GENERAL RELATIVITY Jan 2014

Last revised on: April 9, 2024

Contents

1 Lecture 1: Introduction

1.1 Preliminary

1. GR is Einstein's theory of gravity. Well tested experimentally. We believe it is correct at least macroscopically even at the cosmological scales. It has a length scale because it has a dimensional parameter. (Dimensional analysis and units.) Newton's constant. G_N. Since GM^2/R is energy, G is $ER/M^2 \sim R/m \sim R^2$ - so L^2 - L is length. In units where $\hbar = c = 1$, this length is called Planck length l_p and is about 10^{-33}cm. $1/l_p$ defines m_p - the Planck mass scale 10^{19} Gev.

2. Gravity is very weak. People say weak compared to EM. One depends on mass, the other depends on charge, how does one compare? Take a typical "elementary particle" - electron, proton etc. Compare gravitational and el forces - gr is much weaker.

3. In units where $\hbar = c = 1$ charge is dimensionless. Whereas mass is not. So the gravitational dimensionless charge for a particle of mass M is κM, where $\kappa = \sqrt{G}$. κm_p is the gravitational charge of the proton which is much less than the electric charge: $\frac{e^2}{4\pi\hbar c} = 1/137$. So $e \approx .3$. Whereas $\kappa m = \frac{m_p}{10^{19}} \approx 10^{-19}$. This is the sense in which gravity is weak - for atomic physics, solid state physics, particle physics, nuclear physics.

4. Because it is spin2 , likes attract. And since mass is always positive, it always adds (and is attractive). Which is why macroscopically it is so important, whereas the EM force cancels out between opposite charges - becomes effectively short ranged. The weak and strong force also are unimportant because they are short ranged.

5. Before string theory, particle physicists were not much interested in gravity (being so weak) - except for a small community of supergravitists. String theory made the two unified. Also D-branes further unified things because they have both a gravity description and a Yang-Mills description. Then AdS/CFT made them very directly related and useful - because gravity techniques could be used in Yang-Mills problems. Now gravity occupies centre stage for a large section of HEP.

6. In this course we will do classical gravity - Einstein's theory. Follow Hartle's book but add more mathematical content from Wald. MTW is also a good source - especially for the geometrical pictures. I would like to spend at least some time on AdS space time because of the string theory interest and also dS for cosmology.

7. The course will end with some lectures on cosmology - as in Hartle's book.

8. Quantum Gravity is an unsolved problem. String theory is a good candidate to explain things. But a good formalism is still lacking.

9. We will have two midterms (20%+20%). One final (40%) (or final 20% and term paper 20%). And HW + quizzes (20%). Maybe also some unannounced quizzes. HW are meant to educate, not test. You will also occasionally have some tedious algebra - useful practice for research! You are allowed to discuss with anyone, but in the end you have to work it out yourself.

10. There will be one or two revision lectures on special relativity - assuming that people have seen it in earlier courses. Chapter 4 and 5 of Hartle's book.

1.2 Gravity is Geometry

1. Einstein took seriously the idea that all particles fall with the same acceleration. i.e.

$$ma = G\frac{Mm}{R^2} \implies a = G\frac{M}{R^2} \equiv g \tag{1}$$

The m's cancel. On the LHS the parameter m is the **inertial mass** - the one that occurs in "$F = ma$" -Newton's 2nd law. On the RHS it specifies the gravitational pull of the Earth using Newton's law of gravitation - **gravitational mass**. It is an observational fact that they are equal. So all particles experience the same acceleration and they follow the same trajectory. So it can be taken as a property of space time itself.

2. What property of space time are we using? The geometry. The notion of a **geodesic** - shortest curve between two points is a basic notion in geometry. So one can say all particles move along geodesics. In flat space the shortest curve is a straight line. Not in curved space.

3. Simple example is the surface of the Earth - S^2 - the surface of a sphere. The geodesics are the "great circles" - circles that have centre at the centre of the Earth. The equator is one such. All the longitudes are also great circles.

4. Some interesting properties (See Figure (1)):

   The sum of the angles of a triangle add up to more than 180.

   Ratio of circumference to radius is not 2π. It is $2\pi \frac{sin(\frac{a}{R})}{\frac{a}{R}}$

5. Geometry is specified by defining a **metric**. Gives the distance between two points. So in flat space

$$s^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 \tag{2}$$

   In curved space one defines the line element - distance between two nearby points. For eg. $(x^1 = x, x^2 = y)$

$$ds^2 = \sum_{i,j} g_{ij} dx^i dx^j = g_{11} dx^2 + (g_{12} + g_{21}) dx dy + g_{22} dy^2 \tag{3}$$

   The (symmetric) matrix of functions $g_{ij}(x^k)$ specify the geometry. $g_{ij} = g_{ji}$. Integrate ds along a geodesic to get the distance between points that are a finite distance from each other.

6. Given a surface - eg the surface of the Earth - the geodesics, and distance between points are well defined quantitites. So typically airplanes fly along geodesics to minimise fuel consumption. But the description of the geodesic depends on coordinate system used. So if you change coordinate system g_{ij} will have to change because ds^2 is a fixed geometrical or physical quantity.

   Eg:

$$ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\theta^2 \tag{4}$$

   We are in flat space here. Different coordinate systems.

   On S^2 we have seen that circumference along a latitude depends on $sin(a/R)$. So distance is (using r instead of a)

$$ds^2 = dr^2 + R^2 sin(r/R)^2 d\theta^2 \tag{5}$$

   Different geometry from (4). Note that when $r << R$ (**??**) reduces to (4).

1.3 Inertial Frames and Galilean Relativity

1. Newton's first law **defines an inertial frame**. $\frac{d\vec{v}}{dt} = 0$ in the absence of forces. Imagin a particle at rest or moving with a constant velocity. This can be used as the origin of a coordinate system. Attach three rigid sticks orthogonally to the particle so they move along with it. Make sure there is no rotation. How? Attach three gyroscopes! This system defines an "inertial frame".

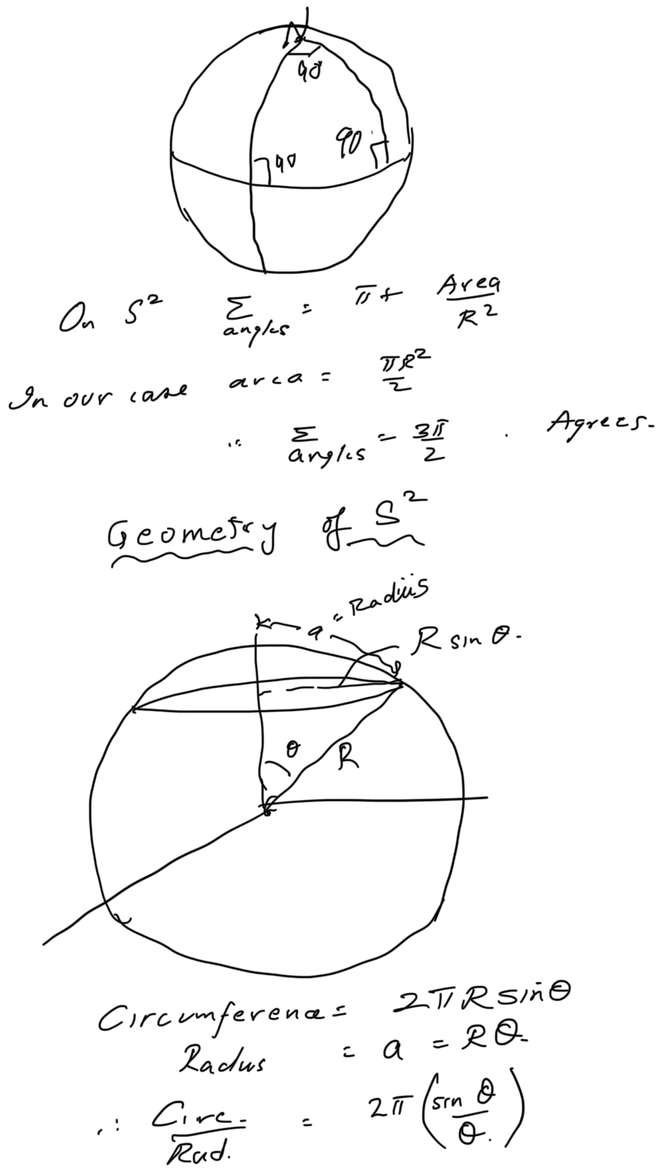


Figure 1: Geometry of S^2

2. Once the frame is fixed, we immediately realize that this is not the only inertial frame. They can be translated:

$$x' = x - a, \quad y' = y, \quad z' = z$$

rotated:

$$x' = xcos\theta + ysin\theta, \quad y' = ycos\theta - xsin\theta, \quad z' = z$$

or boosted:

$$x' = x - vt, \quad y' = y, \quad z' = z$$

Note: $t' = t$ in all these frames. Newton's first law continues to hold.

3. **Galilean Relativity:** More generally in an inertial frame, not only Newton's first law, but all laws hold - i.e. the equations look the same. That means "One cannot distinguish one frame from another by any experiment". eg moving train, space ship - how would you tell if you are moving? This is Galilean Relativity. Valid when $v << c$. Otherwise special relativity.

4. What is a non inertial frame? Accelerating w.r.t what? Newton's answer: There is an absolute space. Mach said: acceleration is w.r.t distant stars. In GR, given the mass distribution, the inertial frame at any point is determined. All masses contribute. The local inertial frame is where the usual laws are valid. This is the result of the equivalence principle.

5. **The equivalence principle:** Starts with the fact that inertial mass (m in $F = ma$) and gravitational mass (m in $F = Gm^2/r^2$) are equal always. So all objects have the same acceleration. Someone in a freely falling elevator feels no weight. It is as if there is no gravity. So you can get rid of the gravitational force by going to a particular frame. Conversely, if you feel a force, it could be due to gravity or acceleration in a gravity free region.

6. So **stationary in gravitational field means non inertial. Freely falling is gravitational field means inertial.** So Einstein said in a freely falling frame in a gravitational field all laws of physics would look as in an inertial frame - i.e usual special relativity. This allowed him to deduce how the equations have to be written down in the presence of gravity.

7. Freely falling: objects follow "geodesics" of space time. Can get Newton's first law from this - particles go in a straight line at constant speed. Principle of least proper time/distance. Works for light rays also. Related to least action principle in mechanics.

8. There are tricky issues because the statement is a local one. What about charges rotating the Earth in freely falling orbits. Do these charges radiate? Yes!

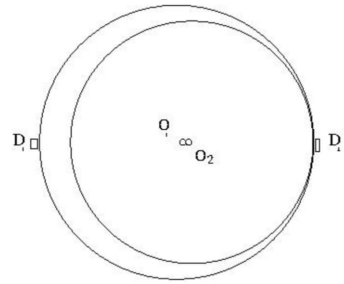


Figure 2: Relativity of Simultaneity

2 Lecture 2: Special Relativity

1. Philosophically - unified space and time into one "space-time". Draw space and time as 4-cartesian coordinates and points correspond to events.

2. No information can travel faster than light, which is finite and the basic postulate is that "the speed of light is the same in every inertial frame". That means a point source emits a circular wave front of light. Every one including those who are moving w.r.t this source sees it as a circular front.

3. One immediate consequence is that events simultaneous in one frame are not so in other frames. **"Relativity of Simultaneity".**

   First how does one synchronize clocks that are far away? Take a light source at the midway point $O_1$and send out a light front. It hits the two detectors D_1, D_2 placed at the location of the clocks at the same time. That defines simultaneous events in one frame - of O_1. Now consider a person O_2 moving towards one detector. The light front looks circular to him also. (See fig 1.) It clearly reaches the nearer detector D_1 before the other one. So the same event is not simultaneous in his frame.

4. **Time Dilation:** Consider a clock made of a light ray reflecting off two mirrors and going back and forth. Time between two ticks is the time it takes for the ray to start from one mirror and return. In the rest frame this is just $2L/c$. But the same tick as observed by a person moving will take longer - because he sees the mirrors moving at a velocity v, and the light ray traversing a distance $\frac{2\sqrt{v^2(t'/2)^2+L^2}}{c}$ where t' is the time he measures between two ticks tick. So he says

   $$t' = \frac{2\sqrt{v^2(t'/2)^2 + L^2}}{c}$$

   Solving for t' we get

   $$t' = 2L/c \frac{1}{\sqrt{1 - v^2/c^2}} \equiv 2L/c \times \gamma$$

   He thinks it takes longer since $\gamma \geq 1$. So he says the clock is running "slow". Hence the statement that "moving clocks slow down".

5. The lifetime of a muon travelling at the speed of light is much more than that of one at rest. This is experimentally verified. So its internal clock slows down. All internal clocks slow down. So we might as well call it time dilation, rather than attribute it to some mechanical explanation

6. **Lorentz Transformation:** Mathematically we want the eqn

   $$x^2 + y^2 + z^2 = c^2t^2$$

to remain the same in another frame i.e.

$$x'^2 + y'^2 + z'^2 = c^2 t'^2$$

The general solution to this are the Lorentz transformation. (Historically Lorentz had noticed that Maxwell's equation had this invariance.) The transformation is

$$\begin{aligned} x' &= \gamma(x - \beta c t) \\ t' &= \gamma(t - \beta \frac{x}{c}) \\ y' &= y \\ z' &= z \end{aligned} \tag{6}$$

where $\beta = \frac{v}{c}$ and $\gamma = \frac{1}{\sqrt{1-\beta^2}}$. When you set $c = 1$ it looks better.

$$\begin{aligned} x' &= \gamma(x - \beta t) \\ t' &= \gamma(t - \beta x) \\ y' &= y \\ z' &= z \end{aligned} \tag{7}$$

7. Let us derive the **time dilation** expression using this: We have two events separated in time by Δt and in space by Δx

$$\begin{aligned} \Delta x' &= \gamma(\Delta x - \beta c \Delta t) \\ \Delta t' &= \gamma(\Delta t - \beta \frac{\Delta x}{c}) \end{aligned} \tag{8}$$

   Say in the $x - t$ frame $\Delta x = 0$. Then $\Delta t' = \gamma \Delta t$. $\Delta x' \neq 0$ so this is the fixed frame. This clock (i.e. in the fixed frame) reads a larger time than the moving clock. So he thinks the moving clock has slowed down.

8. Yet another way of deriving this is to look at the space time diagram. See Figure (3). The light ray reaches mirror. This event is called A. A is shown in the figure. For the moving observer (with his moving clock) this event is on his time axis at some time $\Delta t'$. This is also the proper time interval OA.

   For the obsrver O at rest the event A happens after time AB, Δt. The distance OB is $\Delta x = \beta \Delta t$. So he sees a proper time for OA as $\Delta t^2 - \beta^2 \Delta t^2$. Since proper times will agree for both,

$$\Delta t'^2 = \Delta t^2 (1 - \beta^2) = \frac{\Delta t^2}{\gamma^2}$$

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

   O sees that O' 's clock is showing less ($\Delta t' < \Delta t$) than O's clock. He concludes that O' 's clock is slower than his.

   So the moving clock sees a smaller interval - moving clock is slower.

   Eg If a muon is moving along with the moving observer and the event A' is its decay, then the lab observer sees this happen at time Δt which is very large compared to $\Delta t'$. So he thinks the muon has lived longer.

