpha\beta} = \bar{h}_{\alpha\beta} - \frac{1}{2} \eta_{\alpha\beta} \bar{h}$. From Eq. B.7 we read off $\bar{h} = \bar{h}^{00} - \sum_{i} \bar{h}^{ii} = \bar{h}^{00}$, so

$$
\begin{align*}
h_{00} &= \bar{h}^{00} - \frac{1}{2} \eta_{00} \bar{h} = \frac{1}{2} \bar{h} \\
h_{11} &= \bar{h}^{11} - \frac{1}{2} \eta_{11} \bar{h} = \frac{1}{2} \bar{h} = h_{22} = h_{33} \\
h_{0i} &= \frac{1}{c} A^{i}
\end{align*}
$$  \hfill (B.8)

This produces the metric to first order in $|\vec{u}/c| \ll 1$,

$$
ds^2 = \left(1 + \frac{2\Phi}{c^2}\right) c^2 dt^2 + \sum_{i} \frac{2A^{i}}{c} c dt dx^{i} - \left(1 - \frac{2\Phi}{c^2}\right) (dx^2 + dy^2 + dz^2) \hfill (B.9)$$
This generalizes the weak field isotropic metric studied in the textbook to stationary rotating sources which produces the new $g_{0l}$ term.

These expressions can be simplified further for a rotating uniform rigid sphere of radius $R$. Suppose that one samples the gravitational field far from the source, $|\vec{x}| \gg R$. Then the leading term in $h_{00}$ is,

$$h_{00} = -\frac{2GM}{c^2 r} \quad (B.10)$$

where $r$ is the spherical polar coordinate, the distance from the origin which we take as the center of the sphere. This is a very familiar result. In order to evaluate $A_l$ we need the spatial distribution of the velocity, $\vec{\nu}(\vec{y}) = \vec{\omega} \times \vec{y}$ where $\vec{\omega}$ will be chosen to point along the fixed $z$ axis. In terms of Cartesian coordinates, $\nu^k(\vec{y}) = \sum_{lm} \epsilon^{klm} \omega_l y^m$. Again, we take $r \gg R$ so we can expand the denominator in Eq. B.6,

$$|\vec{x} - \vec{y}|^{-1} = \left( x^2 - 2 \vec{x} \cdot \vec{y} + y^2 \right)^{-1/2} \approx \frac{1}{r} \left( 1 + \frac{\vec{x} \cdot \vec{y}}{r^2} + \ldots \right) \quad (B.11)$$

So,

$$h_{0k}(\vec{x}) = -\frac{4G}{c^3} \sum_{klm} \epsilon_{klm} \omega^l \int \frac{\rho(\vec{y})}{r} y^m \left( 1 + \frac{\vec{x} \cdot \vec{y}}{r^2} + \ldots \right) d^3 y$$

But $\int \rho(\vec{y}) y^m d^3 y = 0$ for a mass distribution which is symmetric with respect to the three spatial directions. Only even powers of the components of $\vec{y}$ contribute. So,

$$h_{0k}(\vec{x}) = -\frac{4G}{c^3 r^3} \sum_{lm} \epsilon_{klm} \omega^l x^m \int \rho(\vec{y}) (y^m)^2 d^3 y \quad (B.12)$$

Since $\omega^l$ is non-zero only for component $l = 3$, only the $m = 1$ and 2 terms contribute to Eq. B.12,

$$h_{0k}(\vec{x}) = -\frac{2G}{c^3 r^3} \sum_{m} \epsilon_{k3m} \omega^3 x^m \int \rho(\vec{y}) [(y^1)^2 + (y^2)^2] d^3 y \quad (B.13)$$

where we used the cylindrical symmetry of $\rho(\vec{y})$, $\int \rho(\vec{y}) (y^1)^2 d^3 y = \int \rho(\vec{y}) (y^2)^2 d^3 y$.

Recognize the moment of inertia about the z-axis, $I_z$, here and the angular momentum $J^3 = \omega^3 I_z$ about the z-axis,
\[ h_{0k}(\vec{x}) = -\frac{2G}{c^3r^3}J^3 \sum m \epsilon_{k3m}x^m \]  
(B.14)

Finally, \( h_{0k} \) contributes a piece to the line element,

\[
ds^2 = 2 \sum_k h_{0k} c dt \, dx^k + \text{(diagonal contributions)}
\]

\[
= 2h_{01} c dt \, dx^1 + 2h_{02} c dt \, dx^2 + \ldots
= -\frac{4GJ}{c^3r^3} (x^2 dx^1 - x^1 dx^2) c dt + \ldots
\]

(B.15)

But in spherical coordinates \( x^1 dx^2 - x^2 dx^1 = r^2 \sin^2 \theta \, d\varphi \), so the metric becomes,

\[
ds^2 = +\frac{4GJ}{c^3r} \sin^2 \theta \, d\varphi c dt + \text{(diagonal contributions)}
\]

(B.16)

This metric, its physics and formalism, will be discussed further below.

Collecting our weak field metric outside a rotating stationary body, we have the invariant differential line element,

\[
ds^2 = \left(1 - \frac{2GM}{c^2r}\right) c^2 dt^2 - \left(1 + \frac{2GM}{c^2r}\right) (dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta \, d\varphi^2) + \frac{4GJ}{c^3r^3} (x dy - y dx) c dt
\]

(B.17)

We want to illustrate the effects of the off-diagonal term that is proportional to the angular momentum \( J \) carried by the mass \( M \). Recall that this term is magnetic in character – its origin was from the mass current, \( \rho u^0 u^k \) carried by the spinning object. It has no non-relativistic analog – in Newton’s world the gravitational potential due to a localized mass is unaffected by its rotational velocity.

Let’s imagine a gyroscope falling along the z-axis toward the rotating planet [4,9,10]. Let the gyroscope point in the x-y plane. We want to see if the rotation of the planet causes the gyroscope to precess. We choose this configuration of the gyroscope since it illustrates the effects we are interested in with a minimum of algebra. Since we drop the gyroscope down the z-axis, the gyroscope’s rest frame is an inertial frame throughout its descent. If the gyroscope precesses, then this means that inertial frames are rotating in the vicinity of the rotating mass!
The spin $s^\mu$ carried by the gyroscope travels on a geodesic. In the gyroscope’s locally inertial frame $u^\mu$ is purely time-like and the spin $s^\mu$ is purely space-like. Therefore, $\sum_\mu s^\mu u_\mu = 0$ all along the gyroscope’s path, which indicates that the spin parallel transports as the gyroscope falls. It therefore satisfies the parallel transport differential equation,

$$\frac{ds^\mu}{d\tau} + \sum_{\alpha \beta} \Gamma^\mu_{\alpha \beta} s^\alpha u^\beta = 0 \quad (B.18)$$

where $\Gamma^\mu_{\alpha \beta}$ are the Christoffel symbols implied by the metric Eq. B.17. Next we need various Christoffel symbols on the z-axis, $x = y = 0$ and $r = z$. Recall the formula for the Christoffel symbols,

$$\Gamma^\mu_{\alpha \beta} = \frac{1}{2} \sum_\rho g^{\mu \rho} \left( \partial_\alpha g_{\rho \beta} + \partial_\beta g_{\alpha \rho} - \partial_\rho g_{\alpha \beta} \right) \quad (B.19a)$$

which simplifies to first order, $g_{\alpha \beta} = \eta_{\alpha \beta} + h_{\alpha \beta}$,

$$\Gamma^\mu_{\alpha \beta} = \frac{1}{2} \left( \partial_\beta h^\mu_\alpha + \partial_\alpha h^\mu_\beta - \partial^\mu h_{\alpha \beta} \right) \quad (B.19b)$$

Eq. B.18 and B.19b appear complicated with many possible terms. However, let’s consider only non-relativistic motion of the gyroscope so $u^0 \approx c$ and $u^3 \approx 0$ to leading order in $1/c$. Then the possible non-trivial terms in Eq. B.18 are,

$$\frac{d}{d\tau}s^1 = -\Gamma^1_{10}s^1 - \Gamma^1_{20}s^2$$
$$\frac{d}{d\tau}s^2 = -\Gamma^2_{10}s^1 - \Gamma^2_{20}s^2 \quad (B.20)$$

where we chose the spin to lie initially in the plane transverse to the z axis. From Eq. B.19b and B.17, one computes,

$$\Gamma^1_{20} \approx \frac{2GJ}{c^3 r^3} \quad \Gamma^2_{10} \approx -\frac{2GJ}{c^3 r^3} \quad (B.21)$$

to leading order in $1/c$. $\Gamma^1_{10}$ and $\Gamma^2_{20}$ are $O(c^{-2})$ smaller and can be ignored. Eq. B.20 becomes

$$\frac{d}{d\tau}s^1 = -\frac{2GJ}{c^2 r^3} s^2, \quad \frac{d}{d\tau}s^2 = \frac{2GJ}{c^2 r^3} s^1 \quad (B.22)$$

which we recognize as simple precession.
\[
\frac{d}{d\tau} \vec{s} = \vec{\omega} \times \vec{s} \tag{B.23}
\]

with \(\vec{\omega} = \frac{2GJ}{c^2 r^3} \hat{\kappa}\): the spin \(\vec{s}\) rotates around the z-axis with an angular velocity given by \(\vec{\omega} = \frac{2GJ}{c^2 r^3}\).