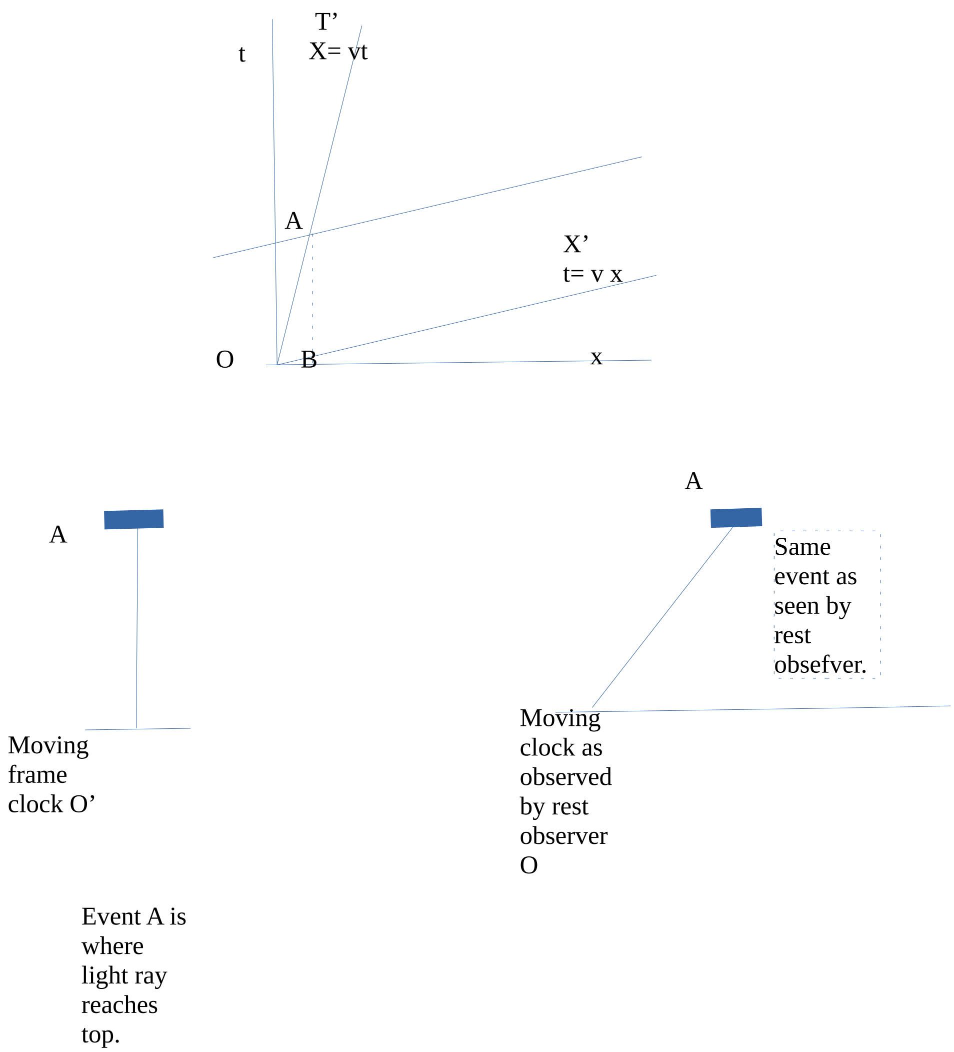


Figure 3: Time Dilation

9. Let us derive **Lorentz contraction**:

$$\begin{aligned} \Delta x' &= \gamma(\Delta x - \beta c \Delta t) \\ \Delta t' &= \gamma(\Delta t - \beta \frac{\Delta x}{c}) \end{aligned} \tag{9}$$

   Let Δx be the distance between two points of a rod. This rod is moving. What is the length observed by x' who sees this rod with x moving? He sets $\Delta t' = 0$. So $\Delta t = \beta \frac{\Delta x}{c}$. Then $\Delta x' = \gamma(\Delta x - \beta^2 c \frac{\Delta x}{c}) = \frac{\Delta x}{\gamma}$. He sees it smaller.

   Figure (3) can also be used to derive this. Do this.

10. The differential form of (6) is

$$\begin{aligned} dx' &= \gamma(dx - \beta c dt) \\ dt' &= \gamma(dt - \beta \frac{dx}{c}) \\ dy' &= dy \\ dz' &= dz \end{aligned} \tag{10}$$

   This gives

$$\frac{dx'}{dt'} = \frac{dx - \beta c dt}{dt - \beta \frac{dx}{c}} = \frac{\frac{dx}{dt} - v}{1 - \beta \frac{dx}{cdt}}$$

   Thus

$$V'_x = \frac{V_x - v}{1 - \frac{vV_x}{c^2}}$$

   Similarly

$$V'_y = \frac{V_y}{\gamma(1 - \frac{vV_x}{c^2})} \quad ; \quad V'_z = \frac{V_z}{\gamma(1 - \frac{vV_x}{c^2})}$$

   This is the relativistic law of addition of velocities.

11. Objects that transform like (x, y, z, t) are called four-vectors denoted x^μ, $\mu = 0, 1, 2, 3$ $(x^0 = t)$. The Lorentz transformation is $x'^\mu = \Lambda^\mu{}_\nu x^\nu$. It leaves invariant the dot product $x.x = x^\mu x_\mu = -(x^0)^2 + \vec{x}.\vec{x}$.

12. Another four vector is p^μ with $p^0 = E/c$. And $p^2 = m^2c^2$. So when $\vec{p} = 0$ we get $E = mc^2$. m is the "rest mass".

13. Define bases: $\hat{e}_\mu$ are orthonormal basis vectors:

$$\hat{e}_\mu . \hat{e}_\nu = \eta_{\mu\nu}$$

   - "Minkowski metric".

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\tilde{a} \equiv (a^0, a^1, a^2, a^3) = a^\mu \hat{e}_\mu = a^0 \hat{e}_0 + a^i \hat{e}_i = a^0 \hat{e}_0 + a^1 \hat{e}_1 + a^2 \hat{e}_2 + a^3 \hat{e}_3$$

   $\tilde{a}$ will denote a 4-vector whereas $\vec{a}$ will denote a 3-vector. Thus $\tilde{a}.\tilde{a} = -(a^0)^2 + \vec{a}.\vec{a}$

   Note that a^μ are only numbers - although we loosely refer to it as a 4-vector. The geometric object is $\hat{e}_\mu$ and also $\tilde{a}$.

14. $$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = -d\tau^2 \tag{11}$$

is invariant under Lorentz transformations. ds is the proper distance. When $ds^2 > 0$ the interval is "space-like". When $ds^2 = 0$ it is "light-like" or "null" and when $ds^2 < 0$ it is "time-like". We have also introduced the proper time, τ. If $ds^2 < 0$ it is better to think in terms of time intervals.

Explain **light cones**. Future light cone of an event - region that can be influenced by the event. Past light cone, region of events that can influence this event. In general one cannot cross the light cone twice - you can only go in one direction.

15. Another useful object is the **4-velocity:** If x^μ is the location of a particle in space-time, the the 4-velocity $\tilde{U}$ is defined by specifying the components:

$$U^\mu = \frac{dx^\mu}{d\tau} = (\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau})$$

where $d\tau$ is the proper time defined by (11). U^μ is a four vector because dx^μ is and $d\tau$ is a scalar - invariant.

We see immediately that

$$d\tau^2 = dt^2 - dx^2 - dy^2 - dz^2$$

$$\frac{d\tau^2}{dt^2} = 1 - v_x^2 - v_y^2 - v_z^2 = 1 - v^2 = \frac{1}{\gamma^2}$$

Here v^i are the usual 3-velocities.

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

and

$$U^i \equiv \frac{dx^i}{d\tau} = \frac{dx^i}{dt}\gamma = \gamma v^i$$

Thus $U^\mu = (\gamma, \gamma\vec{v})$. Also $U^\mu U_\mu = -1$ as an identity. In the rest frame of a particle, since by definition x^i are fixed, so $dx^i = 0$ and thus $U^i = 0$ and also $dt = d\tau$. Thus $d\tau$ **is the time as measured by a clock attached to the moving particle**. Also for this case $U^\mu = (1, 0, 0, 0)$. This is a useful fact to keep in mind. Finally

$$p^\mu = mU^\mu = \gamma m(1, \vec{v})$$

is the 4-momentum of a particle, where m is the rest mass. This gives immediately

$$p^2 = -(p^0)^2 + \vec{p}.\vec{p} = -m^2 \implies p^0 = E = \sqrt{\vec{p}.\vec{p} + m^2} \approx m + \frac{\vec{p}.\vec{p}}{2m} + ...$$

where we have expanded the expression for E in the non relativistic limit to give the usual answer. Note that the equation $E = m$ is the same as $E = mc^2$ since we are working in units where $c = 1$.

16. Dynamics: Newton's law $\vec{F} = m\vec{A}$ has an obvious generalization to

$$\tilde{f} = m\tilde{a}$$

where $\tilde{f}$ is a four vector such that $\vec{f}$ non relativistically is $\vec{F}$ and $\tilde{a} = \frac{d\tilde{U}}{d\tau}$ is the 4-acceleration. (Clearly for small velocities $\vec{a} = \vec{A} = \frac{d^2\vec{x}}{dt^2}$. We will see this more clearly below.)

$2\tilde{a}.\tilde{U} = \frac{d(U.U)}{d\tau} = 0$. Thus all four components of the 4-acc are not independent. Similarly $\tilde{f}.\tilde{U} = 0$ so all three components of the force 4 -vector are also not independent. So the number of equations of motion is still 3 - same as Newtonian mechanics.

$$f^0\gamma - \vec{f}.\vec{V}\gamma = 0 \implies f^0 = \vec{f}.\vec{V}$$

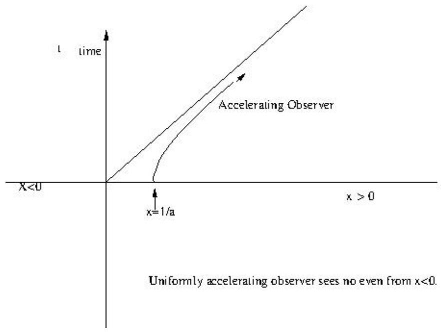


Figure 4: Uniformly accelerating observer

So

$$\tilde{f} = \gamma(\vec{F}.\vec{V}, \vec{F})$$

Here

$$\vec{f} = \frac{d\vec{p}}{d\tau} = \gamma \frac{d\vec{p}}{dt} = \gamma \vec{F}$$

$\vec{p}$ is the relativistic three -momentum.

In the rest frame of the particle $f^0 = 0$, and $dt = d\tau$, so we get

$$\vec{f} = m\vec{a} = m\frac{d^2\vec{x}}{dt^2}$$

which is just Newton's Law. The time component is

$$\frac{dE}{d\tau} = \gamma \vec{F}.\vec{V} \implies \frac{dE}{dt} = \vec{F}.\vec{V}$$

17. Motion with constant 4 acceleration:

    Choose

$$x(\tau) = a^{-1} cosh\ (a\tau), \quad t(\tau) = a^{-1} sinh\ (a\tau)$$

4-vel

$$U^1 = \frac{dx}{d\tau} = sinh\ (a\tau), \quad U^0 = \frac{dt}{d\tau} = cosh\ (a\tau)$$

4 acc

$$a^0 = \frac{dU^0}{d\tau} = asinh\ (a\tau), \quad a^1 = \frac{dU^1}{d\tau} = acosh\ (a\tau)$$

Note that $\tilde{a}.\tilde{a} = (a^1)^2 - (a^0)^2 = a^2$ is constant.

$\frac{U^1}{U^0} = v_x = tanh\ (a\tau) \to \pm 1 \quad as \quad \tau \to \pm\infty$. Approaches velocity of light far in the future and far in the past.

A reference frame using observers moving like this is very commonly used in GR and is called "Rindler Coordinates".

See figure 2: The observer never sees events from $x < 0$! The 45 degree line starting from x=0 is an "event horizon" for this observer.

18. **Photon and Massless particles:**

For a massless particle we cannot parametrize the world line using τ because $d\tau = 0$ for light like motion. So we introduce an arbitrary parameter λ and write

$$U^\mu = \frac{dx^\mu}{d\lambda}$$

Note that $\tilde{U}.\tilde{U} = 0$ A special choice of λ is made by writing

$$x^\mu = U^\mu \lambda$$

with U^μ being a constant so that $\frac{dU^\mu}{d\lambda} = 0$. It is then called an **affine** parameter.

The energy and momentum of a photon also form a four-vector although since the rest mass is zero, one cannot write $p = mU$.

Alternatively one can write $p^\mu = m\frac{dx^\mu}{d\tau}$ and take the limit $m \to 0, d\tau \to 0$ with $m\frac{dx^\mu}{d\tau} = \frac{dx^\mu}{d\lambda}$. So we can call this the four momentum $\tilde{p}$ which satisfies $\tilde{p}.\tilde{p} = 0$. Quantum mechanics tells us that $p^0 = E = \hbar\omega$, $\vec{p} = \hbar\vec{k}$ where ω is the frequency and $\vec{k}$ is the wave vector. So we have $\tilde{p} = \hbar\tilde{k}$. $\tilde{k}.\tilde{k} = 0$. So we have $E = |\vec{p}|$ and $\omega = |\vec{k}|$.

19. **Doppler Shift:**

Let ω, k_x be the frequency and wave number in the rest frame of the source. Observer is moving to the right with $\beta \geq 0$. So $x' = \gamma(x - \beta t)$, $t' = \gamma(t - \beta x)$. And the inverse is $x = \gamma(x' + \beta t')$, $t = \gamma(t' + \beta x')$. So $\omega = \gamma(\omega' + \beta k')$. Also for a photon $|k'| = |\omega'|$. So $\omega = \gamma\omega'(1+\beta)$. So

$$\omega' = \frac{\omega}{\gamma(1+\beta)} = \omega\sqrt{\frac{1-\beta}{1+\beta}} \leq \omega$$

Suppose the photon is going at an angle α' in the observer's frame. $k_{x'} = k' cos\ \alpha'$. Then

$$\omega = \gamma(\omega' + \beta k' cos\ \alpha') = \gamma\omega'(1 + \beta cos\ \alpha')$$

$$\omega' = \frac{\omega}{\gamma(1 + \beta cos\ \alpha')} = \frac{\omega\sqrt{1-\beta^2}}{1 + \beta cos\ \alpha'}$$

Another (better!) way: If $\tilde{k}$ is the 4-vector of a photon and $\tilde{U}$ is the 4-vector of an observer, then $-\tilde{U}.\tilde{k}$ is the frequency of the photon as observed by the observer: In his frame $U^\mu = (1,0,0,0)$, $k'^\mu = (\omega', \vec{k}')$. So $-\tilde{U}.\tilde{k} = \omega'$. But the dot product is an invariant and can be evaluated in any frame.

In the rest frame of the source $U^\mu = (\gamma, \gamma\beta)$ and $k^\mu = (\omega, k)$ So dot product is $-U.k = -\gamma\omega + \gamma\beta\omega = \omega\frac{1-\beta}{\sqrt{1-\beta^2}} = \omega\sqrt{\frac{1-\beta}{1+\beta}}$.

20. **Relativistic Beaming:** Source emits light uniformly in all directions. It is moving with velocity β in our frame.

The angle of the photon is $cos\ \alpha = \frac{k_x}{|k|} = \frac{k_x}{\omega}$ in our frame and $cos\ \alpha' = \frac{k'_x}{|k'|} = \frac{k'_x}{\omega'}$ in the rest frame of the source. The Lorentz transformations are

$$k_x = \gamma(k'_x + \beta\omega'), \quad \omega = \gamma(\omega' + \beta k'_x)$$

So

$$cos\ \alpha = \frac{\gamma(k'_x + \beta\omega')}{\gamma(\omega' + \beta k'_x)} = \frac{\omega'\ cos\ \alpha' + \beta\omega'}{\omega' + \beta\omega' cos\ \alpha'} = \frac{cos\ \alpha' + \beta}{1 + \beta\ cos\ \alpha'}$$

When $\beta = 1$, $cos\ \alpha = 1\ \forall\alpha'$! All the light rays appear to come in the forward direction. This is relativistic beaming - purely kinematical.

21. **Cosmic Ray Background Cutoff - GZK cutoff:** High energy protons in cosmic rays collide with low energy photons in CMB and produce pions:

$$p^+ + \gamma \to n + \pi^+, \quad or \quad p^+ + \gamma \to p^+ + \pi^0$$

In the CMB frame (where CMB is isotropic) $\omega_\gamma \sim 6 \times 10^{-4} e.v.\ (3^o K)$.

Question: What is the minimum energy the proton must have before this reaction becomes possible? Above that energy there won't be any protons - they will lose energy and come down to that value.

Work in the Centre of Mass frame: Let $\tilde{p}'$ be the 4-mom. Then $\vec{p}'_\gamma + \vec{p}'_{p^+} = 0$. Thus

$$\omega'_\gamma + \omega'_{p^+} = m_n + m_\pi, \quad \vec{p}'_\gamma + \vec{p}'_{p^+} = 0$$

This is because energy is a minimum when the particles are produced at rest. So the initial energy is just equal to the sum of the rest masses of the neutron and pion. So $(p'_\gamma + p'_{p^+})^2 = (m_n + m_\pi)^2 \Longrightarrow (p_\gamma + p_{p^+})^2 = (m_n + m_\pi)^2$. Expand LHS

$$-\omega_p^2 - \omega_\gamma^2 - 2\omega_p\omega_\gamma + (\vec{p}_p)^2 + (\vec{p}_\gamma)^2 + 2\vec{p}_p.\vec{p}_\gamma = -(m_n + m_\pi)^2$$

Use $\omega_\gamma^2 = \vec{p}_\gamma^2, \quad \omega_p^2 - \vec{p}_p^2 = m_p^2$. Get

$$-2\omega_\gamma\omega_p + 2\vec{p}_\gamma.\vec{p}_p = m_p^2 - (m_n + m_\pi)^2$$

Now, since they are relativistic $2\vec{p}_\gamma.\vec{p}_p = -2\omega_\gamma\omega_p$. Also $m_n \approx m_p$. So

$$-4\omega_\gamma\omega_p = -2m_n m_\pi + m_\pi^2$$

Use $m_n >> m_\pi$ to get

$$\omega_p = \frac{2m_n m_\pi + m_\pi^2}{4\omega_\gamma} = \frac{m_n m_\pi}{2\omega_\gamma}$$

$m_n \approx 1Gev = 10^9 ev, \quad m_\pi = 130Mev \approx 10^8 ev, \quad \omega_\gamma \approx 3^o K \approx 3 \times 10^{-4} ev.$

So $\omega_p \approx 10^{20} ev = 10^{11} Gev$. So enrgies larger than this will degrade *provided* there are collisions. Area of cross section of proton is $(10^{-15} m)^2 = 10^{-26} cm^2$. EM Scattering cross section $\sim 1/100\times$ geometric cross section $\approx 10^{-28}$. No of photons /cc ≈ 400. So mean free path $= \frac{1}{\sigma N} = 10^{25} cm$. This is 10 million light years - local group of about 30 galaxies.(eg Andromeda is a 725kpc= 2.5 million light years.) So protons coming from further away than this cannot have energies greater than 10^{11}Gev. Conversely, if you find such particles, they are from local group. Cosmic rays have been detected up to $10^{20} Gev$.

3 Lecture 3: Equivalence Principle and Curved Space Time

3.1 Equivalence Principle

1. A person in a freely falling elevator holds a cup in his hand. Because all objects accelerate with the same g he feels no weight on his hand. Thus effects of gravity can be removed by going to an accelerating (freely falling) frame. This is the starting point.

   Similarly if a spaceship, far removed from all gravitational fields, were to accelerate, people inside would feel a gravitational like force in the opposite direction. Thus gravity can be created by acceleration.

2. If acceleration due to gravity g, were uniform, this would be all there is to it. But it is not and can never be. g varies with space and time. So one modifies the statement to say that in small regions of space-time this equivalence holds. So that means at each point in space we need to go to a different frame to undo the effects of gravity.

   However does this mean that all there is to gravity can be reproduced by some complicated choice of accelerating frames?

3. **Analogy with EM:** Let us make an analogy. The equivalence principle as an analogy in EM. The effects of an em vector potential A, can be reproduced by changing the frame in the complex basis of wave functions. Thus consider (Dirac Hamiltonian)

   $$\bar{\Psi}\gamma^i\partial_i\Psi(x)$$

   In the presence of A it becomes

   $$\bar{\Psi}(x)\gamma^i(\partial_i - iA_i)\Psi(x)$$

   Suppose A is uniform. Then I can get the same effect by changing basis

   $$|x\rangle \to e^{i\theta(x)}|x\rangle \tag{12}$$

   (Remember $\langle x|\Psi\rangle \equiv \Psi(x)$.) In the new basis, $\Psi(x) = e^{-i\theta(x)}\Psi'(x)$.

   $$\bar{\Psi}(x)\partial_i e^{-i\theta(x)}\Psi'(x) = \bar{\Psi}'(x)\gamma^i(\partial_i - i\partial_i\theta(x))\Psi'(x)$$

   Choose $\theta(x) = A_i x^i$.

   So we have reproduced the effect of a uniform A by a change of basis! This is the "equivalence principle" of EM: Uniform $A \leftrightarrow$ uniform g.

   In EM it is called "**gauge principle**".

   Now ask the same question: Can we reproduce the effect of non uniform A by making $\theta(x)$ more complicated?

   The answer is No! $\theta(x)$ is one function. $A_i(x)$ is a set of *three* functions. So not possible.

   Note: It would be possible in one space dimension!

   In standard language $\theta(x)$ is only the pure gauge part! It mimics in *form* the em fld but not in *substance*.

   But we can use this to figure out what the equations should look like : Replace $\partial_i\theta$ by A_i. More precisely, choose the eqn such that when A is changed to $A_i - \partial_i\theta$ it cancels the effect of the basis change (12). This procedure is called "making the action *gauge invariant*".

4. Equivalence principle does not say that all effects of gravity can be reproduced by coordinate changes. But it tells you that you can do it at one point. And it tells you what the gravitational field should "look" like in an equation. It gives you information about the *form* of the equation. Einstein showed that the equations can be written in a "*coordinate invariant*" way by introducing a gravitational field appropriately.

5. Let us consider a very simple example of this. Consider the equation

$$F = m\frac{d^2z}{dt^2} \tag{13}$$

This is invariant under

$$z = z' - vt, \quad F' = F$$

z' is the coordinates used by an observer moving downwards with a velocity v. But it is not invariant under

$$z = z' - \frac{1}{2}at^2, \quad F = F' \tag{14}$$

The force is a physical thing and we would like to get the same answer. Here z' is the coordinate of an observer accelerating downward with an acceleration a.

Can we introduce a new field - a gravitational field g into the equation to make it invariant?

Modify the first equation (13)

$$F = m\frac{d^2z}{dt^2} + mg \tag{15}$$

Make the transformation

$$z = z' - \frac{1}{2}at^2, \quad g = g' + a, \quad F = F' \tag{16}$$

Now the equation becomes

$$F' = m\frac{d^2z'}{dt^2} + mg' \tag{17}$$

Clearly (15) and (17) look the same.

To see what is happening consider a special case.

Let O be at rest. O sees a gravitational field g and a weight mg.

Now, for instance let $a = g$. Then O' is a freely falling observer. So he sees an object accelerating upwards with g. But he sees no gravitational field ($g' = 0$). He ascribes the force to acceleration. Both are correct. By introducing a gravitational field we have managed to make the equations invariant - or covariant!

So a *constant uniform* gravitational field is mimicked by going to a uniformly accelerating frame. This is the equivalence principle.

We will see later that the number of independent functions available from coord transf is less than the number required to describe grav fld. So you can't explain all gravitational phenomena by saying it is just a coordinate change.

But again in one dimension, that is all there is to it - and this is almost true in 2 dimension.

3.2 Example of Equivalence Principle in Newtonian Mechanics

1. Let us apply the eq principle to understand the effect of a gravitational field on clocks. So consider a rocket in outer space where there is no grav field. See figures 1 and 2.

   A light ray enters through one window and exits through the other. The person on the ground sees the light ray going in a straight line and an accelerating rocket. The person in the rocket thinks he is in a gravitational field and that light is bending due to gravity.

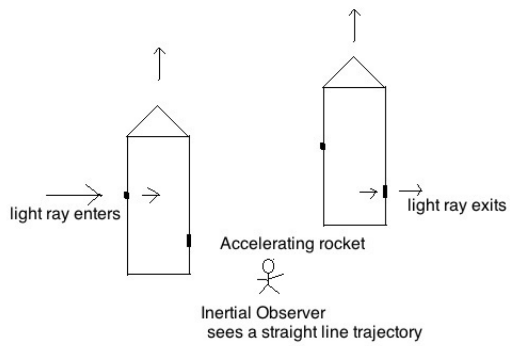


Figure 5: Equivalence Principle: Inertial Observer's Viewpoint

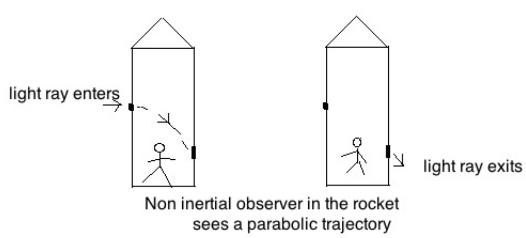


Figure 6: Equivalence Principle: Non Inertial Observer's Viewpoint

2. We will use the equivalence principle to study the effect of uniform gravity on clocks. Consider a rocket accelerating upward with an acc of g. By the eq pr. this simulates a gravitational fld of g. We are going to use Newtonian mechanics with absolute time.

   Let the distance between the observers A and B be h. Signal is emitted at time 0. It is received at time t. Distance travelled by light from A is ct. Distance travelled by rocket (and thus distance travelled by B) is $\frac{1}{2}gt^2$. So

$$ct + \frac{1}{2}gt^2 = h \tag{18}$$

   Now consider the next signal. It is sent at time T and received at time $t + T'$. Distance travelled by ray is $c(t + T' - T)$. Distance travelled by rocket (and thus B) is $(gT)(t + T' - T) + \frac{1}{2}g(t + T' - T)^2$. Thus

$$c(t + T' - T) + (gT)(t + T' - T) + \frac{1}{2}g(t + T' - T)^2 = h$$

   Let us drop terms of order T^2, T'^2.

$$c(t + T' - T) + gTt + \frac{1}{2}gt^2 + gt(T' - T) = c(t + T' - T) + \frac{1}{2}gt^2 + gtT' = h \tag{19}$$

   Combining (18) and (19) we get

$$c(T' - T) + gtT' = 0$$

$$1 - \frac{T}{T'} = -g\frac{t}{c} = -\frac{gh}{c^2}$$

   We have assumed that $ct >> \frac{1}{2}gt^2$. Thus

$$\frac{T}{T'} = 1 + \frac{gh}{c^2}, \qquad \frac{T'}{T} = 1 - \frac{gh}{c^2} \tag{20}$$

   Equivalently, frequency being inversely prop to time intervals,

$$\frac{\omega'}{\omega} = 1 + \frac{gh}{c^2} \tag{21}$$

   Thus the frequency seen by B is higher than that seen by A. This is like saying the photon acquires an energy $mgh = \frac{E}{c^2}gh$ as it comes down in a gravitational field.

   B's ticks are spaced closer than A's. So he will conclude that photon has higher frequency. On the other hand if he is using it to define his time, we can say that his proper time interval between ticks is less than for A. So A will say that B's clock has slowed down. In the same sense that we say a moving clock slows down.

3.3 Curved Space Time in Newtonian Mechanics

1. We said that uniform gravity can be simulated by a constant acceleration but this cannot be true in general. There is more to it than a coordinate change. What is more? Coordinates are drawn on a space time manifold. Changing coordinates changes the labeling of points but does not change the **geometry**. So what is more is the geometry.

   When clocks are affected by gravitational field, instead of ascribing this to the mechanical properties of the clock (which would be very complicated because all clocks regardless of mechanism have to slow down) we might as well assign it to time itself changing. More precisely the metric of space-time.

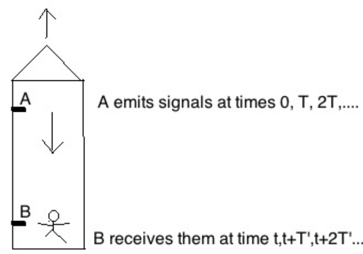


Figure 7: Effect of gravitational field on clocks using the equivalence principle

2. Let us illustrate this in Newtonian mechanics. Geometry is specified by giving a metric. Consider

$$ds^2 = -(1 + \frac{2\phi(x)}{c^2})(cdt)^2 + (1 - \frac{2\phi(x)}{c^2})(dx^2 + dy^2 + dz^2) \tag{22}$$

t is coordinate time. Proper time is $-ds^2$. The trajectory of light rays obey $ds^2 = 0$. This gives

$$\frac{dx}{dt} = c\frac{(1 + \frac{2\phi(x)}{c^2})}{(1 - \frac{2\phi(x)}{c^2})}$$

The shape of the world line is clearly independent of t. They follow parallel curves. (see figure (20)).

So the photon that leaves x at $t = 0$ reaches x' at t and the photon that leaves x at $t = \delta t$ reaches x' at $t + \delta t$. So the coordinate time separation δt is the same at x and x'. Thus if proper time intervals are T, T' at two different points x, x', their ratio is clearly,

$$\frac{T}{T'} = \sqrt{\frac{(1 + \frac{2\phi(x)}{c^2})}{(1 + \frac{2\phi(x')}{c^2})}} \approx 1 + \frac{\phi(x) - \phi(x')}{c^2} \approx 1 + \frac{gh}{c^2}$$

which reproduces the time dilation found above.

3. We have thus illustrated that a modified geometry produces the required effect. This metric is taken as the weak field limit of a metric in GR. Clearly the coefficient of dx^2 doesn't play a big role and to this extent many metrics can produce the required effect.

4. In addition to trajectories of light rays that satisfies $ds^2 = 0$, this metric singles out some world lines. These are **geodesics**. Geodesics are curves that connect two points and give the shortest (proper) distance between them. In flat Minkowski space these are straight lines. Thus Newton's first law is essentially saying that particles follow geodesics when no forces act on them.

   When gravitational forces are present they do not go along straight lines. However we expect them to undergo "free fall". These trajectories are in fact geodesics of the new geometry. This will be clear later. Let us test this with our metric.

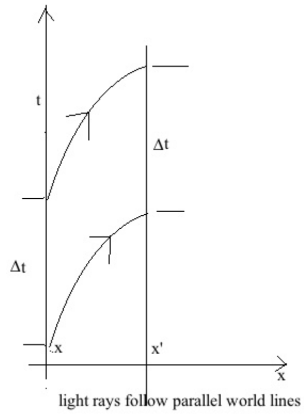


Figure 8: Light rays follow identical trajectories shifted vertically by coordinate time interval Δt

Proper time along a curve can be written as

$$\tau_{AB} = \int_A^B d\tau = \int \frac{d\tau}{dt} dt = \int \sqrt{(1+\frac{2\phi(x)}{c^2}) - \frac{1}{c^2}(1-\frac{2\phi(x)}{c^2})[(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2 + (\frac{dz}{dt})^2]}dt$$

If we neglect higher orders in $\frac{1}{c^2}$ we get

$$\tau_{AB} = \int_A^B d\tau = \int \frac{d\tau}{dt} dt = \int \sqrt{(1+\frac{2\phi(x)}{c^2}) - \frac{1}{c^2}[(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2 + (\frac{dz}{dt})^2]}dt$$

Extremizing this is equivalent to extremizing:

$$\int dt[-\phi(x) + \frac{1}{2}[(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2 + (\frac{dz}{dt})^2]]$$

which is the non relativistic action for a particle in a gravitational field.

5. What we need to do is to make all this consistent with special relativity.

4 Lecture 4: Basic Introduction to Tensors and Differential Geometry

4.1 Vectors, Covectors, Tensors

1. Tangent Vectors:

   In flat space we introduced 4-vectors and expressed them in terms of basis vectors. We can do the same thing in curved space. However things have to be done locally. Thus on S^2 a finite length vector from equator to pole is not a vector. A small -infinitesimal - vector connecting two close by points is a vector. It is tangent to the surface - so it is a <u>tangent</u> vector.

2. At each point there is a vector space. We can introduce a basis - $\vec{e}_\mu$. Any vector $\vec{A}$ at that point can be writes as $\vec{A} = A^\mu \vec{e}_\mu$.