Eq. B.23 illustrates the precession effect discovered by Lense and Thirring in 1918 for a special (and simple) setup. The Lense-Thirring effect [11] was measured by NASA’s Gravity Probe B satellite experiment. The web page for this experiment can be found linked to the previous Supplementary Lecture 5. The experimental discovery of the Lense-Thirring precession was a testament to accurate, sophisticated satellite technology. The Lense-Thirring effect is an order of magnitude smaller than geodesic precession that was also measured in the same experiment. See Special Topic J for more details.

### Special Topic C. Gravito-Magnetism

Linearized gravity can be approached as a problem in special relativity. A thorough discussion would follow the textbook's derivation of Maxwell equations from Gauss' Law, special relativity and locality. The final result would be the so-called Einstein-Maxwell equations. Their limitations are noteworthy: they are consequences of special relativity so they only work for first order effects in a background Minkowski space time. They reproduce the results of the previous section of this supplementary lecture, but do not provide a pathway to the non-linear effects that are essential for a good understanding of general relativity. However, they show that velocity dependent effects such as frame-dragging for weak gravitational fields are consequences of special relativity alone where velocity dependent forces, magnetic effects, follow from first principles and are not exclusive to electromagnetism.

Let’s continue to use the foundations and formalism of general relativity and work with \(h^{\alpha \beta}\). For a static, weak source of gravity we had [4],

\[
\Phi(\vec{x}) = -G \int \frac{\rho(\vec{y})}{|\vec{x} - \vec{y}|} d^3 y \\
A^i(\vec{x}) = -\frac{4G}{c^2} \int \frac{h^{0i}(\vec{y})}{|\vec{x} - \vec{y}|} d^3 y \tag{C.1}
\]

and \(h^{00} = h^{11} = h^{22} = h^{33} = \frac{2\Phi}{c^2}, h^{0i} = A^i / c\) and,
\[ ds^2 = \left(1 + \frac{2\Phi}{c^2}\right)c^2 dt^2 + \sum \frac{2A_i}{c} c dt dx^i - \left(1 - \frac{2\Phi}{c^2}\right)(dx^2 + dy^2 + dz^2) \] \hspace{1cm} (C.2)

Recall from the textbook discussion of general relativity that a test mass outside the source will follow a trajectory given by the geodesic equation,

\[
\frac{d}{d\tau} u^\sigma + \sum_{\alpha\beta} \Gamma^\sigma_{\alpha\beta} u^\alpha u^\beta = 0 \hspace{1cm} (C.3)
\]

where \( u^\sigma = dx^\sigma / d\tau \). We only want to treat the particle’s motion to first order so the first order expression Eq. B.19b for the Christoffel symbols applies,

\[
\Gamma^\mu_{\alpha\beta} = \frac{1}{2} \left( \partial_\beta h^\mu_\alpha + \partial_\alpha h^\mu_\beta - \partial^\mu h_\alpha\beta \right)
\]

We will only consider non-relativistic motion so,

\[
u^\sigma = \gamma (c, \vec{u}) \approx (c, \vec{u}) \hspace{1cm} (C.4)
\]

and \( \gamma \approx 1 \) to second order and \( \frac{d}{d\tau} \) can be replaced with \( \frac{d}{dt} \) to this accuracy in Eq. 30. Then it becomes,

\[
\frac{d^2}{dt^2} x^i \approx -\left( c^2 \Gamma^i_{00} + c \sum_j \Gamma^i_{0j} u^j \right) \hspace{1cm} (C.5)
\]

Next we need \( \Gamma^i_{00} \) and \( \Gamma^i_{0j} \) to first order,

\[
\Gamma^i_{00} = \frac{1}{2} \left( \partial_0 h^i_0 + \partial_0 h^i_0 - \partial^i h_00 \right) = -\frac{1}{c^2} \partial^i \Phi \hspace{1cm} (C.6a)
\]

\[
\Gamma^i_{0j} = \frac{1}{2} \left( \partial_j h^i_0 + \partial_0 h^i_j - \partial^i h_0j \right) = \frac{1}{2c} \left( \partial_j A^i - \partial^i A_j \right) \hspace{1cm} (C.6b)
\]

and we see the appearance of the “gravito-electric” field here, \( \vec{E} = -\vec{\nabla} \Phi \), and the “gravito-magnetic” field, \( \vec{B} = \vec{\nabla} \times \vec{A} \). Substituting into Eq. C.5 (and exercising care with signs!),

\[
\frac{d^2}{dt^2} \vec{x} \approx \vec{E} + \vec{u} \times \vec{B} \hspace{1cm} (C.7)
\]

to first order. This is the “gravito-Lorentz force law”. The magnetic term is the prize here: it is the \( O(1/c) \) correction to Newton’s law of gravity. A more elementary derivation can be achieved by copying the developments in Chapters 4-8 of the textbook beginning with Newton’s law of gravity instead of Coulomb’s law of electrostatics and making the obvious replacements.
of charges→mass, k→G and being especially careful with signs (gravity is always attractive and only positive masses exist). A major difference, however, is that the source of the electromagnetic field is the charge current \( J^\mu \) and the charge is Lorentz invariant. The energy-momentum tensor \( T^{\alpha \beta} \) is the source of gravity and the total energy is the zeroth component of a four vector and is \textit{not} a Lorentz invariant. These differences limit the electromagnetism-gravitational analogy to first order effects in the fields and to low velocities, \( v/c \ll 1 \).

Let’s consider another example of a non-relativistic source of gravity inspired from electrodynamics: a uniform mass current in the z-direction as shown in Fig. C.1, the gravitational analogue of a fixed, external current in a long wire.

![Figure C.1. Gravito-Magnetic field \( \vec{B} \) around a mass current in the z direction](image)

Let the velocity of the masses on the z axis be \( v \) and assume that their motion is non-relativistic, \( v/c \ll 1 \). Then,

\[
A^z = -\frac{4G}{c^2} v \int \frac{\rho(y)}{|x-y|} d^3y = \frac{4}{c^2} v \Phi
\]  

(C.8)
Then,

\[ B^x = \frac{\partial}{\partial y} A^z = \frac{4}{c^2} v \frac{\partial}{\partial y} \Phi = \frac{4}{c^2} v E^y \]

\[ B^y = \frac{\partial}{\partial x} A^z = -\frac{4}{c^2} v \frac{\partial}{\partial x} \Phi = \frac{4}{c^2} v E^x \]  

(C.9)

As shown in the figure, \( \vec{B} \) circulates around \( \rho \vec{u} \) following the right hand rule of electrodynamics. So, if the test particle has a positive velocity in the z direction, the \( \vec{u} \times \vec{B} \) term in Eq. C.7 points toward the wire and the gravito-magnetic force is attractive.

Another illustration we could do is to consider a gyroscope, the gravitational analogue of a magnetic moment in electrodynamics, place it in a uniform gravito-magnetic field \( \vec{B} \) and verify that its equation of motion predicts precession, Eq. B.23. This is the Lense-Thirring effect in this language. Our quantitative result of the previous section could be obtained here as well.

**Special Topic D. Rotating Stars, Black Holes and an Introduction to the Kerr Metric**

The textbook studied strong gravity in the context of spherically symmetric massive stationary objects. Their mass \( M \) determined their metric, and we obtained and studied the Schwarzschild metric and its black hole. In the case of the black hole we discovered an event horizon from which nothing could escape. According to the physics of the evolution of stars, black holes are the late stage evolutionary results of the burning of massive stars [9]. Since most stars rotate, we would expect most black holes to be characterized by their angular momentum \( J \) in addition to their mass \( M \). This is in fact the case. The exterior of rotating stars and black holes are described by another famous metric, the Kerr metric which was discovered by Roy Kerr in 1963 [12], forty seven years after the Schwarzschild solution.
The Kerr metric has some amazing features which figure prominently in astrophysics [4,9,10]. Unfortunately the Kerr metric is complicated and the analysis of its properties is too lengthy for a thorough, satisfying discussion here. We will only discuss some introductory properties to emphasize the new physics the Kerr metric presents us with. The ambitious student should consult the references [9,10].

We are interested in stationary metrics which are cylindrically symmetric. We suppose that we are considering a star which is rotating uniformly with a fixed axis in the z direction. The angle of rotation about the z axis will be denoted \( \varphi \). The two other spacelike coordinates will be labelled \( x^1 \) and \( x^2 \) and could be Cartesian coordinates or the cylindrical polar coordinates \( r \) (distance to the z-axis) and a polar angle \( \theta \). A general metric \( g_{\alpha\beta} \) will depend only on \( x^1 \) and \( x^2 \) because the cylindrical symmetry excludes dependence on \( \varphi \) and the stationary aspect of the problem excludes dependence on \( t \). Our experience with similar problems suggests that there will be two conserved quantities: energy, \( p_0 \), implied by the time independence of the problem and angular momentum, \( -p_\varphi = L \), implied by the \( \varphi \) independence of the problem. (See Appendix B for background and formalism.) The metric \( g_{\alpha\beta} = g_{\alpha\beta}(x^1, x^2) \) and the invariant line element will have the form [1,4],

\[
ds^2 = g_{tt} c^2 dt^2 + 2 g_{t\varphi} c dt d\varphi + g_{\varphi\varphi} d\varphi^2 + g_{rr} dr^2 + g_{\theta\theta} d\theta^2 \tag{D.1}
\]

The mixed term \( c dt d\varphi \) parametrizes the angular motion of the spinning object. Eq. D.1 can be written in a more suggestive form, taking inspiration from our analysis of rotating coordinate systems in lecture 2,

\[
ds^2 = \left( g_{tt} + \frac{g_{t\varphi}^2}{g_{\varphi\varphi}} \right) c^2 dt^2 + g_{\varphi\varphi} (d\varphi - \omega dt)^2 + g_{rr} dr^2 + g_{\theta\theta} d\theta^2 \tag{D.2}
\]

where \( \omega = -\frac{g_{t\varphi}}{g_{\varphi\varphi}} \). If we define \( \varphi' = \varphi - \omega t \) and note that \( \omega = \omega(r, \theta) \), then \( d\varphi' = d\varphi - \omega dt \) and the second term in Eq. D.2 becomes \( g_{\varphi\varphi} d\varphi'^2 \). If \( \varphi' \) is held fixed, then \( \omega(r, \theta) = \frac{d\varphi}{dt} \).