3. We can also have as usual the dual vector space *at each point* - the space of maps : that map a vector to a real number. This is spanned by a dual basis of co-vectors $\underline{\omega}^\nu$ that obey

$$\langle \underline{\omega}^\nu, \vec{e}_\mu \rangle = \delta^\nu_\mu \tag{23}$$

   So a general covector is $\underline{W} = W_\nu \underline{\omega}^\nu$

4. **Tensors**: One can introduce product spaces spanned by $\vec{e}_\mu \otimes \vec{e}_\nu ...$ which give us "contravariant" tensors $\mathbf{A} = A^{\mu\nu ..} \vec{e}_\mu \otimes \vec{e}_\nu ....$ Similarly $\underline{\omega}^\mu \otimes \omega^\nu ....$ span the dual space and we have "covariant" tensors $\mathbf{W} = W_{\mu\nu ..} \underline{\omega}^\mu \otimes \omega^\nu .....$ Finally one can have mixed tensors $\mathbf{A} = A^\mu_\nu \vec{e}_\mu \otimes \omega^\nu$ etc.

   Note: All this is done *locally* - at each point.

5. These tensors are geometric objects and have a meaning independent of the coordinate system. Using the inner product (23) , one gets a number for a tensor and a set of vectors and covectors. This inner product is independent of coordinate systems. Thus if, for instance, $\mathbf{A} = A^\mu_\nu \vec{e}_\mu \otimes \omega^\nu$ and $\vec{b} = b^\mu \vec{e}_\mu$ and $\underline{w} = w_\nu \underline{\omega}^\nu$, then

$$\mathbf{A}(\vec{b}, \underline{w}) = A^\mu_\nu b^\nu w_\mu$$

6. We will see more formally later that one can associate a derivative operator with a vector. In flat space take $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$. Consider $\frac{\partial}{\partial x}\vec{r} = \hat{i}$. Similarly $\frac{\partial}{\partial y}\vec{r} = \hat{j}$ and $\frac{\partial}{\partial z}\vec{r} = \hat{k}$. Thus

$$\frac{\partial}{\partial x} \leftrightarrow \hat{i}, \quad \frac{\partial}{\partial y} \leftrightarrow \hat{j}, \quad \frac{\partial}{\partial z} \leftrightarrow \hat{k}$$

   More generally one could have $\vec{e}_x = v^x(x,y,z)\frac{\partial}{\partial x}, \quad \vec{e}_y = v^y(x,y,z)\frac{\partial}{\partial y}, \quad \vec{e}_x = v^z(x,y,z)\frac{\partial}{\partial z}$.

   Even more generally a vector is just some differential operator.

7. In the special case where we have a coordinate system such that the basis vectors are $\frac{\partial}{\partial x^\mu}$, these are called a coordinate basis.

8. The same can be said for the dual vector space of co vectors. A coordinate basis would be $\underline{dx}^i = (\underline{dx}, \underline{dy}, \underline{dz})$. But in general $\underline{\omega}^i$ need not be d of anything.

9. A coordinate basis for vectors and co-vectors is:

$$\langle dx^i, \frac{\partial}{\partial x^j} \rangle = \delta^j_i$$

10. An example of such a geometric tensor is the metric tensor:

$$\mathbf{g} = g_{\mu\nu}\underline{\omega}^\mu \otimes \underline{\omega}^\nu$$

The coefficients $g_{\mu\nu}$ can be evaluated as

$$\mathbf{g}(\vec{e}_\mu, \vec{e}_\nu) = g_{\mu\nu}$$

Informally one writes:

$$\vec{e}_\mu . \vec{e}_\nu = g_{\mu\nu}$$

The same tensor can be written in a coordinate basis. The coefficients are different:

$$\mathbf{g} = g'_{\mu\nu}\underline{dx}^\mu \otimes \underline{dx}^\nu$$

Thus

$$\mathbf{g}(\frac{\vec{\partial}}{\partial x^\mu}, \frac{\vec{\partial}}{\partial x^\nu}) = g'_{\mu\nu}$$

11. One can evaluate

$$\mathbf{g}(\vec{A}, \vec{B}) = g_{\mu\nu}A^\mu B^\nu$$

One can also define the covariant

$$\mathbf{g^{-1}} = g^{\mu\nu}\vec{e}_\mu \otimes \vec{e}_\nu$$

The fact that $g^{\mu\nu}$ and $g_{\mu\nu}$ are inverses as matrices is evident.

12. **Lowering and raising indices**: The metric can be used to raise and lower indices.

$$A^\mu g_{\mu\nu} \equiv A_\nu$$

$\underline{A} = A_\mu \underline{\omega}^\mu$ can be thought of as a covector formed from the vector A^μ. It is defined so that the dot product $\vec{A}.\vec{B} = \langle \underline{A}, \vec{B} \rangle$

13. The metric gives the distance between two points. Note that metric structure is over and above the manifold and coordinate system and tangent vector space etc. It is an extra attribute of the manifold - if it exists.

14. A flat Euclidean manifold is where $g_{\mu\nu} = \delta_{\mu\nu}$. For Minkowski, use $\eta_{\mu\nu}$.

15. Vierbein: We have seen that $\vec{e}_\mu . \vec{e}_\nu = g_{\mu\nu}$. On the other hand at every point, one can rescale and choose a basis such that $\hat{e}_a . \hat{e}_b = \eta_{ab}$. This is called an orthonormal basis. The two basis are then related by

$$\vec{e}_\mu = e^a_\mu \hat{e}_a \tag{24}$$

e^a_μ are called **vierbeins** and it is easy to show that:

$$e^a_\mu e^b_\nu g^{\mu\nu} = \eta^{ab}, \quad or \quad e_{a\mu} e_{b\nu} g^{\mu\nu} = \eta_{ab} \tag{25}$$

Index a on e^a_μ are lowered using η_{ab}.

$$e^a_\mu e^b_\nu \eta_{ab} = g_{\mu\nu}$$

is also easy to show.

4.2 Examples: Polar and Cartesian coordinates

(a) Cartesian Coordinates: $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$. So

$$\frac{\vec{\partial}}{\partial x} \leftrightarrow \hat{i}, \quad \frac{\vec{\partial}}{\partial y} \leftrightarrow \hat{j}, \quad \frac{\vec{\partial}}{\partial z} \leftrightarrow \hat{k}$$

Dual space basis: dx, dy, dz. Metric for flat space is $\mathbf{g} = \underline{dx} \otimes \underline{dx} + \underline{dy} \otimes \underline{dy} + \underline{dz} \otimes \underline{dz}$. More commonly $ds^2 = dx^2 + dy^2 + dz^2$. Thus $\frac{\vec{\partial}}{\partial x} \cdot \frac{\vec{\partial}}{\partial x} = \mathbf{g}(\frac{\vec{\partial}}{\partial x}, \frac{\vec{\partial}}{\partial x}) = 1$. Similarly for the others.

(b) One can also write $\vec{r} = r\ cos\ \theta\hat{i} + r\ sin\ \theta\hat{j}$. Then

$$\frac{\vec{\partial}}{\partial \theta} \approx \frac{\partial \vec{r}}{\partial \theta} = \vec{e}_\theta = -r\ sin\ \theta\hat{i} + r\ cos\ \theta\hat{j} \implies \hat{e}_\theta = -\ sin\ \theta\hat{i} +\ cos\ \theta\hat{j}$$

$$\frac{\vec{\partial}}{\partial r} \approx \frac{\partial \vec{r}}{\partial r} = \vec{e}_r =\ cos\ \theta\hat{i} +\ sin\ \theta\hat{j} = \hat{e}_r$$

This tells us how to change basis from Cartesian to Polar. Thus

$$\vec{e}_\theta . \vec{e}_\theta = r^2, \quad \vec{e}_r . \vec{e}_r = 1$$

This gives the metric in polar coordinates.

(c) Polar coordinates metric: $ds^2 = dr^2 + r^2 d\theta^2$ or $\mathbf{g} = dr \otimes dr + r^2 d\theta \otimes d\theta$. Thus we can use $\underline{\omega}^r = \underline{dr}$, $\underline{\omega}^\theta = r\underline{d\theta}$ as our basis co vectors and $\mathbf{g} = \underline{\omega}^r \otimes \underline{\omega}^r + \underline{\omega}^\theta \otimes \underline{\omega}^\theta$. Dual basis is $\hat{e}_\theta = \frac{1}{r}\frac{\vec{\partial}}{\partial \theta}$, $\hat{e}_r = \frac{\vec{\partial}}{\partial r}$. Thus

$$\mathbf{g}(\hat{e}_\theta, \hat{e}_\theta) = 1, \quad \mathbf{g}(\hat{e}_r, \hat{e}_r) = 1$$

Clearly this is an orthonormal basis. whereas

$$\mathbf{g}(\frac{\vec{\partial}}{\partial \theta}, \frac{\vec{\partial}}{\partial \theta}) = g_{\theta\theta} = r^2, \quad \mathbf{g}(\frac{\vec{\partial}}{\partial r}, \frac{\vec{\partial}}{\partial r}) = g_{rr} = 1$$

(d) Similarly

$$\mathbf{g^{-1}}(\underline{dr}, \underline{dr}) = g^{rr} = 1, \quad \mathbf{g^{-1}}(\underline{d\theta}, \underline{d\theta}) = g^{\theta\theta} = \frac{1}{r^2}$$

(e) Let us define vierbeins to go from Polar coordinates to Cartesian $\frac{\vec{\partial}}{\partial \theta} = e^x_\theta \hat{i} + e^y_\theta \hat{j}$. Thus

$$e^x_\theta = -r\ sin\ \theta, \quad e^y_\theta = r\ cos\ \theta$$

$\frac{\vec{\partial}}{\partial r} = cos\ \theta\hat{i} + sin\ \theta\hat{j}$. So

$$e^x_r = cos\ \theta, \quad e^y_r = sin\ \theta$$

One can verify that $e^x_\theta e^x_\theta + e^y_\theta e^y_\theta = g_{\theta\theta} = r^2$ and so on.

4.3 Volume Element, Coordinate Transformations

16. Volume element: Consider a metric $ds^2 = g_{xx}dx^2 + g_{yy}dy^2$ in two dimensions. Then the area element is clearly

$$dA = \sqrt{g_{xx}}dx\ \sqrt{g_{yy}}dy = \sqrt{det\ g_{\mu\nu}}\ dxdy$$

e.g. In polar coordinates

$$\sqrt{det\ g_{\mu\nu}}\ dxdy \sim r\ drd\theta$$

This can be generalized to any dimension: $dV = \sqrt{g}\ d^n x$, where $g = det\ g_{\mu\nu}$. For Minkowski metric we write $dV = \sqrt{-g}\ d^n x$

17. Change of coordinates: Let $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$ be a metric in the coordinate system x^μ. Suppose we change coordinates to $x^{'\mu}(x)$. Proper distances are not supposed to change. So

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = g'_{\mu\nu}dx^{'\mu}dx^{'\nu}$$

$$= g'_{\mu\nu}\frac{\partial x^{'\mu}}{\partial x^\alpha}\frac{\partial x^{'\nu}}{\partial x^\beta}dx^\alpha dx^\beta$$

So

$$g_{\alpha\beta} = g'_{\mu\nu}\frac{\partial x^{'\mu}}{\partial x^\alpha}\frac{\partial x^{'\nu}}{\partial x^\beta}$$

Similar transformation rules apply for any rank tensors. For eg we write

$$\vec{A} \equiv A^\mu \frac{\vec{\partial}}{\partial x^\mu} = A^\mu \frac{\partial x^{'\alpha}}{\partial x^\mu}\frac{\vec{\partial}}{\partial x^{'\alpha}} \equiv A^{'\alpha}\frac{\vec{\partial}}{\partial x^{'\alpha}}$$

thus $A^{'\mu} = \frac{\partial x^{'\mu}}{\partial x^\alpha}A^\alpha$.

In many books these transformation properties define tensors.

18. Can any metric be made into $\eta_{\mu\nu}$ by a sufficiently complicated coord transf? A general coord transf is of the form $x^{'\alpha}(x^\mu)$ which can be Taylor expanded about any point (say 0):

$$x^{'\alpha}(x) = x^{'\alpha}(0) + \underbrace{\frac{\partial x^{'\alpha}}{\partial x^\mu}|_0}_{4\times 4=16} x^\mu + \frac{1}{2!}\underbrace{\frac{\partial^2 x^{'\alpha}}{\partial x^\mu \partial x^\nu}|_0}_{4\times\frac{4.5}{1.2}=40} x^\mu x^\nu + \frac{1}{3!}\underbrace{\frac{\partial^3 x^{'\alpha}}{\partial x^\mu \partial x^\nu \partial x^\sigma}|_0}_{4\times\frac{4.5.6}{1.2.3}=80} x^\mu x^\nu x^\sigma + ...$$

The metric itself can be expanded as

$$g_{\mu\nu}(x) = g_{\mu\nu}(0) + \underbrace{\frac{\partial g_{\mu\nu}}{\partial x^\rho}|_0}_{10\times 4=40} x^\rho + \underbrace{\frac{\partial^2 g_{\mu\nu}}{\partial x^\rho \partial x^\sigma}|_0}_{10\times\frac{4.5}{1.2}=100} x^\rho x^\sigma + ... + ...$$

The sixteen first derivatives can be used to make $g_{\mu\nu}(0) = \eta_{\mu\nu}$. (A symmetric matrix can be diagonalized by an orthogonal transformation - 6 parameters, then the eigen values can be made into 1 by four rescalings - 4 parameters.) The remaining 6 are Lorentz transformations that leave $\eta_{\mu\nu}$ invariant.

The 40 second derivatives can be used to make the 40 first derivatives of $g_{\mu\nu}$ at 0, to zero.

80 of the 100 second derivatives of $g_{\mu\nu}$ can be set to zero using the 80 third derivatives of x. **The remaining 20 cannot be set to zero and are physical. They constitute the Riemann Curvature tensor.** One can do Lorentz transformations (6 parameters) on them. So of these 20 only 14 are independent.

This is why coordinate transformations cannot mimic gravity in general.

Thus: Locally (i.e. at any one point) on a manifold one can choose coordinates such that, not only $g_{\mu\nu} = \eta_{\mu\nu}$ but also

$$\frac{\partial g_{\mu\nu}}{\partial x^\rho}|_{x=x_0} = 0$$

So it looks like flat space in a small neighbourhood. This is geometry. But it is also the equivalence principle - in a local inertial frame the laws of flat space hold. That is why gravity is described geometrically.

19. A possible source of confusion: Vierbeins can be used to change basis so that the metric is flat. Thus compare $\eta_{\alpha\beta} = g'_{\mu\nu}\frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial x'^{\nu}}{\partial x^{\beta}}$ with $e_{a\mu}e_{b\nu}g^{\mu\nu} = \eta_{ab}$. The former cannot in general be done (everywhere- it can be done at one point) whereas the latter can. What is the difference? The former is a coordinate change and is defined by *four* functions $X'^{\mu}(X^{\alpha})$. Cannot transform a *ten* component object. The latter is defined by *sixteen* functions $e_{a\mu}(X)$, which is more than enough to do the job. The extra six are local - (local, because these are functions of X)- Lorentz transformations, that do not affect η_{ab}.

20. Simple example of Curved Spaces: Spheres.

    Consider a two-sphere S^2 sitting in $\mathbb{R}^3$. Metric in $\mathbb{R}^3$ is

    $$ds^2 = dx^2 + dy^2 + dz^2 \tag{26}$$

    and the equation for the sphere is $x^2+y^2+z^2 = R^2$. We can parametrize the sphere as follows: Let ρ, ϕ be polar coordinates for the xy-plane. Then $\rho^2 = x^2 + y^2 = R^2 - z^2$. Also $\rho\ cos\ \phi = x, \quad \rho\ sin\ \phi = y$. We know that

    $$dx^2 + dy^2 = d\rho^2 + \rho^2 d\phi^2, \quad ds^2 = d\rho^2 + \rho^2 d\phi^2 + dz^2 \tag{27}$$

    Also $2\rho d\rho = -2zdz$. Thus

    $$dz^2 = \frac{\rho^2}{z^2}d\rho^2 = \frac{\rho^2}{R^2-\rho^2}d\rho^2$$

    Combining this with (27) we get

    $$ds^2 = \frac{d\rho^2}{1-\frac{\rho^2}{R^2}} + \rho^2 d\phi^2 \tag{28}$$

    Notice that near $\rho = 0$(North pole of the sphere) it is just a plane. Near $\rho = R$ the coordinate system is singular and the metric coefficient diverges. But there is nothing unusual happening there geometrically. This is a **coordinate singularity**. A coordinate singularity also happens at the horizon of a black hole in some coordinate systems. Also at $\rho = 0$ the coordinate system breaks down because the value of ϕ is not determined.