It is also handy to compute the contravariant components of the metric. Since there is a \( c dt d\varphi \) term, there is a \( 2 \times 2 \) matrix to invert in the \( t - \varphi \) sector. The algebra gives,

\[
g^{rr} = \frac{1}{g_{rr}}, \quad g^{\theta\theta} = \frac{1}{g_{\theta\theta}}, \quad g^{tt} = \frac{g_{t\varphi}}{D}, \quad g^{t\varphi} = \frac{\omega g_{t\varphi}}{D}, \quad g^{\varphi\varphi} = \frac{g_{tt}}{D} \tag{D.3}
\]
where \( \omega = - g_{t\phi} / g_{\phi\phi} \) again and \( D \) is the determinant, \( D = g_{tt} g_{\phi\phi} - g_{t\phi}^2 \). We will need these results in later developments.

In order to measure \( \omega(r, \theta) \) by doing experiments in the space time described by Eq. D.2, consider the motion of a test particle with vanishing angular momentum, \( p_{\phi} = 0 \), [4]. (Recall our discussion of particle motion in the Schwarzschild metric: its time independence implied that the covariant component \( p_0 \) is the conserved energy and the \( \phi - \) independence implied that the covariant component \( -p_{\phi} = L \) is the particle’s conserved angular momentum.) If \( p_{\phi} = 0 \), then the particle’s covariant components are, using the metric of Eq. D.1,

\[
p^\phi = \sum_{\mu} g^{\phi\mu} p_\mu = g^{\phi t} p_t, \quad p^t = \sum_{\mu} g^{t\mu} p_\mu = g^{tt} p_t
\]

But \( p^\mu = \frac{du^\mu}{d\tau} \), so \( p^t = \frac{dt}{d\tau} \) and \( p^\phi = \frac{d\phi}{d\tau} \). Therefore,

\[
\frac{d\phi}{dt} = \frac{p^\phi}{p^t} = \frac{g^{\phi t}}{g^{tt}} \tag{D.4}
\]

But reading off Eq. D.3, we have,

\[
\frac{d\phi}{dt} = \frac{g^{\phi t}}{g^{tt}} = \frac{(\omega g_{\phi\phi}/D)}{(g_{\phi\phi}/D)} = \omega(r, \theta) \tag{D.5}
\]

So, we learn that \( \omega(r, \theta) \) is the angular velocity of a particle with vanishing angular momentum in the \( z \) direction. Apparently when a particle is dropped straight along a radial spoke to the source, it acquires a non-zero angular velocity \( \omega(r, \theta) = g^{\phi t} / g^{tt} \). This effect is called “frame dragging” [4]. In light of the Lense-Thirring precession derived in the previous section above, this effect of rotation does not come as a complete surprise. It is indicative of a “gravitomagnetic” field generated by a rotating mass. We will expand on this point below. Since the particle dropped along a radial spoke travels on a geodesic, the particle’s rest frame is inertial. So, inertial frames in the vicinity of a rotating mass rotate with an angular velocity \( \omega(r, \theta) \).

Now consider rotating black holes. Recall from the discussion of Schwarzschild black holes in lecture 12, that coordinate singularities appeared at the Schwarzschild radius \( r_s = 2GM/c^2 \). At \( r = r_s \), \( g_{tt} \) vanished and \( g_{rr} \) diverged. These two conditions had two separate implications. Recall from our discussion of the redshift in time-independent environments, that when light is
emitted at frequency $v_E$ at fixed spatial coordinates $E$ and the light is then observed at fixed spatial coordinate $O$, its frequency $v_O$ is given by,

$$\frac{v_O}{v_E} = \sqrt{\frac{g_{tt}(E)}{g_{tt}(O)}}$$  \quad (D.6)

So, if $g_{tt}(E) \to 0$, then $v_O \to 0$ and the light has suffered an infinite redshift. (Eq. D.6 is reviewed and derived in Special Topic F below.) The Schwarzschild radius is a surface of infinite redshift. The Schwarzschild radius was also seen to be an event horizon which meant that the light cone at each point on the Schwarzschild surface lies entirely on one side of the surface and is tangent to the spatial surface at that point. This property implied that particles and light could cross it in one direction but not the other. This was guaranteed by having $g^{rr} = 0$, or equivalently $g_{rr} = \infty$ on the surface, called an “event horizon”.

In the Schwarzschild black hole the infinite redshift surface coincides with the event horizon. However, if the black hole is rotating this degeneracy is broken and some new, perplexing phenomena ensue. To understand these phenomena in general terms, consider a light ray emitted from a point of fixed spatial coordinates outside a rotating mass. If the light ray is pointed in the $\varphi$ direction it propagates along a null line element $ds^2 = 0$ parametrized by,

$$ds^2 = g_{tt}dt^2 + 2g_{t\varphi}dtd\varphi + g_{\varphi\varphi}d\varphi^2 = 0$$  \quad (D.7)

Dividing through by $dt^2$, we can solve for the angular velocity,

$$\frac{d\varphi}{dt} = -\frac{g_{t\varphi}}{g_{\varphi\varphi}} \pm \sqrt{\left(\frac{g_{t\varphi}}{g_{\varphi\varphi}}\right)^2 - \frac{g_{tt}}{g_{\varphi\varphi}}}$$  \quad (D.8)

This formula is particularly interesting on a surface of infinite redshift where $g_{tt} = 0$. Then the two solutions of Eq. D.8 reduce to,

$$\frac{d\varphi}{dt} = -\frac{2g_{t\varphi}}{g_{\varphi\varphi}} = 2\omega, \quad \frac{d\varphi}{dt} = 0$$  \quad (D.9)

In the first case the light ray was initially pointed in the direction of motion and in the second case it was opposed to the direction of rotation. In that case “frame dragging” is so intense that the light ray is stationary!. If the light ray were replaced by a massive particle, then in either case the particle rotates with the source. Recall that in the case of the Schwarzschild black hole
stationary behavior is not possible inside the event horizon. (This particular result can be obtained simply, as we noted in lecture 12. We imagined that the particle was at rest at fixed \((r, \theta, \varphi)\) so \(u^\mu = (u^t, 0, 0, 0)\). But \(u^\mu\) has fixed length, \(u \cdot u = \Sigma \mu u^\mu u_\mu = c^2\), which implies \(g_{tt}(u^t)^2 = c^2\). This relation is impossible if \(g_{tt} < 0\) which leads to our conclusion that stationary motion is impossible inside the Schwarzschild black hole.)

It is finally time to work with the Kerr metric [12]. In Schwarzschild-like coordinates \((t, r, \theta, \varphi)\) the metric describing a stationary axi-symmetric source is,

\[
d s^2 = \left(1 - \frac{2\mu}{r}\right)c^2 dt^2 + \frac{4\mu acr \sin^2 \theta}{\rho^2} c dt d\varphi - \frac{\rho^2}{\Delta} dr^2 - \rho^2 d\theta^2 - \left(r^2 + \frac{2\mu a^2 \sin^2 \theta}{\rho^2}\right) \sin^2 \theta d\varphi^2
\]

where \(\mu = GM/c^2\) and \(a = J/Mc\) and the functions \(\rho^2\) and \(\Delta\) read,

\[
\rho^2 = r^2 + a^2 \cos^2 \theta \quad \Delta = r^2 - 2\mu r + a^2
\]

Eq. D.10a is considerably more complicated than the Schwarzschild metric and its derivation and analysis are challenging. However, one can check that if the rotation of the source vanishes, \(J \rightarrow 0\), then the metric reduces to the Schwarzschild metric, and when \(r \rightarrow \infty\) it reduces to Minkowski space time.

Consider the limit of the Kerr metric when \(\frac{GM}{c^2} \rightarrow 0\) holding \(a = J/Mc\) fixed [4]. Then Eq. D.10 reduces to,

\[
d s^2 = c^2 dt^2 - \frac{\rho^2}{r^2 + a^2} dr^2 - \rho^2 d\theta^2 - (r^2 + a^2) \sin^2 \theta d\varphi^2
\]

The Minkowski metric in Cartesian coordinates is hiding here. Define,

\[
\begin{align*}
x &= \sqrt{r^2 + a^2} \sin \theta \cos \varphi \\
y &= \sqrt{r^2 + a^2} \sin \theta \sin \varphi \\
z &= r \cos \theta
\end{align*}
\]

Then algebra gives,

\[
ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2
\]
as claimed. Now we see, surprisingly(!), that in this limit the value \( r = 0 \) corresponds to a disc of radius \( a \) in the equatorial plane \( (z = 0, \theta = \pi/2) \), \( x^2 + y^2 = a^2 \). Only in the limit \( a \to 0 \) do the coordinates \((r, \theta, \varphi)\) become standard spherical polar coordinates of three dimensional Euclidean space. In general when \( M \) and \( J \) are non-zero, \((r, \theta, \varphi)\) are not the familiar variables we have used before!