    We cannot parametrize the whole sphere with one coordinate patch. We need at least two. All coordinate patches become singular somewhere on the sphere.

    Another common example is the spherical coordinate system using θ, ϕ - longitudes and latitudes. The metric is

    $$ds^2 = R^2(d\theta^2 + sin^2\ \theta d\phi^2) \tag{29}$$

    This also breaks down at $\theta = 0, 180$ - the poles - because the value of ϕ is not defined.

21. Hypersurfaces: The above was an example of a hypersurface. Take a three dimensional space (which happened to be flat R^3 in the above example). Define a two dimensional surface in that three dimensional space by specifying a relation between x, y, z. We can do the same thing in 3+1 dimensional space-time. Thus an equation of the form

    $$t = f(x, y, z)$$

    divides space time into three dimensional hypersurfaces labelled by the time coordinate. Eg in 1+1 dimension one can write $t = x$. Defines a line in space time. $t = x + 1$ defines another line. The set of all lines of the form $t = x + a$ can cover the whole 1+1 dimenional space time. Another eg is $t = 0$. This is just the x-axis. Similarly $t = 1$. Thus the set of lines $t = a$ cover the space time. Each line is a hypersurface.

    In the first example, the line was a trajectory of a light ray. Each line is a "null" line. In the second example it was spacelike. So each line can be thought of as space.

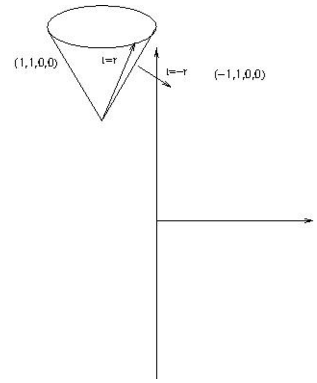


Figure 9: An example of a null hypersurface is a lightcone.

More generally we can write a hypersurface as

$$f(x,y,z,t) = 0$$

We can ask whether it is spacelike, null etc. Let $\vec{t}_1, \vec{t}_2, \vec{t}_3$ be tangent vectors along the hypersurface. Let $\vec{n}$ be a normal defined by $\vec{n}.\vec{t}_i = 0, \quad \forall i$. If

$$\vec{n}.\vec{n} < 0, \quad \vec{t}_i.\vec{t}_i > 0$$

then surface is space-like.

If $\vec{n}.\vec{n} > 0$ then the hyper surface is time like. Then one of the tangents is time like.

If $\vec{n}.\vec{n} = 0$ then the hyper surface is null. But $\vec{t}$ is defined by $\vec{n}.\vec{t} = 0$. This means that $\vec{n}$ lies on the hypersurface! The normal to a null hypersurface lies on the hypersurface.

Example 1: Null hypersurface.

Light cone in flat space time. The light cone is defined by $t = r$. Consider $l = (1,1,0,0)$ where we choose (t, r, θ, ϕ) coordinates. $l.l = 0$ is a null tangent. $(0,0,1,0)$ and $(0,0,0,1)$ are the other two space-like tangents. Note that $(1,-1,0,0)$ is also null. It is not orthogonal to l. So it cannot be considered to be a normal. But it is the 4th direction.

Example 2: Spacelike hypersurface.

A surface given by:

$$r^2 - t^2 = -a^2$$

$r = a\ sinh\ \chi,\ t = a\ cosh\ \chi$ is a parametrization of this surface (in addition to θ, ϕ).

$$ds^2 = a^2[d\chi^2 + sinh^2\chi(d\theta^2 + sin^2\theta d\phi^2)]$$

This is a spatial slice in "Anti deSitter space". The time part is $-a^2 cosh^2\chi d\tau^2$.

The tangent vector $\vec{t} = (\frac{\partial t}{\partial \chi}, \frac{\partial r}{\partial \chi}, 0, 0) = (a\ sinh\ \chi, a\ cosh\ \chi, 0, 0)$ satisfies $\vec{t}.\vec{t} = a^2 > 0$. Also the vector $\vec{n} = (cosh\ \chi, sinh\ \chi)$ is clearly orthogonal to $\vec{t}$ and also to e_ϕ, e_θ and so is a normal. It satisfies $\vec{n}.\vec{n} = -1 < 0$ and so we say this hypersuface is space-like.

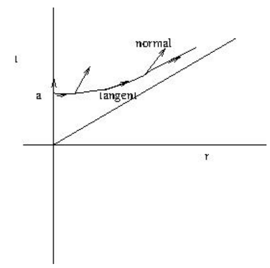


Figure 10: An example of a spacelike hypersurface.

A very important example of a coordinate transformation is the Penrose diagram. The Penrose diagram for flat space is given here.

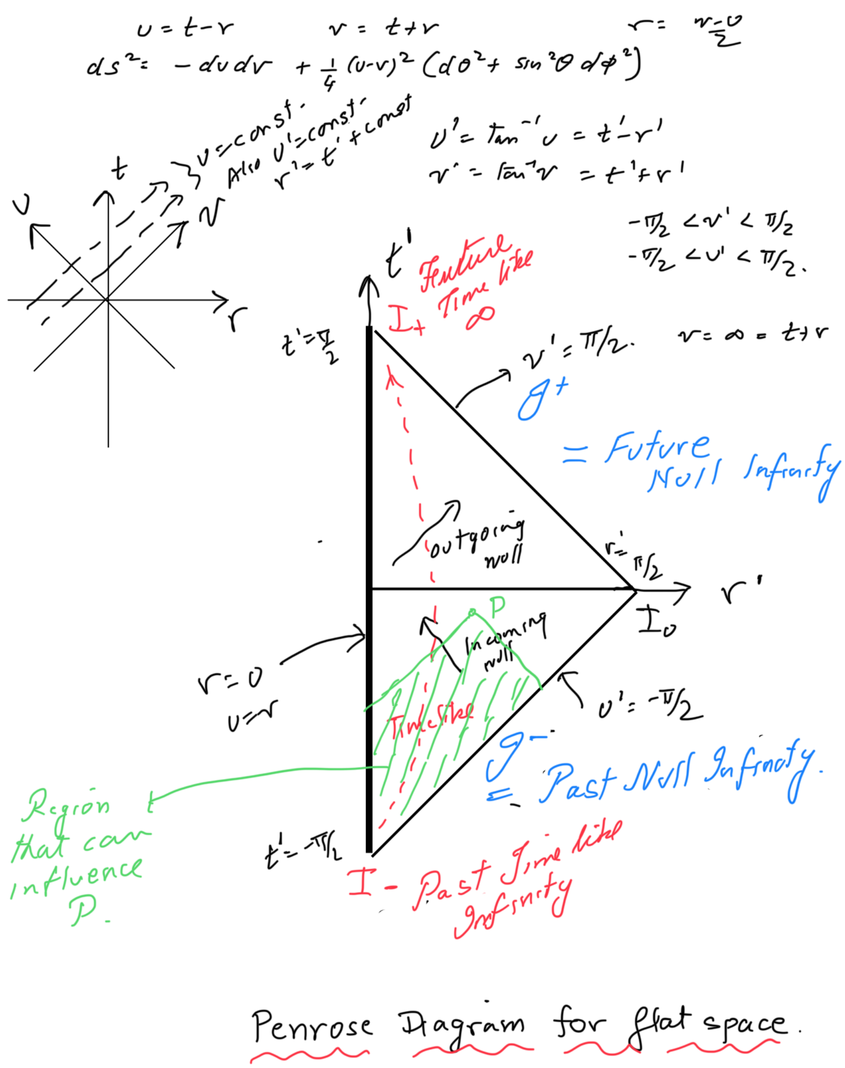


Figure 11: Penrose diagram for flat space

5 Lecture 5: Geodesics

5.1 Geodesic Equation

1. Newton's First law states that free particles move in straight lines at constant velocity. This can be understood as extremizing the proper time

$$\tau = \int_A^B d\tau = \int_A^B \sqrt{dt^2 - d\vec{x}.d\vec{x}} = \int_A^B \sqrt{1 - \frac{d\vec{x}}{dt}.\frac{d\vec{x}}{dt}}dt \equiv \int dt\sqrt{L}$$

The equations are

$$\frac{d}{dt}(L^{-\frac{1}{2}}\frac{dx^i}{dt}) = 0$$

Using $d\tau = \sqrt{L}dt$ we can write this as

$$\frac{d^2x^i}{d\tau^2} = 0$$

This gives the solution: $U^i \equiv \frac{dx^i}{d\tau} = a^i$ and $x^i = a^i\tau + b^i$. One can solve for t also:

$$\frac{d\tau}{dt} = \sqrt{L} = \sqrt{1 - (\frac{dx^i}{dt})^2} \implies U^0 \equiv \frac{dt}{d\tau} = \frac{1}{\sqrt{1-v^2}} \equiv \gamma$$

Also dividing both sides by $\frac{d\tau}{dt}$ and squaring we get

$$1 = (U^0)^2 - (U^i)^2$$

A manifestly Lorentz invariant way of doing this is to write

$$\tau = \int_A^B \sqrt{(\frac{dx^0}{ds})^2 - \frac{d\vec{x}}{ds}.\frac{d\vec{x}}{ds}}ds \equiv \int ds\sqrt{L}$$

where s is an arbitrary parameter, and one gets

$$\frac{d^2x^\mu}{d\tau^2} = 0$$

One can use the time component of the equation to say that $x^0 = a^0\tau + b^0$ and get the same equations as before for the i components.

This holds when there are no forces, including gravity.

2. If one starts with a Lagrangian

$$L = \int ds[-(\frac{dx^0}{ds})^2 + \frac{d\vec{x}}{ds}.\frac{d\vec{x}}{ds}] = \int ds\ \eta_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds}$$

which is a generalization of a free particle Lagrangian in non relativistic mechanics to a Lorentz invariant one, we obtain the same equations:

$$\frac{d^2x^\mu}{ds^2} = 0$$

It is often simpler to work with this, especially in the quantum theory, because there is no square root.

3. Einstein's idea is that gravitational forces can be included by curving space time. So particles still follow geodesics, but now in some different space time specified by a metric. In fact in the last section Newton's EOM in a gravitational field was reproduced by a particular metric. If there are non gravitational forces eg electric forces, then the particle deviates from the geodesic. It is also assumed here that the particle is small enough that it doesn't modify space time geometry. So the generalization would be to extremize

$$\tau = \int_A^B d\tau = \int_A^B \sqrt{-g_{\mu\nu}(x)dx^\mu dx^\nu} = \int_A^B \sqrt{-g_{\mu\nu}(x)\frac{dx^\mu}{ds}\frac{dx^\nu}{ds}}ds = \int_A^B \sqrt{L}ds \tag{30}$$

One expects two terms, one involving the second derivative $\frac{d^2x^\mu}{ds^2}$ as before, and also one term from differentiating $g_{\mu\nu}(x)$ w.r.t x, which will involve the product of two first derivatives $\frac{dx^\mu}{ds}$. Thus we expect

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} = 0 \tag{31}$$

4. Christoffel Coefficients: Γ's are called Christoffel Coefficients.

   Let us evaluate the Γ's:

   Let $l = \sqrt{L}$. We need various derivatives:

$$\begin{aligned}\frac{\partial l}{\partial x^\gamma} &= \frac{1}{2}L^{-\frac{1}{2}}\frac{\partial L}{\partial x^\gamma} = \frac{1}{2l}[-g_{\alpha\beta,\gamma}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds}] \\ \frac{\partial l}{\partial(\frac{dx^\gamma}{ds})} &= \frac{1}{2}L^{-\frac{1}{2}}[-2g_{\alpha\gamma}\frac{dx^\alpha}{ds}] = -\frac{1}{l}g_{\alpha\gamma}\frac{dx^\alpha}{ds} \\ \implies \frac{d}{ds}\frac{\partial l}{\partial(\frac{dx^\gamma}{ds})} &= -\frac{d}{ds}(\frac{1}{l}g_{\alpha\gamma}\frac{dx^\alpha}{ds})\end{aligned} \tag{32}$$

   Let us now use $lds = d\tau$, so

$$= -l\frac{d}{d\tau}(g_{\alpha\gamma}\frac{dx^\alpha}{d\tau}) \tag{33}$$

   So combining (32),(33) we get Lagrange's EOM

$$l[-\frac{1}{2}g_{\alpha\beta,\gamma}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} + \frac{d}{d\tau}(g_{\alpha\gamma}\frac{dx^\alpha}{d\tau})] = 0 \tag{34}$$

   Using $\frac{d}{d\tau}g_{\alpha\gamma} = g_{\alpha\gamma,\beta}\frac{dx^\beta}{d\tau}$ we get

$$[-\frac{1}{2}g_{\alpha\beta,\gamma}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} + g_{\alpha\gamma,\beta}\frac{dx^\beta}{d\tau}\frac{dx^\alpha}{d\tau} + g_{\alpha\gamma}\frac{d^2x^\alpha}{d\tau^2}] = 0 \tag{35}$$

   The second term can be symmetrised to get

$$g_{\alpha\gamma}\frac{d^2x^\alpha}{d\tau^2} + \frac{1}{2}(-g_{\alpha\beta,\gamma} + g_{\alpha\gamma,\beta} + g_{\beta\gamma,\alpha})\frac{dx^\beta}{d\tau}\frac{dx^\alpha}{d\tau} = 0 \tag{36}$$

$$\implies \frac{d^2x^\mu}{d\tau^2} + \underbrace{g^{\mu\gamma}\frac{1}{2}(-g_{\alpha\beta,\gamma} + g_{\alpha\gamma,\beta} + g_{\beta\gamma,\alpha})}_{\Gamma^\mu_{\alpha\beta}}\frac{dx^\beta}{d\tau}\frac{dx^\alpha}{d\tau} = 0 \tag{37}$$

   This is the equation for a geodesic. The coefficients $\Gamma^\mu_{\alpha\beta}$ are called Christoffel symbols or C..connection coefficients.

   This equation is **covariant**: This means that if we go to a different coordinate system we will get an equation that looks the same except with a transformed Γ. The new Γ is obtained using the transformed metric using the same formula as given in (37). Thus once we agree to this defn of how Γ changes under a coordinate change, we can say that the equation is covariant.

5. In practice the best way to obtain the coefficients is by using the same method, rather than use the general formula. Example: Flat space in polar coordinates.

$$S_{AB} = \int_A^B dS = \int_A^B d\sigma \; [(\frac{dr}{d\sigma})^2 + r^2(\frac{d\phi}{d\sigma})^2]^{\frac{1}{2}}$$

EOM:

$$\frac{d}{d\sigma}\frac{\partial L}{\partial(\frac{\partial x^\alpha}{\partial\sigma})} - \frac{\partial L}{\partial x^\alpha} = 0$$

$$\frac{d}{d\sigma}(\frac{1}{l}\frac{dr}{d\sigma}) - \frac{r}{l}(\frac{d\phi}{d\sigma})^2 = 0 \implies \frac{d}{dS}(\frac{dr}{dS}) - r(\frac{d\phi}{dS})^2 = 0$$

Thus $\Gamma^r_{\phi\phi} = -r$. And $\Gamma^r_{\phi r} = 0 = \Gamma^r_{rr}$.

Similarly

$$\frac{d}{dS}(r^2\frac{d\phi}{dS}) = 0 \implies \Gamma^\phi_{r\phi} = \frac{1}{r} \quad , \quad \Gamma^\phi_{rr} = 0 = \Gamma^\phi_{\phi\phi}$$

6. Note that if we change the parameter from S to some arbitrary $\sigma(S)$, then since $\frac{d}{dS} = \frac{d\sigma}{dS}\frac{d}{d\sigma}$, the first term $\frac{d^2x}{dS^2}$ in the geodesic equation will get a term $\frac{d^2\sigma}{dS^2}\frac{dx}{d\sigma}$. Thus we end up getting an additional term in the geodesic eqn involving one velocity. This σ is then not an "affine" parameter. Affine parameters give geodesic eqn in the standard form. Any linear transformation preserves affineness: $\sigma = aS + b$.