To get a global understanding of the Kerr metric let’s find its surfaces of infinite redshift and its event horizons [4]. The surfaces of infinite redshift satisfy \( g_{tt} = 0 \). We read off Eq. D.10,

\[
g_{tt} = 1 - \frac{2\mu r}{r^2 + a^2 \cos^2 \theta} = 0
\]  

(D.14)

So, the \( r \) coordinates where \( g_{tt} = 0 \) are,

\[
r_{inf}^{\pm} = \frac{GM}{c^2} \pm \sqrt{\left(\frac{GM}{c^2}\right)^2 - \left(\frac{J}{Mc}\right)^2 \cos^2 \theta}
\]  

(D.15)

These are sensible solutions as long as \( J \) is not too large for a given \( M \), \( J^2 < \frac{G^2M^4}{c^2} \).

Let’s compare these surfaces with the event horizons where \( g_{rr} = \infty \),

\[
g_{rr} = \frac{\rho^2}{\Delta} = -\frac{r^2 \left(\frac{J}{Mc}\right)^2 \cos^2 \theta}{r^2 - 2\frac{GM}{c^2} \left(\frac{J}{Mc}\right)^2}
\]  

(D.16)

Setting the denominator in Eq. D.16 to zero, we solve for \( r \),

\[
r_{\pm} = \frac{GM}{c^2} \pm \sqrt{\left(\frac{GM}{c^2}\right)^2 - \left(\frac{J}{Mc}\right)^2}
\]  

(D.17)

Unlike the Schwarzschild geometry, in the Kerr geometry the non-vanishing \( J \) breaks the coincidences of the event horizons and the surfaces of infinite redshift: we see that \( r_+ < r_{inf}^{+} \) except at the poles! This inequality leads to some very curious and important astrophysics.

The Kerr metric is also singular when \( \rho^2 = r^2 + \left(\frac{J}{Mc}\right)^2 \cos^2 \theta = 0 \). This is a real geometric singularity: the Riemann curvature diverges there. To achieve \( \rho^2 = 0 \) we need \( r = 0 \) and \( \theta = \pi/2 \). We learned above that \( r = 0 \) represents a disc of coordinate radius \( a = J/Mc \) in the equatorial plane. Note that \( r_{inf}^{-} = 0 \) at \( \theta = \pi/2 \). We plot the singular surfaces of the Kerr
metric in Fig. D.1.

We note that \( r_{\inf +} \geq r_+ > r_- \geq r_{\inf -} \). The region between \( r_{\inf +} \) and \( r_+ \), called the “ergosphere”, is particularly important in astrophysics. In this region where \( g_{tt} < 0 \), an observer cannot maintain a fixed position at \((r, \theta, \varphi)\) as we described above. However, we can consider an observer rotating in the same direction as the black hole [4]. In that case,

\[
u^\mu = u^t (1,0,0,\Omega)
\]  

(D.18)

where \( \Omega = d\varphi/dt \) is the observer’s angular velocity as measured by another observer at large \( r \).

Since we must have \( u \cdot u = \Sigma \mu u^\mu u_\mu = c^2 \), we can evaluate,

\[
u \cdot u = g_{tt}(u^t)^2 + 2g_{t\varphi}u^tu^\varphi + g_{\varphi \varphi}(u^\varphi)^2 = (u^t)^2(g_{tt} + 2g_{t\varphi}\Omega + g_{\varphi \varphi}\Omega^2) = c^2
\]  

(D.19)

This implies that \( g_{tt} + 2g_{t\varphi}\Omega + g_{\varphi \varphi}\Omega^2 > 0 \) which vanishes at,
\[
\Omega_{\pm} = -\frac{g_{t\varphi}}{g_{\varphi\varphi}} \pm \sqrt{\left(\frac{g_{t\varphi}}{g_{\varphi\varphi}}\right)^2 - \frac{g_{tt}}{g_{\varphi\varphi}}} = \omega \pm \sqrt{\omega^2 - \frac{g_{tt}}{g_{\varphi\varphi}}} \quad (D.20)
\]

where we identified \( \omega = -\frac{g_{t\varphi}}{g_{\varphi\varphi}} \) from Eq. D.4 and D.5. Note that on the surface of infinite redshift, \( r_{\text{inf} \pm} \), where \( g_{tt} = 0 \), \( \Omega_+ = 2\omega \) and \( \Omega_- = 0 \). Another interesting situation occurs at \( r = r_+ \), the outer event horizon. There \( \Delta = r^2 - 2\frac{GM}{c^2} r + \left(\frac{J}{Mc}\right)^2 = 0 \) and \( \omega^2 = \frac{g_{tt}}{g_{\varphi\varphi}} \), so \( \Omega_{\pm} = \Omega = \omega \), so all observers on circular orbits rotate with a common \( \Omega = \omega(r, \theta) = \frac{J/M}{2\frac{GM}{c^2} r_+} = \frac{Jc^2}{2GM^2 r_+} \).

This value of \( \omega \) is the maximum possible in the ergosphere.

Kerr black holes can interact with their environments and exchange energy and angular momentum \([4,9,10]\). This is so because particles can enter and leave the surface \( r_{\text{inf} +} \) and interact with the Kerr black hole. For example, there are Penrose processes where energy and angular momentum can be extracted from the Kerr black hole. In this case a particle enters the ergosphere, decays into two particles and one of them propagates through the event horizon \( r_+ \) while the second one is ejected back through \( r_{\text{inf} +} \). The Penrose process can work to extract energy and angular momentum out of the rotating black hole because: 1. particles can enter and leave through the surface of infinite redshift at \( r_{\text{inf} +} \), and 2. \( g_{tt} < 0 \) inside the ergosphere so time-like and space-like components of four vectors like \( p^\mu \) are effectively interchanged. The reader should consult the references for the curious details \([9,10]\).

The extreme dynamics within the ergosphere creates important and dramatic astrophysical phenomena. Material will be attracted to the Kerr black hole. It will form accretion discs as it penetrates the surface of infinite redshift and falls toward the event horizon \( r_+ \). The material will be dragged and will co-rotate with \( J \). Realistic calculations and computer simulations show that turbulent viscosity among the in-falling particles cause them to lose angular momentum and to radiate light thus losing gravitational energy. The fraction of the rest mass energy that can be released this way can range as high as \( (\sqrt{3} - 1)/\sqrt{3} \approx 42\% \) in principle and around 30\% for realistic astrophysical black holes. This is a huge energy efficiency! Recall that the efficiency of nuclear explosions is typically a fraction of a percent. The radiation of material falling into a rotating black hole illustrates the impressive gravitational binding energy in these strong fields. The reader is encouraged to learn more following the references \([9]\).
The first picture of an ergosphere of a rotating supermassive black hole was obtained by the Event Horizon Telescope in 2019 and announced on April 10, 2019.

Fig. D.2 The image of a Black Hole, from the Event Horizon Telescope.

Here is the [web site](#) of the announcement of this grand discovery where you can learn more about the general relativity and astrophysics behind the image.

**Special Topic E. Newton in Orbit around a Schwarzschild Black Hole**

Throughout this and other lectures we have imagined a spaceship maintaining a fixed position outside a black hole. How large a thrust is required? Will such adventures ever be practical? What kind of propulsion mechanism would be required? This lecture investigates these pressing questions.
Suppose that Isaac Newton’s great great great ... grand daughter was a Star Fleet commander who is studying the outer region of a black hole. She needs to know the thrust required to maintain her starship at a constant coordinate $r$ so she can run experiments. She knows general relativity like every starship commander, but she also likes to understand physics in the terms of her great great great...grand father, Sir Isaac.

When the starship is at fixed $(r, \theta, \varphi)$, she knows that the starship must have a radial thrust to counter the ship’s natural free fall on a radial geodesic. She likes to think in terms of four forces, in deference to her heritage, rather than geometry. She knows that the attraction is radial and has magnitude $f^r = GMm/r^2$ since the starship is holding its position fixed. But since she knows general relativity, she realizes that this is not the force her famous great great great...grand father would have equated to the required thrust to hold the starship of mass $m$ at a fixed position outside the black hole of mass $M$. He would have done experiments using an orthonormal set of coordinate basis vectors [9]. He would want $f^r$ calculated with respect to such a basis. The basis vector in the $r$ direction used in the Schwarzschild coordinate system, is,

$$e_r \cdot e_r = (0,1,0,0) \cdot (0,1,0,0) = g_{11}(r)$$

(E.1)

Here $g_{11}(r) = \left(1 - \frac{2GM}{c^2r}\right)^{-1}$ and it accounts for the spatial curvature outside the black hole. So, the normalized vector Commander Newton wants is,

$$\frac{1}{\sqrt{g_{11}(r)}} e_r = \left(0, \frac{1}{\sqrt{g_{11}(r)}}, 0, 0\right)$$

(E.2)

Finally, in this basis the radial force which balances the attraction of gravity is,

$$\sqrt{g_{11}(r)}f^r = \left(1 - \frac{2GM}{c^2r}\right)^{-1/2} \frac{GMm}{r^2}$$

(E.3)

Now the commander knows the thrust required to maintain the starship at constant $r$, safe from falling through the Schwarzschild horizon and being lost to the outer region where the planet Earth and Cambridge University reside. The geometric factor $\left(1 - \frac{2GM}{c^2r}\right)^{-1/2}$ shows that an arbitrarily large thrust is required as $r$ approaches the Schwarzschild radius from above!