5.2 Solving the Geodesic Equation: Symmetries and Conserved Quantities

7. Geodesic eqns are four 2nd order DE's. The equations are not always easy to solve because they are non linear with non constant coefficients. If there are integrals of the motion, they simplify.

8.
$$U^\alpha U_\alpha = -1 \implies g_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = -1$$

This is constant along the trajectory and gives one conserved quantity always.

9. If there are symmetries we get addnl conserved quantities.

Thus suppose metric does not depend on x^1. Then $x^1 \to x^1 + \xi^1$ is a symmetry of the space-time. Thus

$$\frac{\partial L}{\partial x_1} = 0 \implies \frac{d}{d\sigma}\frac{\partial L}{\partial(\frac{dx^1}{d\sigma})} = 0$$

$$\implies \frac{\partial L}{\partial(\frac{dx^1}{d\sigma})} = \frac{1}{l}g_{1\beta}\frac{dx^\beta}{d\sigma} = g_{1\beta}\frac{dx^\beta}{d\tau} = constant$$

More generally

$$\xi^\alpha U^\beta g_{\alpha\beta} = \vec{\xi}.\vec{U} = constant$$

(or $\vec{\xi}.\vec{p} = constant$).

10. Let us apply this to the polar coordinate system. The final answer will give straight lines in a rather unusual parametrization: $\vec{U}$ is space-like:

$$\vec{U}.\vec{U} = (\frac{dr}{dS})^2 g_{rr} + (\frac{d\phi}{dS})^2 g_{\phi\phi} = 1 \implies (\frac{dr}{dS})^2 + r^2(\frac{d\phi}{dS})^2 = 1$$

Metric is independent of ϕ so

$$g_{\phi\phi}(\frac{d\phi}{dS}) = r^2(\frac{d\phi}{dS}) = constant \; = c$$

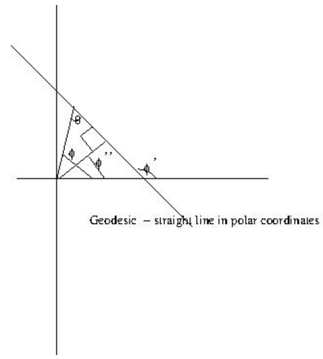


Figure 12: Geodesic is a straight line.

Note: If $c = 0$ the equation gives $\phi = const$ and $r = S$ which describes a straight line thru the origin. More generally: $(\frac{d\phi}{dS}) = \frac{c}{r^2}$ and $(\frac{dr}{dS}) = \sqrt{1 - \frac{c^2}{r^2}}$.

$$(\frac{d\phi}{dr}) = \frac{c}{r^2}\frac{1}{\sqrt{1-\frac{c^2}{r^2}}} \implies \int d\phi \ = c\int \frac{dr}{r^2}(1-\frac{c^2}{r^2})^{-\frac{1}{2}}$$

Let $\frac{c}{r} = sin\ \theta$. Then $-l\frac{dr}{r^2} = cos\ \theta d\theta$. So

$$\int d\phi \ = -\int d\theta \implies (\phi - \phi') = -\theta = sin^{-1}\frac{c}{r}$$

$$c = rsin(\phi - \phi') = rcos(\phi - \phi")$$

This is the eqn of a line where $\phi"$ has the interpretation of the angle made by the normal to the line at the origin. See Fig.9.

11. Geodesics on S^2.

$$dS^2 = d\theta^2 + sin^2\theta d\phi^2$$

Extremize

$$S = \int dS = \int d\sigma\sqrt{(\frac{d\theta}{d\sigma})^2 + sin^2\theta(\frac{d\phi}{d\sigma})^2} \equiv \int d\sigma\ \sqrt{L}$$

EOM

$$\frac{1}{2\sqrt{L}}2sin\ \theta\ cos\ \theta(\frac{d\phi}{d\sigma})^2 - \frac{d}{d\sigma}(\frac{1}{2\sqrt{L}}2\frac{d\theta}{d\sigma})$$

$$\implies [sin\ \theta\ cos\ \theta(\frac{d\phi}{dS})^2 - \frac{d}{dS}(\frac{d\theta}{dS})] = 0$$

For ϕ

$$\frac{d}{d\sigma}\frac{1}{2\sqrt{L}}2sin^2\ \theta(\frac{d\phi}{d\sigma}) = 0 \implies sin^2\ \theta\frac{d^2\phi}{dS^2} + 2sin\ \theta cos\ \theta\frac{d\theta}{dS}\frac{d\phi}{dS} = 0$$

$$\implies \frac{d^2\phi}{dS^2} + 2cot\ \theta\frac{d\theta}{dS}\frac{d\phi}{dS} = 0$$

Thus $\Gamma^\theta_{\phi\phi} = -\frac{1}{2}sin\ 2\theta$, $\Gamma^\theta_{\theta\phi} = \Gamma^\theta_{\theta\theta} = 0$. Also $\Gamma^\phi_{\phi\phi} = \Gamma^\phi_{\theta\theta} = 0$. And $\Gamma^\phi_{\theta\phi} = cot\ \theta$.

Special case: $\frac{d\phi}{dS} = 0$. Then $\frac{d}{dS}(\frac{d\theta}{dS}) = 0$. So $(\frac{d\theta}{dS}) = c \implies \theta = cS$ is a geodesic. The $\vec{U}.\vec{U} = 1$ constraint gives $c = 1$. Thus $\theta = S, \phi = const$ - great circles thru poles are geodesics.

5.3 Null Geodesics

What is the trajectory of a massless particle such as light ray? $d\tau^2 = 0$ - so it cannot be extremized! So start with **flat** space and covariantize:

In flat space we have say,

$$x = t$$

Can write this as

$$x^1 = U^1\lambda \quad ; \quad x^0 = U^0\lambda$$

with $U^0 = U^1$. Thus $\vec{U} = u(1,1,0,0)$ and $\vec{U}.\vec{U} = 0$. So more generally:

$$x^\alpha = U^\alpha\lambda \quad ; \quad \vec{U}.\vec{U} = 0$$

Thus

$$\frac{d^2x^\alpha}{d\lambda^2} = 0 = \frac{dU^\alpha}{d\lambda}$$

is the defining equation in flat space. What is the covariant curved space equation that reduces to this in flat space? By analogy with the earlier equation we write:

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

The only difference is that we have a constraint $\vec{U}.\vec{U} = 0$ rather than $\vec{U}.\vec{U} = -1$. And λ is not the proper time or proper distance - just a parameter along the trajectory. Thus for a photon we can take $U^\alpha = p^\alpha$ where $\vec{p}$ is the 4-momentum.

This equation is "covariant" in the sense defined earlier: if we write the equation in a different coordinate system, then we will get the same equation! Only the Γ's will be obtained from the transformed metric.

5.4 Riemann Normal Coordinates

12. The Riemann Normal Coordinates (RNC) is an example of a frame of a freely falling observer and is locally an inertial frame.

13. Construct a coordinate system at P as follows: P is the origin - coordinates $y^i = 0$. Consider any point Q. Draw geodesics from P thru Q. The geodesic is defined by the tangent vector at P. Call it ξ^i - assume that it is normalized to 1. Then along this geodesic let Q be at a parameter value s, choose s as proper distance. Assign coordinates

$$y^i = s\xi^i \tag{38}$$

Note that $\vec{\xi}$ is a *geometric* object - tangent vector at P to a geodesic and has a meaning independent of coordinate systems.

14. A point along the same geodesic at twice the distance will have y scaled by 2. So geodesic looks like a straight line if we plot it in Cartesian y coordinates system. In fact

$$\frac{d^2y^i}{ds^2} = 0$$

This implies that

$$\bar{\Gamma}(y)^i_{jk}\frac{dy^j}{ds}\frac{dy^k}{ds} = \bar{\Gamma}^i_{jk}\xi^j\xi^k = 0 \tag{39}$$

We use a bar to indicate that these are being evaluated in the RNC coordinates. This is a restriction on $\bar{\Gamma}$ at *all* points. However at any given point, $\vec{\xi}$ is fixed and not arbitrary. It is given by the slope (at the origin) of the geodesic that goes through that particular point. So only those components of $\bar{\Gamma}$ are zero.

At P the $\vec{\xi}$ are unrestricted and can point in any direction. So the above constraint implies that $\bar{\Gamma}^i_{jk}(P) = 0$. Again at P, in the RNC, since the geodesic equation is $\frac{d^2y^i}{ds^2} = 0$, which is the equation of a free particle in an inertial frame, the RNC constitute an inertial or freely falling frame at P.

15. Coordinate system breaks down when geodesics cross, so it is valid only in a finite domain.

16. Since $\vec{\xi}$ is a geometric object so is $\vec{y}$.

17. (See Petrov, Eisenhart) Suppose we begin with a coordinate system x. And P is at x_0. Then a point such as Q on the geodesic has an expansion:

$$x^i(s) = x^i_0 + \xi^i s + \frac{1}{2}\frac{d^2x^i}{ds^2}|_0 s^2 + \frac{1}{3!}\frac{d^3x^i}{ds^3}|_0 s^3 + ...$$

From the geodesic equation $\frac{d^2x^i}{ds^2}|_0 = (\Gamma^i_{jk})_0$ and also

$$\frac{d^3x^i}{ds^3} + \Gamma^i_{jkl}\frac{dx^j}{ds}\frac{dx^k}{ds}\frac{dx^l}{ds} = 0$$

where

$$\Gamma^i_{jkl} = \frac{1}{3}P[\partial_l\Gamma^i_{jk} - \Gamma^i_{ak}\Gamma^a_{jl} - \Gamma^i_{ja}\Gamma^a_{kl}] \tag{40}$$

and P denotes cyclic permutations,

$$\Gamma^i_{jklm} = \frac{1}{4}P[\partial_m\Gamma^i_{jkl} - \Gamma^i_{ajk}\Gamma^a_{lm} - \Gamma^i_{lak}\Gamma^a_{jm} - \Gamma^i_{lja}\Gamma^a_{km}] \tag{41}$$

and so on. Thus we get:

$$x^i(s) = x^i_0 + \xi^i s - \frac{1}{2}(\Gamma^i_{jk})_0\xi^j\xi^k s^2 - \frac{1}{3!}(\Gamma^i_{jkl})_0\xi^j\xi^k\xi^l s^3 + ...$$

In the y coordinates (RNC) the Taylor expansion stops at the linear term

$$y^i = \xi^i s$$

Thus going to RNC involves setting all these Γ's to zero. At P ($x = 0$), any $\vec{\xi}$ is allowed, so the constraint involves setting the Γ's to zero.

$$(\Gamma^i_{jk})_0\xi^j\xi^k = (\Gamma^i_{jkl})_0\xi^j\xi^k\xi^l = ... = 0$$

Since the ξ^i can point in any direction, it must be that the symmetrised Gamma's vanish: $\Gamma^i_{(jkl...)} = \tilde{\Gamma}^i_{jkl...} = 0$ Note that there are 40 Γ^-jk's and 80 Γ^i_{jkl} - exactly the number of parameters we found in coordinate parametrizations. Note that $\frac{\partial^n x'^i}{\partial x^j\partial x^k...}$ has the same index structure as the symmetrised Γ.

Alternately in the RNC these can be thought of as the Taylor expansion of constraints (39). Thus as mentioned above in the RNC also we have

$$(\bar{\Gamma}^i_{jk})_0 = 0 = (\bar{\Gamma}^i_{jkl})_0 = ... = 0$$

From (40) and (41) we see that the higher Γ's involve $(\bar{\Gamma}^i_{jk})_0$ in the second term onwards. So all these are zero. Thus

$$\bar{\Gamma}^i_{jkl} = \frac{1}{3}P[\partial_l\bar{\Gamma}^i_{jk}], \quad \Gamma^i_{jklm} = \frac{1}{4}P[\partial_m\bar{\Gamma}^i_{jkl}], ...$$

They are all symmetrized derivatives of Γ^i_{jk}.

18. Scalar functions can be expanded in a Taylor series and coefficients are tensors at the origin:

$$\Phi(y) = \Phi(0) + y^i \partial_i \Phi(0) + \frac{1}{2} y^i y^j \partial_i \partial_j \Phi(0) + ...$$

y^i are vectors at the origin, so the derivatives of Φ are all tensors. In a general coordinate system these would become Covariant derivatives plus curvature tensors. See Petrov.

19. Example: As an example we consider unit S^2 and choose for P the North Pole $\theta = 0$. Geodesics thru P are $\phi = const$ and $\theta = cs$. From $\frac{d\theta}{ds}\frac{d\theta}{ds} = 1$ we get $c = 1$. Thus $\theta = s$.

    We can now choose Cartesian coordinates: $y^i = s\xi^i$ where $\vec{\xi} = (cos\ \phi, sin\ \phi)$. $\vec{y} = s(cos\ \phi, sin\ \phi) = \theta(cos\ \phi, sin\ \phi)$. Note that at $\theta = \pi$ this coordinate system breaks down, because we get different $y = \pi(cos\ \phi, sin\ \phi)$ labelled by ϕ, for the South Pole. At the N Pole since $\theta = 0$ this problem is not there. The RNC coordinate system breaks down when geodesics cross.

    Alternately one can use polar coordinates r, χ and set $r = s$. So $\vec{\xi} = (\xi^r = 1, \xi^\chi = 0)$. But Polar coordinates are not well defined at the Poles so there is some ambiguity.

20. **Freely falling frame:** If a particle is in free fall we can construct a freely falling frame *along his trajectory*. Take for t his proper time. At any τ along the world line. construct space like geodesics and use RNC for them. Attach gyroscopes along these directions. At a later time use these gyroscopes to define the directions and construct RNC in the spatial directions. Thus $x^\mu = (\tau, sn^i)$. Where s is the spatial proper distance along the spacelike geodesic. Thus $\frac{d^2x^\mu}{ds^2} = 0$ and $\frac{d^2x^\mu}{d\tau^2} = 0$ all along the world line, i.e. at each point along the trajectory, all the coordinate axes are along geodesics, so $\Gamma = 0$ *at each point along the world line.*

    This coordinate system or frame is the closest one can get to an inertial frame in curved space-time.

    In practice we enclose a freely falling particle in an elevator cabin and fall with it. Make sure that our space coordinates are not rotating by observing three gyroscopes with orthogonal orientations.

5.5 Basis Vectors for accelerated observers

21. Consider motion of an observer with constant 4 acceleration that has been discussed earlier (reproduced here):

    Choose

$$x(\tau) = a^{-1} cosh\ (a\tau), \quad t(\tau) = a^{-1} sinh\ (a\tau) \tag{42}$$

    4-vel

$$U^1 = \frac{dx}{d\tau} = sinh\ (a\tau), \quad U^0 = \frac{dt}{d\tau} = cosh\ (a\tau) \tag{43}$$

    4 acc

$$a^0 = \frac{dU^0}{d\tau} = asinh\ (a\tau), \quad a^1 = \frac{dU^1}{d\tau} = acosh\ (a\tau) \tag{44}$$

    Note that $\tilde{a}.\tilde{a} = (a^1)^2 - (a^0)^2 = a^2$ is constant.

    What basis vectors should he choose? ANS: At any instant consider an inertial observer who has the same velocity and is right next to him. This person's clock ticks exactly the same way as that of the accelerated observer - for *small enough time intervals* - as long as their world lines coincide. Similarly the lengths they measure will be the same - for small enough rods. So use the same basis vectors. Let them be orthonormal. In particular let $\hat{e}_0 = \vec{U}$ their instantaneous velocity. So from (43) we see that

$$\hat{e}_0 = \tilde{U} = (cosh\ (a\tau), sinh\ (a\tau), 0, 0)$$

in the basis used by the accelerating observer. This is the same $\hat{e}_0$ used by the inertial comoving observer. It is obtained by a Lorentz boost from $(1,0,0,0)$ - which is the time direction for the inertial observer at rest. For $\hat{e}_1$ we need to use $\hat{e}_1.\hat{e}_0 = 0$ to get

$$\hat{e}_1 = (sinh\ (a\tau), cosh\ (a\tau), 0, 0)$$

We have also set $\hat{e}_1.\hat{e}_1 = 1$. Thus $\hat{e}_1 = \frac{1}{a}\tilde{a}$. (Remember that $\tilde{a}.\tilde{U} = 0$.) Finally

$$\hat{e}_2 = (0,0,1,0), \quad \hat{e}_3 = (0,0,0,1)$$

Again these are given in the basis of the inertial observer at rest and coincide with the bases of the comoving inertial observer.

Since the observer is accelerating his basis keeps changing - as the explicit τ dependence shows. Note that $\hat{e}_2, \hat{e}_3$ are independent of time. This is because a Lorentz boost along the x-axis is all that is being done. No spatial rotation. So the transverse basis vectors are untouched. In practice one can attach gyroscopes pointing in x, y, z directions to ensure that they do not undergo spatial rotation. The Lorentz boost is performed because the observer is speeding up.

22. **Fermi-Walker Transport**

This can be generalized to arbitrary observers: From each instant to the next perform the Lorentz boost from one co moving observer to the next. So the basis vectors of one co moving observer after boosting becomes the basis vectors of the comoving observer at the next instant. No spatial rotations - only boosts. Since the direction of the boost keeps changing we construct a rotation operator that always boosts in the direction of the instantaneous acceleration: For any vector b^μ:

$$\frac{db^\mu}{d\tau} = -\Omega^{\mu\nu} b_\nu, \quad \Omega^{\mu\nu} = a^\mu U^\nu - a^\nu U^\mu \tag{45}$$

Consider the case of acceleration in the x-direction. Using (43) and (44),

$$\Omega^{01} = -(a^0 U^1 - a^1 U^0) = a$$

Let $\tilde{b} = (b^0, b^1, 0, 0)$. Then the equation gives

$$\begin{aligned} \frac{db^0}{d\tau} &= -\Omega^{01} b_1 = a b_1 = a b^1 \\ \frac{db^1}{d\tau} &= -\Omega^{10} b_0 = -a b_0 = a b^0 \end{aligned}$$

Thus we get

$$\delta b^0 = a\delta t b^1 \quad ; \quad \delta b^1 = a \delta t b^0$$

Noting that $a\delta t = \delta v$ is the increase in velocity of the boosted frame, we see that the change above is a Lorentz boost on a 4-vector, as required.

One can check that $\Omega^{\mu\nu} U_\nu = -a^\mu$ as it should be. Also if $\tilde{w}$ is orthogonal to $\tilde{U}$ and $\tilde{a}$ (i.e. it is a spacelike vector that does not lie in the plane of the boost - like $\hat{e}_2, \hat{e}_3$ in the earlier example), then $\Omega^{\mu\nu} w_\nu = 0$. Again no spatial rotation is performed. In practice - attach gyroscopes. Any vector that obeys this rule (45) is said to be **Fermi-Walker** transported.

This tetrad of an accelerated observer (as described above) is thus Fermi-Walker transported.

23. **Stationary Observer**

Consider an observer who is stationary (in some coordinate system). A stationary observer is a special kind of accelerated observer. Note that anyone who is not in free fall is accelerated. In general his frame is not inertial (except in flat space). What is a good coordinate system for him to use? He can use $\hat{e}_0 = \tilde{U}_{obs}$.i.e.

$$\tilde{U}_{obs} = \underbrace{U^0}_{=1} \hat{e}_0 + 0\hat{e}_i$$

Let

$$\hat{e}_0.\hat{e}_i = 0 \quad ; \quad \hat{e}_i.\hat{e}_j = \delta_{ij}$$

But these are not inertial coordinates.

6 Schwarzschild Metric

6.1 Introduction

1. The exterior geometry of a spherically symmetric star of mass M is described by the following metric:

$$ds^2 = -(1 - \frac{2GM}{c^2 r})(cdt)^2 + \frac{dr^2}{(1 - \frac{2GM}{c^2 r})} + r^2(d\theta^2 + sin^2\,\theta d\phi^2) \quad (46)$$

We will use units where $c = 1$ and also set $G = 1$. G has dimensions of $length^2$. So we have defined a unit of length (and also therefore mass).

$$ds^2 = -(1 - \frac{2M}{r})dt^2 + \frac{dr^2}{(1 - \frac{2M}{r})} + r^2(d\theta^2 + sin^2\,\theta d\phi^2) \quad (47)$$

2. **Interpretation of r:** r is a radial coordinate in a polar coordinate system but since the metric is different, one cannot interpret r as the physical distance from the origin. The physical distance is $\int_0^r \frac{dr}{\sqrt{1-\frac{2GM}{r}}}$. It is much easier to interpret it in terms of area: $A = 4\pi r^2$ is the area of a spherical surface of radius r surrounding the star.

Interpretation of t: This is a parameter that labels time. For the stationary observer at infinity t is the physical time.

Interpretation of M: For large r the space time becomes that of a static Newtonian metric introduced earlier with $\phi = -\frac{GM}{r}$. This metric reproduced Newtonian EOM in a gravitational potential. From this we see that M is the mass as seen by an observer at ∞. This is a general technique - to interpret parameters in GR, we go far away and make contact with Newtonian mechanics. Then we can see the physical significance of the parameter. Otherwise it is just a parameter in a solution.

"Schwarzschild radius": We can define $R_s = \frac{2GM}{c^2}$. This is where the metric turns singular. It is called the "Schwarzschild radius". For the sun $R_s = 2.95km << R_{sun} = 6.9 \times 10^5 km$. Since the metric is valid only outside the surface of the star, this singularity does not bother us. If the star radius is less than R_s, then we have to worry. That would be called a "black hole".

Note that $\frac{GM}{c^2} \sim L$ has dimensions of length. Also since Mc has dimensions of momentum, $\frac{\hbar G}{c^3} \sim L\frac{\hbar}{Mc} \sim L^2$. The length defined by $G, c, \hbar$ is called the Planck length l_P. In units where $c = \hbar = 1$, $G \approx l_P^2$. G thus introduces a length scale into physics (unlike electric charge).

In GR we often use geometrical units where $G = 1$ - this is a choice of the unit of length. Thus GM defines a length. It is, up to a factor of 2, the Schwarzschild radius. Thus if we give M in km, it means we are really talking of $\frac{GM}{c^2}$. Thus we can say the mass of the sun is about 1.5 km! This should not be confused with $\frac{\hbar}{Mc}$ (Compton wavelength) which also gives a length associated with a mass.

Let us compare Compton Wavelength and Gravitational length of particles. Take a proton. The gravitational length (say R_G), is Gm_{proton}. Let us work in units of Gev for mass and $(Gev)^{-1}$ for length (this is around a fermi). $G = l_p^2$. So $R_G = m_{proton} l_p^2$. $m_{proton} = 1Gev$ and l_p is $10^{-19}Gev^{-1}$. Thus $R_G = 10^{-38}Gev^{-1} = 10^{-19}l_p$. Thus the gravitational length or Schwarzschild radius of a proton is much less than l_p. What about the Compton Wavelength: This is just $\frac{1}{m_{proton}} = 1Gev^{-1} = 10^{19}l_p$. Thus the Compton Wavelength of a proton is larger than its Schwarzschild radius by a factor of 10^{38}. A planck mass particle has both lengths of the same order. Thus a planck mass particle is very likely a black hole. It could also be that for a particle what is relevant is its actual size and not the Compton wavelength. In that case if we think of an electron as having zero size (so far there is no experimental evidence for a finite size up to scales of several hundred Gev), it is clearly less than its Schwarzschild radius and then we would think of all point particles as black holes! Clearly these are open questions.

3. **Killing Vectors**

The metric coefficients do not depend on t. Furthermore it has spherical symmetry. This is the unique solution of Einstein's equations that has these properties.

These are symmetries. We have already seen that when $x^\alpha \to x^\alpha + \xi^\alpha$ is a symmetry of the metric coefficient we have conserved quantities along geodesics. $\vec{\xi}$ is a diffeomorphism of our manifold. Diffeomorphisms are generated by vector fields. In this case we can call it $\vec{\xi}$. The formal statement of symmetry is that the Lie Derivative of the metric should vanish

$$\mathcal{L}_\xi g_{\mu\nu} = 0 = \xi^\alpha g_{\mu\nu,\alpha} + \xi^\alpha{}_{,\mu} g_{\alpha\nu} + \xi^\alpha{}_{,\nu} g_{\mu,\alpha} \tag{48}$$

For simple translations $x^\alpha \to x^\alpha + \xi^\alpha$ where ξ is constant, this reduces to the statement that the metric coefficients should be independent of x^α. But (48) is the more general definition of symmetry. It can be shown that $\vec{\xi}.\vec{U}$ is a constant along the geodesic as we did earlier. ξ is called a Killing Vector field.

Let us work in the (t, r, θ, ϕ) Schwarzschild coordinate system, with basis vectors $\frac{\tilde{\partial}}{\partial t}, \frac{\tilde{\partial}}{\partial r}, \frac{\tilde{\partial}}{\partial \theta}, \frac{\tilde{\partial}}{\partial \phi}$. These are not normalized. Thus for instance $\frac{\tilde{\partial}}{\partial t}.\frac{\tilde{\partial}}{\partial t} = g_{tt} = -(1 - \frac{2M}{r})$.

$$\tilde{\xi} = (1, 0, 0, 0) = \frac{\tilde{\partial}}{\partial t}$$

is a Killing vector corresponding to time translations. Similarly it is easy to see that $\phi \to \phi + \eta$ is a symmetry clearly. This is part of the spherical symmetry. Thus

$$\eta = (0, 0, 0, 1) = \frac{\tilde{\partial}}{\partial \phi}$$

is a Killing vector. The other two rotational Killing vectors are not so obvious.

Corresponding to these two ξ, η, we have two conserved quantities:

$$e = -\tilde{U}.\tilde{\xi}, \qquad l = \tilde{U}.\tilde{\eta}$$

It is easy to see that e is the energy per unit mass of the particle ($-\tilde{p}.\tilde{\xi}$ is energy of the particle), and l is angular momentum in the ϕ direction per unit mass of the particle.

6.2 Orbits of Particles

4. **Gravitational Redshift:**

Use the fact that $\tilde{p}.\tilde{\xi}$ is conserved along a photon orbit, where $\tilde{p}$ is the 4-momentum of the photon. The energy of the photon as measured by an observer with 4-velocity $\tilde{U}_{obs}$ is $E = -\tilde{U}_{obs}.\tilde{p}$. This is clear from the fact that for the observer $E = -\hat{e}_0.\tilde{p}$ by defn. where $\hat{e}_0$ is the basis vector in the time direction in his orthonormal basis vector set and $\hat{e}_0 = \tilde{U}_{obs}$. In his orthonormal system $\tilde{U}_{obs} = (1, 0, 0, 0)$.

Let us write $\tilde{U}_{obs}$ in the Schwarzschild coordinate system. Assume the observer is stationary. This means

$$U^r = U^\theta = U^\phi = 0$$

Also

$$U^t U^t g_{tt} = -1 \implies U^t = \frac{1}{\sqrt{1 - \frac{2M}{r}}}$$

Thus

$$\tilde{U} = \frac{1}{\sqrt{1 - \frac{2M}{r}}}(1, 0, 0, 0) = \frac{1}{\sqrt{1 - \frac{2M}{r}}}\tilde{\xi}$$

Thus

$$E(r) = \tilde{p}.\tilde{U} = \frac{1}{\sqrt{1-\frac{2M}{r}}}\tilde{p}.\tilde{\xi}$$

Since $\tilde{p}.\tilde{\xi}$ is conserved along the photon trajectory, we see that the observed energy depends on the radial location of the observer as:

$$\frac{E(r)}{E(\infty)} = \frac{1}{\sqrt{1-\frac{2M}{r}}} \approx 1 + \frac{M}{r} for\ large\ r$$

The larger the r, the smaller the energy (and frequency) of the photon. This is the gravitational redshift.

5. **Particle Orbits:**

Let $\tilde{U} = (\frac{dt}{d\tau}, \frac{dr}{d\tau}, \frac{d\theta}{d\tau}, \frac{d\phi}{d\tau})$ be the 4-velocity of the particle. Conserved quantities:

$$e = -\tilde{\xi}.\tilde{U} = (1-\frac{2M}{r})U^0 = (1-\frac{2M}{r})\frac{dt}{d\tau} = energy/rest\ mass \tag{49}$$

We have written it using the Schwarzschild coordinates.

$$l = \tilde{\eta}.\tilde{U} = r^2 sin^2\ \theta\frac{d\phi}{d\tau} = r^2\frac{d\phi}{d\tau} = ang\ mom/rest\ mass \tag{50}$$

(If you want to put back c: $l \to \frac{l}{c}$ and $\frac{2GM}{r} \to \frac{2GM}{c^2 r}$. $\frac{GM}{c^2}$ is a length.)

Conservation of l means the plane of the orbit is fixed - as in Newt. mech. Thus consider a particle whose orbit has $\frac{d\phi}{d\tau} = 0$ at some instant. Then it is moving along a longitude. $\theta \neq 0$. Thus $l = 0$ at that instant. But then $l = 0$ *always*. So $\frac{d\phi}{d\tau} = 0$ always. So it always moves along a longitude. Thus the plane of this orbit is fixed. Once this is proved for this orbit, by rotational symmetry it is true for all orbits. For simplicity we pick the equatorial plane with $\theta = \frac{\pi}{2}$ from now on.

We have three equations for three unknowns: U^t, U^r, U^ϕ ($U^\theta = 0$).

$$U^\phi = \frac{d\phi}{d\tau} = \frac{l}{r^2}, \quad U^t = \frac{dt}{d\tau} = \frac{e}{(1-\frac{2M}{r})}$$

$$-\frac{e^2}{1-\frac{2M}{r}} + \frac{(U^r)^2}{(1-\frac{2M}{r})} + \frac{r^2 l^2}{r^4} = -1$$

$$\implies \underbrace{\frac{1}{2}(\frac{dr}{d\tau})^2}_{K.E.} + \underbrace{[-\frac{M}{r} + \frac{l^2}{2r^2} - \frac{Ml^2}{r^3}]}_{V_{eff}(r)=P.E.} = \underbrace{\frac{e^2-1}{2}}_{Total\ energy} \tag{51}$$

If we put back the c, G we get

$$\implies \underbrace{\frac{1}{2c^2}(\frac{dr}{d\tau})^2}_{K.E.} + \underbrace{\frac{1}{c^2}[-\frac{GM}{r} + \frac{l^2}{2r^2} - \frac{GMl^2}{c^2 r^3}]}_{\frac{1}{c^2}V_{eff}(r)=P.E.} = \underbrace{\frac{e^2-1}{2}}_{Total\ energy} \tag{52}$$

The first term in V_{eff} is the usual Newtonian grav potential. The second term is the usual centrifugal barrier. The last term $\frac{GMl^2}{c^2 r^3}$ is the deviation from the Newtonian $\frac{1}{r^2}$ force law. The Runge-Lenz vector is thus not conserved and the axis of the ellipse precesses.