**Starship’s Log**: The Commander knew that $f^r = GMm/r^2$ in Schwarzschild coordinates. Let’s derive this from scratch. Recall the force in a curved geometry from lecture 8,
\[ f^\mu = m \left( \frac{d^2 x^\mu}{d\tau^2} + \sum_{\alpha \beta} \Gamma^\mu_{\alpha \beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} \right) \]  

(E.4)

To apply this formula we need the four velocity of the stationary starship. Its four velocity \( u^\mu \) satisfies the condition \( u \cdot u = \sum_{\alpha \beta} g_{\alpha \beta} u^\alpha u^\beta = c^2 \). But the starship is stationary (and not accelerating) so the only non-zero component of \( u^\mu = dx^\mu / d\tau \) is \( u^0 \). Then the normalization condition reduces to \( u \cdot u = g_{00} u^0 u^0 = c^2 \). But \( g_{00} = \left(1 - \frac{2GM}{c^2 r}\right) \) for the Schwarzschild metric, so \( u^0 = c \left(1 - \frac{2GM}{c^2 r}\right)^{-1/2} \). Now we substitute into Eq. E.4,

\[ f^r = m \Gamma^r_{00} u^0 u^0 = m c^2 \Gamma^r_{00} \left(1 - \frac{2GM}{c^2 r}\right)^{-1} \]  

(E.5)

Finally we look up the Christoffel symbol for the Schwarzschild metric from lecture 12,

\[ \Gamma^r_{00} = \left(1 - \frac{2GM}{c^2 r}\right) \frac{GM}{c^2 r^2} \]

Substituting \( \Gamma^r_{00} \) into Eq. E.5 we find the result our Commander knew,

\[ f^r = \frac{GMm}{r^2} \]  

(E.6)

So, when the rocket is stationary, the force provided by its engines is balanced by the “force” provided by the curvature of space time expressed by the \( \Gamma^\mu_{\alpha \beta} \) in Eq. E.4.

**Special Topic F. Symmetries, Conservation Laws, Exact Redshift**

**Formula for Static Gravity**

We have discussed the gravitational redshift from several perspectives for weak gravitational fields in lecture 4. Now that we have the Schwarzschild metric which describes the time independent static gravitational field outside a mass M, we can do better. As we already noted in lecture 12, the Schwarzschild metric has several special symmetries which simplify the propagation of masses in the environment of \( M \). It is time independent, and 2. It is \( \phi \) independent. We have seen that these symmetries lead to energy and angular momentum
conservation and we used these results to reduce the calculation of the trajectories of test masses to a one-dimensional mechanics problem in the variables \( r \) and \( t \) much like one does in Newton’s world.

Recall the geodesic equation for a covariant four vector,

\[
    m \frac{dp_\mu}{d\tau} = \sum_{\alpha\beta} \Gamma^\alpha_{\beta\mu} p^\beta p^\alpha
\]

from lecture 7. The right hand side of this equation can be simplified using the explicit formula for the Christoffel symbol written in terms of the metric,

\[
    \sum_{\alpha\beta} \Gamma^\alpha_{\beta\mu} p^\beta p^\alpha = \frac{1}{2} \sum_{\alpha\beta\gamma} g^{\alpha\gamma} \left( \partial_\beta g_{\gamma\mu} + \partial_\gamma g_{\beta\mu} - \partial_\mu g_{\beta\gamma} \right) p^\beta p^\alpha
\]

\[
    = \frac{1}{2} \sum_{\beta\gamma} \left( \partial_\beta g_{\gamma\mu} + \partial_\gamma g_{\mu\beta} - \partial_\mu g_{\gamma\beta} \right) p^\beta p^\gamma = \frac{1}{2} \sum_{\beta\gamma} (\partial_\mu g_{\gamma\beta}) p^\beta p^\gamma
\]

where we used the symmetric character of \( p^\beta p^\gamma \) to simplify the algebra in the last step. In summary,

\[
    m \frac{dp_\mu}{d\tau} = \frac{1}{2} \sum_{\beta\gamma} (\partial_\mu g_{\gamma\beta}) p^\beta p^\gamma \tag{F.1}
\]

Note that if we choose \( \mu = 0 \), then \( m dp_0/d\tau \), the right hand side, vanishes since \( g_{\gamma\beta} \) is time independent. We learn that \( p_0 \), the covariant component, is conserved. We recognize this argument as a generalization of the energy conservation argument in Newtonian mechanics. Here the metric has replaced Newton’s potential, but the logic in the derivation of a conservation law is the same.

It is useful to write this conservation law in an explicitly covariant form. The \( \mu = 0 \) component is special here so write it as the zeroth component of a contravariant vector [9] in the Schwarzschild coordinate system,

\[
    K^\mu = (1,0,0,0)
\]

Then the conserved \( p_0 \) can be written,
\[ p_0 = \sum_{\mu} K^\mu p_\mu = \sum_{\mu} K_\mu p^\mu \]

\( \sum_{\mu} K_\mu p^\mu \) is a scalar quantity which is conserved along the particle’s path. Call it \( S \) and write,

\[ S = \sum_{\mu} K_\mu \frac{dx^\mu}{d\tau} = \text{constant} \]

Since \( K^\mu = (1,0,0,0) \), \( K_0 = \sum_{\mu} g_{0\mu} K^\mu = g_{00} K^0 = \left(1 - \frac{2GM}{c^2 r}\right) \) for the Schwarzschild metric, and the conserved quantity \( S \) is,

\[ S = \sum_{\mu} K_\mu \frac{dx^\mu}{d\tau} = g_{00} \frac{dx^0}{d\tau} = c \left(1 - \frac{2GM}{c^2 r}\right) \frac{dt}{d\tau} \]

Now we are ready for our application to the strong gravity redshift. Suppose there is an observer at coordinate \( r \) and she observes a particle (or wave) of four momentum \( p^\mu \). If the observer has four velocity \( u^\mu \), then the energy of the particle in the observer’s rest frame can be written in an invariant form as,

\[ E = \frac{1}{c} u \cdot p \]

Recall how this formula works. In special relativity, or in a freely falling frame in a gravitational field, \( u^\mu = dx^\mu / d\tau \). In Cartesian coordinates, \( u^\mu = (dct / d\tau, dx^i / d\tau) = p^\mu / m = (E/mc, \vec{p}/m) = \gamma(1, \vec{v}) \). In the observer’s rest frame \( u^\mu = (c, 0) \) and \( E = \frac{1}{c} u \cdot p = p_0 \) and all is well. \( u^\mu / c \) has unit length, \( u \cdot u = \sum_{\alpha\beta} g_{\alpha\beta} u^\alpha u^\beta = c^2 \). It’s length is a scalar quantity independent of the coordinate system. In addition, it is convenient to choose a set of four orthonormal basis vectors. For example, if the observer is at coordinate \( r \) outside the star of mass \( M \) and she is at rest, then

\[ u \cdot u = \sum_{\alpha\beta} g_{\alpha\beta} u^\alpha u^\beta = g_{00} u^0 u^0 = \left(1 - \frac{2GM}{c^2 r}\right) u^0 u^0 = c^2 \]

And,
\[ u^\mu = \left( c \left( 1 - \frac{2GM}{c^2r} \right)^{-1/2}, 0,0,0 \right) \]

where we see the Schwarzschild radius “singularity” in this coordinate system. The observer measures the energy of the particle which is moving freely in the gravitational field at position \( x^\mu(r) \),

\[ E = \frac{1}{c} u \cdot p = \frac{1}{c} \sum_{a\beta} g_{a\beta} u^a p^\beta = \frac{m}{c} \sum_{a\beta} g_{a\beta} u^a \frac{dx^\beta}{d\tau} = \frac{m}{c} g_{00} u^0 \frac{dt}{d\tau} \]

But \( \frac{dt}{d\tau} \) is related to the conserved quantity \( S \), \( S = g_{00} \frac{dx^0}{dt} = c \left( 1 - \frac{2GM}{c^2r} \right) \frac{dt}{d\tau} \), so

\[ E = \frac{m}{c} S u^0 = m \left( 1 - \frac{2GM}{c^2r} \right)^{-1/2} S \]

Therefore, the energy of the particle depends on the \( r \) coordinate of the observer,

\[ E(r) \sim (g_{00}(r))^{-1/2} = \left( 1 - \frac{2GM}{c^2r} \right)^{-1/2} \]

In the case of the photon, \( E = \hbar \omega \), so if the photon is observed at two subsequent \( r \) values, \( r_1 \) and \( r_2 \), then

\[ \frac{\omega_2}{\omega_1} = \left( \frac{g_{00}(r_1)}{g_{00}(r_2)} \right)^{1/2} = \left( \frac{1 - \frac{2GM}{c^2r_1}}{1 - \frac{2GM}{c^2r_2}} \right)^{1/2} \]

And we note that the photon is redshifted as it propagates away from the star.

If the field is weak, \( GM/c^2r \ll 1 \), then we have to first order,

\[ \frac{\omega_2}{\omega_1} \approx 1 - \frac{GM}{c^2r_1} + \frac{GM}{c^2r_2} = 1 + \frac{1}{c^2} \left( V(r_1) - V(r_2) \right) \]

where \( V(r) = -GM/r \) is the Newtonian potential, as discussed in lecture 4.

There is one point we have been cavalier about here. In the case of the massless photon we cannot use the proper time \( \tau \) to parametrize its path since \( d\tau^2 = 0 \) along a light ray. So, we
should redo the derivation with another parameter to track the path. However, the parameter cancels out in the ratio $\omega_1/\omega_2$ so this point leaves the final result unchanged.

**Special Topic G. An Accelerating Clock $C'$ and Hyperbolic Motion**

Consider a clock $C'$ accelerating in an inertial frame $S$ as shown in Fig. G.1. Suppose that $C'$ has a velocity $v(t)$ at the time $t$ measured in the frame $S$ and let $S'$ be an inertial frame with a fixed velocity that matches $v(t)$. Then $C'$ is instantaneously at rest in $S'$ at time $t$. The frame $S'$ allows us to use Lorentz Transformations to transform the kinematics of $C'$ to the frame $S$ which is a conventional inertial frame where we make measurements following the rules of special relativity.
First, what would an observer at rest in $S$ measure the acceleration of $C'$ to be? In Chapter 7 of the textbook the transformation law for acceleration was derived. Recall that for the component of the acceleration in the direction of $\mathbf{v}(t)$, call it $a_x$, we had Eq. 7.5, and the textbook,

$$a_x = \frac{a'_x}{\gamma^3(1 + u'_x/c^2)^3}$$

G.1

In this application $a'_x = g$, a fixed, time independent, proper acceleration, and $u'_x = 0$ since $C'$ is instantaneously at rest in $S'$. Also, $\gamma = (1 - v(t)^2/c^2)^{-1/2}$, as always, but with the twist that it is time dependent since $v$ is. So, Eq. G.1 reduces to,

$$a_x = \frac{dv(t)}{dt} = \frac{g}{\gamma^3}$$

G.2

which can be written in a form that can be integrated

$$dt = \frac{1}{g}\gamma^3 dv$$

G.3
Instead of doing that, let’s consider the relation between the time on the accelerating clock and the time on a clock at rest at the origin in the inertial frame S. These two time intervals are related by time dilation,

\[ d\tau = \frac{1}{\gamma} dt \] \hspace{1cm} G.4

Eq. G.3 and G.4 can be combined,

\[ d\tau = \frac{1}{g} \gamma^2 dv \] \hspace{1cm} G.5

which can be integrated conveniently,

\[ \tau = \int d\tau = \frac{1}{g} \int \frac{dv}{1 - v^2/c^2} = \frac{1}{g} \int \frac{dv}{(1 + v/c)(1 + v/c)} = \frac{1}{2g} \int \left( \frac{1}{1 + v/c} - \frac{1}{1 - v/c} \right) dv \]

So,

\[ \tau = \frac{c}{2g} \ln \left( \frac{1 + v/c}{1 - v/c} \right) \] \hspace{1cm} G.6

which can be solved for \( v(\tau) \),

\[ v(\tau) = c \frac{e^{2g\tau/c} - 1}{e^{2g\tau/c} + 1} = c \frac{e^{g\tau/c} - e^{-g\tau/c}}{e^{g\tau/c} + e^{-g\tau/c}} \]

\[ v(\tau) = c \tanh(g\tau/c) \] \hspace{1cm} G.7

which shows very elegantly that \( v(\tau) \) begins at zero and approaches the speed limit \( c \) as \( g\tau/c \) becomes large.

Now combine Eq. G.7 with the time dilation equation Eq. G.4 to relate \( t \) and \( \tau \) directly. Substituting Eq. G.7 into \( \gamma^{-1} \) gives,

\[ \gamma = \left( 1 - \frac{v^2}{c^2} \right)^{-1/2} = (1 - \tanh^2(g\tau/c))^{-1/2} = \cosh(g\tau/c) \] \hspace{1cm} G.8

So Eq. G.4, time dilation, can be written just in terms of \( t \) and \( \tau \),

\[ dt = \cosh(g\tau/c) \, d\tau \] \hspace{1cm} G.9

which can be integrated,

\[ gt/c = \sinh(g\tau/c) \] \hspace{1cm} G.10

We learn that there is a simple and elegant hyperbolic relation between time measurements in the accelerating frame and the inertial frame!
Finally, the position of the clock $C'$ in the inertial frame $S$ can be found using Eq. G.6 and G.9,

$$x = \int v \, dt = \int c \tanh(g\tau/c) \cosh(g\tau/c) \, d\tau = c \int \sinh(g\tau/c) \, d\tau$$

$$x = \frac{c^2}{g} [\cosh(g\tau/c) - 1] \quad \text{G.11a}$$

If we had chosen the initial condition, $x(t = \tau = 0) = c/g^2$, as was done in Supplemental Lecture 3, then,

$$x = \frac{c^2}{g} \cosh(g\tau/c) \quad \text{G.11b}$$

Now identify Lorentz invariant “hyperbolic motion” that was found in the discussion of Rindler Space in Supplemental Lecture 3,

$$x^2(\tau) - c^2 t^2(\tau) = \frac{c^4}{g^2} \quad \text{G.12}$$

where Eq. G.10 and G.11b were used to obtain Eq. G.12.

It is interesting to recall from Supplemental Lecture 3 and the discussion in the textbook that there is a simple physical realization of hyperbolic motion at a constant proper acceleration $g$: put a charge $e$ of mass $m$ into a uniform electric field $E$. In that case $g = eE/m$. The point is that the electric field is unchanged by boosts along its direction, so $g = eE/m$ is the same in all the instantaneous rest frames of the charged particle as it accelerates. Therefore, the initial hypothesis, that $g$, the proper acceleration, is fixed and time independent, is not so outrageous (or contrived) after all!

Note that the discussion above produces many of the relations in our earlier discussions of accelerating reference frames, Rindler Space, in Supplemental Lecture 3. But the emphasis here is on the quantitative relation between the accelerating and the inertial clocks, $gt/c = \sinh(g\tau/c)$. Fig. G.2 shows $g\tau/c$ plotted against $gt/c$ for $g\tau/c$ ranging from 1 to 9. Since $\sinh x \approx \frac{1}{2} e^x$ for $x \gg 1$, the disparity between the times on the inertial clock and accelerating clock can be astonishing,
Of course this disparity follows from the fact that \( v(\tau) \to c \) as \( g\tau/c \) becomes large and the time dilation factor grows without bound. In fact, for \( g\tau/c \gg 1 \),

\[
v(\tau) = c \tanh(g\tau/c) \approx c\left(1 - 2e^{2g\tau/c} + 2e^{-4g\tau/c} \pm \ldots\right) \approx c\left(1 - \frac{2}{(g\tau/2c)^2} + \frac{2}{(g\tau/2c)^4} \pm \ldots\right) \quad \text{G.13a}
\]

and

\[
\ln(g\tau/c) \approx \frac{g\tau}{c} - \ln 2 \quad \text{G.13b}
\]

for \( g\tau/c \gg 1 \).

**The Twins Once Again....**
The hyperbolic relation Eq. G.10 suggests an extreme example of the twin paradox. To illustrate the twin paradox, we need to put the clock $C'$ on a closed path. Then an inertial clock located at the origin of frame $S$ and the clock $C'$ will have the same positions at the beginning and end of the journey and they can compare transit times locally. This is easy to arrange. Suppose that $C'$ travels along the positive $x$ axis at proper acceleration $g$ for proper time $\tau$ and arrives at $(\tau)$, Eq. G.11, with velocity $v(\tau)$, Eq. G.7. Then change $g$ to $-g$ for another interval of $\tau$. $C'$ will decelerate relative to $S$ and, after this second interval of proper time $\tau$ has elapsed, will be at position $2x(\tau)$ with a vanishing velocity relative to $S$, as shown in Fig. G.3. Next, let $C'$ continue to accelerate in the $x$ direction at proper acceleration $-g$ for a third interval of $\tau$. Then $C'$ will find itself at distance $d(\tau)$ again with a velocity $-v(\tau)$. Finally, if $C'$ accelerates at a rate $+g$ for a fourth and final proper time interval of $\tau$, then it will come to rest at the origin of $S$ where it can compare times with the original stay-at-home clock. The clock $C'$ has a reading of $4\tau$ which should be compared to $4t$, the time that has passed at the origin of $S$, and $\tau$ and $t$ are related by the hyperbolic expression of Eq. G.10. Inspecting Fig. G.3, we see that the velocity of $C'$ is a smooth, continuous function, unlike the path used in the original discussion of the twin paradox in the textbook, where the traveler’s velocity jumped abruptly from a fixed $+v$ to $-v$ at the midpoint of the closed journey where the traveler jumped from an outgoing spaceship to an incoming one. In this case, if we choose $g\tau/c \gg 1$, then the values of Fig. G.2 apply: the accelerating traveler ages an exponentially smaller amount than her stay-at-home relative!
Fig. G.3. Closed spatial path of the accelerating traveler
Special Topic H. Uniform Proper Acceleration

Fig. H.1 Hyperbola of constant proper acceleration and line of constant $\tilde{t}$.

Now let’s return to Eq. G.12, the invariant hyperbola, $x^2(t) - c^2 \ell^2(t) = \frac{c^4}{g^2}$, the path in space time shown in Fig.H.1 taken by an accelerating rocket carry the clock $C'$. The invariant hyperbola has a simple interpretation. Imagine the rocket beginning its flight at $t = 0$ from an initial position on the $x$ axis of $x_0 = c^2 / g$. At the same time launch a light ray toward the rocket but starting at the origin. Will the light ray catch and overtake the rocket? The light ray travels at the speed limit from the origin, but the rocket accelerates from a starting position at $x_0 = c^2 / g$. (Imagine that $\alpha$ is a “typical” acceleration like the acceleration at the surface of the earth, 9.80 m/s$^2$, and verify that $x_0$ is a huge distance.) The outcome of the race is clear from the
Minkowski diagram: the light ray never catches the rocket! We are interested in the distance between the light ray and the rocket, \(x(t) - ct\), and this can be read off the hyperbola,

\[x^2 - c^2t^2 = (x - ct)(x + ct) = c^4/g^2 \quad \text{H.1}\]

We need \(x\) as a function of \(t\) in order to simplify Eq. H.1. From Eq. F.10 and F.11b we find,

\[x = \frac{c}{g^2} \cosh(\frac{g\tau}{c}) = \frac{c}{g^2} \sqrt{1 + \sinh^2(\frac{g\tau}{c})} = \frac{c}{g^2} \sqrt{1 + (\frac{gt}{c})^2} \quad \text{H.2a}\]

So, we can solve Eq. H.1 for \(x - ct\) and find,

\[x - ct = \frac{c^4/g^2}{x + ct} = \frac{c^4/g^2}{ct\left(\sqrt{1+(c/ct)^2}+1\right)} \sim \frac{1}{2} \frac{c^2}{g} \left(\frac{c}{gt}\right) \quad \text{H.2b}\]

where we took \(gt/c \gg 1\) in the last expression to calculate that asymptotically \((x + ct) \sim t^{-1}\): so the light ray approaches but never catches the rocket. Note that the light ray in Fig. 3 is the asymptote of the hyperbola. The hyperbola approaches the light ray in the limit \(g \to \infty\). In this sense light propagation occurs with a “divergent proper acceleration”. Another way to express this result is to note that boosts translate points on the hyperbola. So, the point \(B\) in Fig. 3 can be boosted to point \(A\) and we learn that in the instantaneous rest frame of the rocket, the invariant distance-squared to the light ray is \(c^4/g^2\) for all \(t\).

**Special Topic I. An Accelerating Coordinate System: The Rindler Wedge**

Now let’s get more ambitious and use what we have learned to make an accelerating grid, an accelerating reference frame where we could run experiments and make space-time measurements. Eq, G.12 suggests that we consider the class of hyperbolas,

\[x^2 - c^2t^2 = (\bar{x} + c^2/g)^2 \quad \text{(I.1)}\]

for \(\bar{x} > 0\) as shown in Fig. I.1. The \(\bar{x}\) axis lies on the \(x\) axis at \(t = 0\).
Instead of thinking of each hyperbola as the space-time path of a particular rocket, we can think of $\tilde{x}$ as measuring the distance off the floor of a huge “mother” ship with markings of increasing $\tilde{x}$ as shown in Fig. I.2.
The mother ship could be taken as long as one likes. It is also convenient to put synchronized clocks (synchronized in the mother ship’s instantaneous rest frame) at each marker $\tilde{x}$ telling time $\tilde{t}$.

To better understand the accelerating grid $(c\tilde{t}, \tilde{x})$ let’s differentiate Eq. I.1 with respect to $t$ for fixed $\tilde{x}$ and $g$,

$$2x \frac{dx}{dt} - 2c^2 t = 0 \quad (I.2a)$$

Or, identifying $dx/dt = v$,

$$t = \frac{v}{c^2} x \quad (I.2b)$$

But the Lorentz transformation to the ship having instantaneous velocity $v(t)$ reads,

$$\tilde{x} = \gamma (x - vt)$$

$$\tilde{t} = \gamma \left( t - \frac{v}{c^2} x \right)$$

So, the line $t = \frac{v}{c^2} x$, Eq. I.2.b, is a line of constant $\tilde{t}$. These are rays from the origin, the $\tilde{x}$ axes, at velocity $v$. We recognize this as an example of the Relativity of Simultaneity emphasized in
Chapter 2 and 3 of the textbook. A line of constant $\tilde{t}$ is shown in Fig. I.3 where we also show that this is a line of constant $dx/dt = v$, the tangents to the hyperbolas. This shows that all the points $\tilde{x}$ along the mother ship have the same velocity at any given time $\tilde{t}$ in the mother ship’s rest frame. So, the mother ship accelerates without internal stresses and it provides a useful coordinate grid for the accelerating, non-inertial frame [1]. In other words, there is a common velocity $v(t)$ of the entire ship, or, equivalently, the accelerating coordinate grid.

Fig. I.3  Rindler coordinates showing lines of constant $v(t)$ and constant $\tilde{t}$. 

161
The grid is static in the variable \( \bar{\xi} \) - the points of fixed \( \bar{x} \) maintain the same distance between them that they had at \( t = 0 \) when the entire mother ship was simultaneously at rest in the inertial frame \( S \). Note that the top of the mother ship, points of large \( \bar{x} \), accelerate less than points near its bottom, \( \bar{x} \) small. From Eq. I.1 we could define,

\[
\frac{c^2}{g(\bar{x})} = \bar{x} + \frac{c^2}{g}
\]

where \( \bar{x} > 0 \) and \( g(0) = g \). This guarantees that the Lorentz contraction of the ship is given by the global transformation law, \( \Delta x = l_0/\gamma(v) \) where \( l_0 \) is the proper length of the ship and \( v = v(t) \) is its instantaneous velocity: since the ship contracts as measured in the frame \( S \), its top must accelerate less than its bottom to accommodate a constant velocity.

As discussed in Eq. I.1, the generalizations of Eq. G.10 and G.11b to non-zero \( \bar{x} > 0 \) should be [1],

\[
ct = \left( \bar{x} + \frac{c^2}{g} \right) \sinh\left( \frac{g\bar{\xi}}{c} \right)
\]

\[
x = \left( \bar{x} + \frac{c^2}{g} \right) \cosh\left( \frac{g\bar{\xi}}{c} \right)
\]

These expressions reproduce the class of hyperbolas, \( x^2 - c^2t^2 = (\bar{x} + c^2/g)^2 \) discussed in Eq. I.1.

**The Equivalence Principle of General Relativity, Yet Again**

Let’s end this discussion with two points of interest in general relativity. First, what is the metric in the non-inertial, accelerating coordinate grid \((c\bar{t}, \bar{x})\)? This is an accelerating environment so the Equivalence Principle of general relativity states that that its local properties should be the same as those of a uniform gravitational field with a gravitational acceleration \( g = \alpha \) at \( \bar{x} = 0 \) and a gravitational potential \( V(\bar{x}) = g\bar{x} = \alpha \bar{x} \) for weak fields, \( V(\bar{x})/c^2 \ll 1 \). Second, the space, called the Rindler Wedge after its inventor, W.Rindler, has past and future horizons which make it a toy model of black holes!

To begin let’s calculate the metric in these variables [1]. This is easy. We just change variables using Eq. I.3. In the \((ct, x)\) variables we just have the Minkowski metric so,
\[ ds^2 = c^2 dt^2 - dx^2 = \sum_{\mu \nu} g_{\mu \nu} d\tilde{x}^\mu d\tilde{x}^\nu \quad (I.4) \]

To calculate \( g_{\mu \nu} \) we simply change variables following Eq. I.3 and take differentials,

\[ c dt = \sinh (g \tilde{\tau} / c) \ d\tilde{x} + \left( \tilde{x} + \frac{c^2}{g} \right) \frac{g}{c^2} \cosh (g \tilde{\tau} / c) c d\tilde{\tau} \]

\[ c dt = \sinh (g \tilde{\tau} / c) \ d\tilde{x} + \left( 1 + \frac{g \tilde{x}}{c^2} \right) \cosh (g \tilde{\tau} / c) c d\tilde{\tau} \quad (I.5) \]

And similarly,

\[ dx = \cosh (g \tilde{\tau} / c) \ d\tilde{x} + \left( \tilde{\tau} + \frac{g \tilde{x}}{c^2} \right) \sinh (g \tilde{\tau} / c) c d\tilde{\tau} \]

\[ (I.6) \]

The invariant interval is then,

\[ ds^2 = c^2 dt^2 - dx^2 = \left( 1 + \frac{g \tilde{x}}{c^2} \right)^2 c^2 d\tilde{\tau}^2 - d\tilde{x}^2 \quad (I.7) \]

after some algebra that uses the hyperbolic identities stated above. We learn from this that,

\[ g_{00} = \left( 1 + \frac{g \tilde{x}}{c^2} \right)^2 = 1 + 2 \frac{g \tilde{x}}{c^2} + \left( \frac{g \tilde{x}}{c^2} \right)^2 \quad (I.8) \]

describes the metric of the Rindler Wedge!

We saw in Chapter 11 of the textbook that in a static space time, the metric could be written in the form,

\[ ds^2 = e^{2 \Phi / c^2} c^2 dt^2 - dl \cdot dl \]

where the spatial metric is not known and could be a curved 3-manifold. For the Rindler Wedge, we have

\[ e^{2 \Phi / c^2} = \left( 1 + \frac{g \tilde{x}}{c^2} \right)^2 \]

So,

\[ \Phi(\tilde{x}) = c^2 \ln \left( 1 + \frac{g \tilde{x}}{c^2} \right) \]

And the acceleration is,
\[ g(\bar{x}) = \frac{\partial \Phi}{\partial \bar{x}} = \frac{g}{1 + g \bar{x}/c^2} \]

which agrees with our previous arguments for the proper acceleration in the Rindler Wedge. We learn that a stationary observer on the mother ship experiences a proper acceleration which falls as she ascends the ship. She weighs less the higher up the mother ship she goes.

We cannot model a world where \( g(\bar{x}) \) is strictly independent of \( \bar{x} \) as one might have ideally wished for!
Fig. I.4 Rindler space showing horizons, particles passing through horizons and accelerating grid.

Now consider the Minkowski space time in the past light cone IV of the origin \((ct, x) = (0, 0, )\), the future light cone and the Rindler Wedge region I as shown in Fig. I.4. Region II
covers the region \( x > 0 \) between the light cones \( x = \pm ct \). Suppose that a particle \( P \) meanders in the accelerating frame into region II, passing through the light cone \( x = ct \) on the way. Now the particle \( P \) has no way to communicate with an accelerating observer \( O \) on the Rindler Wedge because that would require a signal that could travel in excess of the speed of light! \( P \) itself can never pass back through the light cone and reach observer \( O \). Thus the accelerated frame, the Wedge I, has a simple future horizon and can be described as a model of a black hole. The figure also shows a past horizon where particles can enter Wedge I and be detected but cannot leave the Wedge until \( t > 0 \) and do so through the future horizon.

The hyperbolic coordinates of the Rindler Wedge provide a coordinate system for Minkowski space that is well equipped to describe accelerated reference systems. But the underlying space-time is just flat Minkowski space. The metric in Rindler coordinates reads,

\[ ds^2 = \left( 1 + \frac{g\ddot{x}}{c^2} \right)^2 c^2 d\tilde{t}^2 - d\tilde{x}^2 \]

This is just a change of coordinates from an inertial coordinate system \((ct, x)\) where

\[ ds^2 = c^2 dt^2 - dx^2 \]

These are the same space-times. Geometric quantities, such as the Ricci scalar \( R \) must be the same in either description. \( R = 0 \) is clearly the case for \( ds^2 = c^2 dt^2 - dx^2 \). How does it happen that \( R = 0 \) for the Rindler coordinate system? Let’s answer the question physically. The Rindler Wedge experiences acceleration in the \( x \) direction: Consider two balls at identical heights but having some transverse separation. They fall to the floor of the mother ship time \( \tilde{t} \) and they maintain their fixed transverse separation. In other words, there are no tidal forces in the Rindler Wedge. But vanishing tidal forces means vanishing curvature, as we learned earlier. We must consider non-uniform accelerations in the transverse directions to find curved space-times.

The reader is invited to calculate the Riemann curvature tensor to verify this explicitly.
Special Topic J. Geodesic Precession in General Relativity

We discussed Lense-Thirring precession in the presence of a rotating mass. But there is a more fundamental and usually much larger precession we should have a look at. This is geodesic precession in a static gravitational field such as the Schwarzschild metric.

Consider a small gyroscope in orbit around a spherical static mass $M$. We want to study the coordination between the gyroscope’s spin and the body’s orbital motion. Are they synchronous?

The body moves on a geodesic,
\[
\frac{du^\mu}{d\tau} + \sum_{\rho\sigma} \Gamma_{\rho\sigma}^\mu u^\rho u^\sigma = 0
\]

Where we are using Schwarzschild coordinates. We need to understand how the spin four vector $s^\sigma$ moves relative to $u^\sigma$. The inner product between the two clearly vanishes, $u \cdot s = \sum_{\sigma\rho} g_{\sigma\rho} s^\sigma u^\rho = 0$ because in the particle’s instantaneous rest frame $u^\rho$ is purely timelike and $s^\sigma$ is purely spacelike. This conservation law, $u \cdot s = 0$ must hold for all $\tau$, so $s^\sigma$ must be parallel transported along the particle’s path. So, the spin four vector must satisfy the parallel transport differential equation,
\[
\frac{ds^\mu}{d\tau} + \sum_{\rho\sigma} \Gamma_{\rho\sigma}^\mu u^\rho u^\sigma = 0
\]

Let’s suppose that the particle is executing a circular orbit. Take the Schwarzschild metric and work in the plane $\theta = \pi/2$ for convenience. We can recall the Christoffel symbols we need from our derivation of the Schwarzschild metric in lecture 12,
\[
\Gamma_{10}^0 = \frac{r_s}{2r^2} (1 - r_s/r)^{-1} \quad \Gamma_{00}^1 = \frac{r_s}{2r^2} (1 - r_s/r) \\
\Gamma_{13}^3 = -r(1 - r_s/r) \quad \Gamma_{13}^3 = \frac{1}{r}
\]

And, the equation of motion for the gyroscope’s spin becomes,
\[
\frac{ds^0}{d\tau} + \Gamma^0_{10} s^1 u^0 = 0 \quad \frac{ds^1}{d\tau} + \Gamma^1_{00} s^0 u^0 + \Gamma^1_{33} s^3 u^3 = 0 \quad \frac{ds^2}{d\tau} = 0 \quad \frac{ds^3}{d\tau} + \Gamma^3_{13} s^1 u^3 = 0
\]

Let’s look up additional results from our study of circular orbits in the Schwarzschild metric. The orbiting body’s four velocity is \( u^\sigma = u^0(1,0,0,\Omega) \) where

\[
u^0 = \frac{dt}{d\tau} = (1 - 3r_s/2r)^{-1/2} \quad \Omega = \frac{d\varphi}{dt} = \left(\frac{c^2r_s}{2r^3}\right)^{1/2}
\]

Next, the fact that \( u \cdot s = 0 \) means that \( s^0 \) and \( s^3 \) are related. Writing out the inner product,

\[
c^1(1 - r_s/r)s^0 u^0 - r^2 s^3 u^3 = 0
\]

In addition,

\[
u^0 = \frac{dt}{d\tau} = \frac{d\varphi}{d\tau} = \frac{d\varphi}{dt} = \Omega
\]

So,

\[
s^0 = \frac{\Omega r^2}{c^2(1 - r_s/r)} s^3
\]

One can check that this relation makes the geodesic equations for \( s^0 \) and \( s^3 \) identical. Now we have,

\[
\frac{ds^1}{d\tau} - \frac{r\Omega}{u^0} u^3 = 0 \quad \frac{ds^2}{d\tau} = 0 \quad \frac{ds^3}{d\tau} + \frac{\Omega}{r} s^1 u^0 = 0
\]

We can decouple the first and last equation in this set by differentiating the first one with respect to \( \tau \) and substituting in the third one. Doing this and using \( u^0 = dt/d\tau \) to go to \( t \) derivatives, we find,

\[
\frac{d^2s^1}{dt^2} + \left(\frac{\Omega}{u^0}\right)^2 s^1 = 0 \quad \frac{ds^2}{dt} = 0 \quad \frac{ds^3}{dt} + \frac{\Omega}{r} s^1 u^0 = 0
\]

We solve the first equation of the three by taking initial conditions \( s^1(0) \neq 0 \) and \( s^2(0) = s^3(0) = 0 \). Then,

\[
s^1(t) = s^1(0) \cos \Omega' t \quad s^2(t) = 0 \quad s^3(t) = -\frac{\Omega}{r\Omega'} s^1(0) \sin \Omega' t
\]
Where,

\[ \Omega' = \Omega / u^0 = \Omega (1 - 3r_s/2r)^{1/2} \]

So, the spatial part of \( s^\sigma, \bar{s} \), rotates relative to the radial direction with a coordinate speed \( \Omega' \) in the negative \( \varphi \) direction. But the orbit, the radial direction itself, rotates with coordinate angular velocity \( \Omega \) in the positive \( \varphi \) direction, as shown in Fig. J.1

Fig. J.1  The orbital position and the orientation of the spin of an orbiting gyroscope.

The difference of the two angular velocities gives the geodesic precession. In particular, in one revolution of the orbit which occurs in time \( t = \frac{2\pi}{\Omega} \), the final direction of \( \bar{s} \) is \( 2\pi + \alpha \), where,
\[
\alpha = \frac{2\pi}{\Omega} (\Omega - \Omega') = 2\pi \left(1 - \frac{\Omega'}{\Omega}\right) = 2\pi \left(1 - \frac{3r_s/2r}{1 - (3r_s/2r)^{1/2}}\right)
\]

The effect is cumulative which makes the geodesic precession effect measureable. In fact NASA’s Gravity Probe B measured it to a few percent in the same experiment where it measured the more challenging Lense-Thirring effect to about ten percent.

A few final comments about this result. In Newton’s world we would expect the gyroscope to maintain its original orientation as it orbits the earth. This means that \(\Omega' = \Omega\) and \(\alpha\) would vanish. (Since \(dt/d\tau = 1\) in Newton’s world \(\Omega' = \Omega\) trivially.) We learn from this discussion and the result \(\Omega' = \Omega/u^0\) that there is a geodesic precession effect because clocks riding with the gyroscope run at a different rate than those at rest and far away. The rates are different for two reasons. First there is the gravitational redshift. Second, the gyroscope is in a circular orbit and is not at rest in the Schwarzschild metric. Both of these effects contribute to the final result for the ratio of the angular frequencies, \(\Omega'/\Omega\), \((1 - 3r_s/2r)^{1/2}\).

**References**