A physical way to understand the third term is as follows: For fixed l the particle has KE $= \frac{1}{2}\frac{l^2}{r^2}m$. This contributes towards the mass of the particle: $\Delta m = \frac{1}{2}\frac{l^2}{r^2 c^2}m$. This in turn makes the gravitational

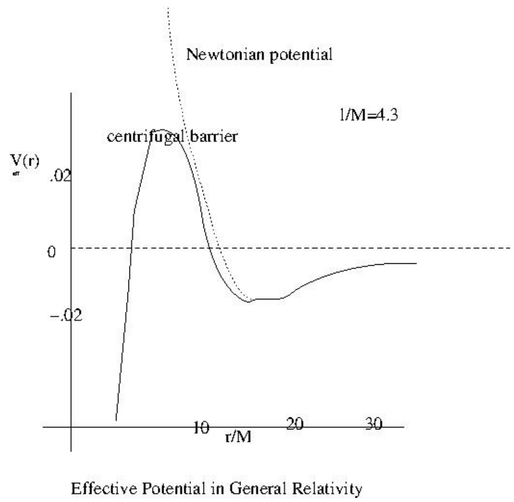


Figure 13: The effective potential in GR has an infinitely deep "pit" unlike in Newtonian mechanics

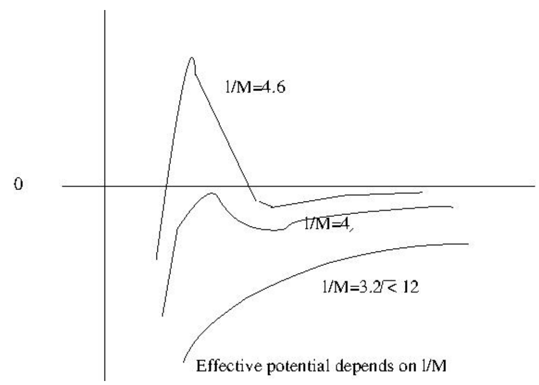


Figure 14: Dependence on l/M

potential energy more negative by $-\frac{GM\Delta m}{r} = -\frac{1}{2}m\frac{GMl^2}{r^3c^2}$. The extra term in the effective potential is just this term (per unit rest mass) - except for a factor of 2 - for which we need a precise theory - GR.

Note that for large r, $V(r)$ goes as $-\frac{1}{r}$ (usual) but for small r it goes as $-\frac{1}{r^3}$ (unusual - usually there would be just the positive centrifugal barrier). See figure 7,8,9 (apologies for the poor quality!).

The extrema of the potential are obtained by $\frac{dV_{eff}}{dr} = 0$ and are at

$$r = \frac{l^2}{2M}[1 \pm \sqrt{1 - 12(\frac{M}{l})^2}] \tag{53}$$

Note that :if $l^2 < 12M^2$ there are no extrema. i.e. need a minimum l - and hence centrifugal barrier - to have extrema.

6. **Radial Plunge Orbits:**

Take $l = 0$ -no barrier. Particle falls in. Take $e = 1$. This means "Total energy" is zero. Particle starts

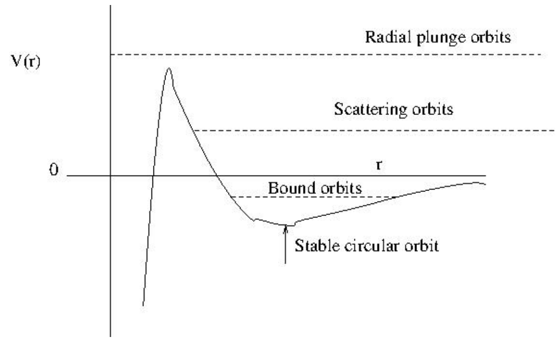


Figure 15: Type of orbits. The dashed line is the value of $e^2 - 1$.

from rest.

$$\frac{1}{2}(\frac{dr}{d\tau})^2 - \frac{M}{r} = 0 \implies \frac{dr}{d\tau} = U^r = -\sqrt{\frac{2M}{r}} \quad (infalling)$$

$$e = 1 \implies \frac{dt}{d\tau} = \frac{1}{1 - \frac{2M}{r}}$$

Thus

$$U^\mu = (\frac{1}{1 - \frac{2M}{r}}, -\sqrt{\frac{2M}{r}}, 0, 0) \tag{54}$$

One can solve for $r(\tau)$:

$$\int dr\ r^{\frac{1}{2}} = -\int \sqrt{2M} d\tau$$

$$\implies r^{\frac{3}{2}} = \frac{3}{2}\sqrt{2M}(\tau_* - \tau) \tag{55}$$

Note that the solution is completely non singular and from any (finite) r the particle reaches $r = 0$ in finite proper time. In contrast let us solve for $r(t)$:

$$\frac{dt}{dr} = \frac{\frac{dt}{d\tau}}{\frac{dr}{d\tau}} = -\frac{1}{1 - \frac{2M}{r}}\sqrt{\frac{r}{2M}}$$

This gives

$$\int dt = -\int dr\ \frac{1}{1 - \frac{2M}{r}}\sqrt{\frac{r}{2M}}$$

Make the substitution $r = y^2$. The final result is

$$t = t^* + 2M[-\frac{2}{3}(\frac{r}{2M})^{\frac{3}{2}} - 2(\frac{r}{2M})^{\frac{1}{2}} + ln\ |\frac{(\frac{r}{2M})^{\frac{1}{2}} + 1}{(\frac{r}{2M})^{\frac{1}{2}} - 1}|] \tag{56}$$

t^* is an integration constant. Clearly the coordinate time taken to reach $r = 2M$ diverges. Thus the outside observer at infinity or far away but at a finite distance (for whom coordinate time is the physical time) never quite sees the particle fall in.

7. **Escape velocity**

Consider an orbit with $l = 0,\ e = 1$. Then at infinity (63) says that since the RHS is zero, and $V_{eff}(\infty) = 0,\ \frac{dr}{d\tau} = 0$. Thus the particle is at rest. Let us evaluate the velocity of this particle as seen by a stationary observer located at r. The 4-velocity of the observer is:

$$U_{obs} = (U^t_{obs}, 0, 0, 0) \quad U^t_{obs} U^t_{obs} g_{tt} = -1 \implies U^t_{obs} = \frac{1}{\sqrt{1 - \frac{2M}{r}}}$$

Let velocity of the particle be $U = (U^t, U^r, 0, 0)$.

$$-U^t g_{tt} = e = 1 \implies U^t = \frac{1}{1 - \frac{2M}{r}}$$

$$U.U = -1 \implies U^t U^t g_{tt} + U^r U^r g_{rr} = -\frac{1}{1 - \frac{2M}{r}} + (U^r)^2 \frac{1}{1 - \frac{2M}{r}} = -1$$

$$\therefore U^r = \pm\sqrt{\frac{2M}{r}} \implies U = (\frac{1}{1 - \frac{2M}{r}}, \pm\sqrt{\frac{2M}{r}}, 0, 0)$$

This gives the 4-velocity of the particle. The energy of the particle as seen by the stationary observer is $-U_{obs}.mU$:

$$-U_{obs}.mU = \frac{1}{\sqrt{1 - \frac{2M}{r}}} \frac{m}{1 - \frac{2M}{r}} (1 - \frac{2M}{r}) = \frac{m}{\sqrt{1 - \frac{2M}{r}}}$$

Using the flat space formula $E = \gamma m = \frac{m}{\sqrt{1-v^2}}$ he would say that the particle has 3-velocity $v = \sqrt{\frac{2M}{r}}$. One can also check this by calculating $\frac{U^r U^r g_{rr}}{U^t U^t g_{tt}} = \frac{2M}{r}$. This is the velocity squared calculated in the orthonormal frame of the **stationary observer**. His orthonormal basis vectors are (as expressed in the Schwarzschild basis: $\partial_t = (1,0,0,0),\quad \partial_r = (0,1,0,0),\quad \partial_\theta = (0,0,1,0),\quad \partial_\phi = (0,0,0,1)$,

$$\hat{e}_0 = (\frac{1}{\sqrt{1 - \frac{2M}{r}}}, 0, 0, 0), \quad \hat{e}_r = (0, \sqrt{1 - \frac{2M}{r}}, 0, 0), \quad \hat{e}_\theta = (0, 0, \frac{1}{r}, 0), \quad \hat{e}_\phi = (0, 0, 0, \frac{1}{rsin\ \theta}) \tag{57}$$

Thus the 4-velocity U of the particle in this orthonormal basis, where he uses his flat space formulae, is:

$$U = (\frac{1}{\sqrt{1 - \frac{2M}{r}}}, \frac{\sqrt{\frac{2M}{r}}}{\sqrt{1 - \frac{2M}{r}}}, 0, 0) \equiv (\gamma, \gamma v^r, 0, 0)$$

This also gives the same value v^r. Note that this is the escape velocity and has the same value as in Netwonian mechanics.

8. **Circular Orbits**

   The minimum of the effective potential is given in (53) to be

$$r_{MIN} = \frac{l^2}{2M}[1 + \sqrt{1 - 12(\frac{M}{l})^2}]$$

   (The negative sign gives a maximum.) The smallest value is when $12(\frac{M}{l})^2 = 1$ or $l^2 = 12M^2$. This gives what is called the "innermost stable circular orbit",

$$r_{ISCO} = 6M$$

The angular velocity $\Omega = \frac{d\phi}{dt}$ is given by

$$\Omega = \frac{d\phi}{dt} = \frac{\frac{d\phi}{d\tau}}{\frac{dt}{d\tau}}$$

Plugging in

$$\frac{dt}{d\tau} = \frac{e}{(1-\frac{2M}{r})} \quad ; \quad \frac{d\phi}{d\tau} = \frac{l}{r^2}$$

$$\Omega = \frac{l}{e}\frac{1}{r^2}(1-\frac{2M}{r}) \tag{58}$$

For a circular orbit $\frac{dr}{d\tau} = 0$. Thus $V_{eff}(r) = \frac{e^2-1}{2}$. Thus

$$1 - \frac{2M}{r} + \frac{l^2}{r^2} - \frac{2Ml}{r^3} = (1-\frac{2M}{r})(1+\frac{l^2}{r^2}) = e^2 \tag{59}$$

We have one more requirement, that r should be a minimum of $V_{eff}(r)$ (which gave the condition (53)). This is

$$\frac{dV_{eff}}{dr} = 0 = \frac{2M}{r^2}(1+\frac{l^2}{r^2}) + (1-\frac{2M}{r})(-\frac{2l^2}{r^3})$$

$$\implies (1+\frac{l^2}{r^2}) = (1-\frac{2M}{r})\frac{l^2}{rM}$$

Inserting this in (59) gives

$$e^2 = (1-\frac{2M}{r})^2\frac{l^2}{rM} \implies \frac{l}{e} = \frac{(Mr)^{\frac{1}{2}}}{(1-\frac{2M}{r})}$$

This gives for Ω from (58)

$$\Omega^2 = \frac{M}{r^3} \tag{60}$$

which is the same form as Kepler's law. This is a coincidence - if we had used proper time to define Ω it would have given a different answer. Thus we can write

$$\tilde{U} = U^t(1,0,0,\Omega)$$

If we require $\tilde{U}.\tilde{U} = -1$ we can evaluate U^t:

$$(U^t)^2 g_{tt} + \Omega^2 (U^t)^2 g_{\phi\phi} = -1$$

$$(U^t)^2[1-\frac{2M}{r} - \Omega^2 r^2] = 1 \implies U^t = \frac{1}{\sqrt{1-\frac{3M}{r}}}$$

(This is to be contrasted with U^t for a stationary particle with $\Omega = 0$, $U^t = \frac{1}{\sqrt{1-\frac{2M}{r}}}$.)

9. **Gravitational Binding vs thermonuclear fusion:** The energy of a particle/unit rest mass is e. This is 1 for a particle at rest at infinity. Otherwise it is $-\xi.U = (1-\frac{2M}{r})U^t$. For a circular orbit at $r = 6M$ we have seen that $U^t = \frac{1}{\sqrt{(1-\frac{3M}{r})}}$. Thus when a particle falls in from infinity to this orbit the energy released is:

$$1 - e = 1 - (1-\frac{2M}{r})\frac{1}{\sqrt{(1-\frac{3M}{r})}} = 1 - \frac{2\sqrt{2}}{3} \approx .057$$

In thermonuclear fusion 4 hydrogen atoms converts to helium: The released energy is about 27 Mev - which is an efficiency of $\frac{27}{4\times 938} \approx .007$. So BH are more efficient. In rotating BH we can get up to .42. So very energetic processes (active galactic nuclei) require BH.

10. **Precession of Orbits**

To get shape we need ϕ vs r.

$$\mathcal{E} = \frac{1}{2}(\frac{dr}{d\tau})^2 + V_{eff}(r) \implies \frac{dr}{d\tau} = \pm\sqrt{2(\mathcal{E} - V_{eff}(r)}$$

Also $\frac{d\phi}{d\tau} = \frac{l}{r^2}$. Thus

$$\frac{d\phi}{dr} = \frac{\pm l}{r^2\sqrt{e^2 - (1-\frac{2M}{r})(1+\frac{l^2}{r^2})}}$$

Thus the angle between r_1 and r_2 is

$$\Delta\phi = \int_{r_1}^{r_2} \frac{d\phi}{dr}dr = \int_{r_1}^{r_2} \frac{\pm l}{r^2\sqrt{e^2 - (1-\frac{2M}{r})(1+\frac{l^2}{r^2})}}dr$$

If we choose r_1, r_2 to be the two turning points, then the total angle from r_1 to r_1 - the deviation from 2π of which gives the precession- is twice the angle from r_1 to r_2. Note that the values of r_1, r_2 are fixed by energy considerations. This is whether or not there is precession. So

$$\Delta\phi_{precession} = 2\int_{r_1}^{r_2} \frac{\pm l}{r^2\sqrt{e^2 - (1-\frac{2M}{r})(1+\frac{l^2}{r^2})}}dr$$

where the limits are the turning points. Turning point is where the denominator vanishes (by defn). Putting back c, G we can rewrite it as

$$2\int_{r_1}^{r_2} \frac{l}{r^2[\underbrace{c^2(e^2-1) + \frac{2GM}{r} - \frac{l^2}{r^2}}_{Newtonian} + \underbrace{\frac{2GMl^2}{c^2r^3}}_{GR\ correction}]^{\frac{1}{2}}}dr \qquad (61)$$

(Note: $\frac{e^2-1}{2} \approx O(\frac{1}{2}(\frac{dr}{d\tau})^2) \approx \frac{1}{c^2}(\frac{dr}{dt})^2$. Thus $c^2(e^2-1) \approx O(1)$.)

Newtonian Approximation: In the Newtonian approximation we should get 2π. Let us see this. The turning points are solutions of

$$V_{eff}(r) = \mathcal{E} \implies \frac{1}{c^2}[-\frac{GM}{r} + \frac{l^2}{2r^2}] = \mathcal{E}_{Newt}$$

$$\frac{1}{r^2} - \frac{2GM}{l^2 r} - \frac{2c^2\mathcal{E}}{l^2} = 0$$

The roots $r_{1,2}$ satisfy

$$\frac{1}{r_1} + \frac{1}{r_2} = \frac{2GM}{l^2}$$

If we define $\frac{l}{r} = v$ then $v_1 + v_2 = \frac{2GM}{l}$.

Now the integral in the Newtonian approximation can be written as

$$2\int_{v_1}^{v_2} dv\ \frac{1}{\sqrt{c^2(e^2-1) + \frac{2GM}{l}v - v^2}}$$

This is of the form

$$2\int_{v_1}^{v_2} dv \frac{1}{[(v_1 - v)(v - v_2)]^{\frac{1}{2}}}$$

Letting $v - v_2 = y$ we get

$$\frac{2}{\sqrt{v_1 - v_2}} \int_0^{v_1 - v_2} dy \frac{1}{\sqrt{y(1 - \frac{y}{v_1 - v_2})}}$$

Finally, let $\frac{y}{v_1 - v_2} = x$ we get

$$2 \int_0^1 dx\ x^{-\frac{1}{2}} (1 - x)^{-\frac{1}{2}} = 2 \frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(1)} = 2\pi$$

GR correction: We go back to (61) and evaluate the first correction. Keep in mind that $e^2 \approx 1 + O(1/c^2)$.

$$2 \int_{r_1}^{r_2} \frac{dr\ l}{r^2 [c^2(e^2 - 1) + \frac{2GM}{r} - \frac{l^2}{r^2} + \frac{2GMl^2}{c^2 r^3}]^{\frac{1}{2}}}$$

$$= 2 \int_{r_1}^{r_2} \frac{dr}{r^2} \frac{l}{[c^2 e^2 - [c^2(1 - \frac{2GM}{c^2 r})(1 + \frac{l^2}{c^2 r^2})]]^{\frac{1}{2}}}$$

$$= 2 \int_{r_1}^{r_2} \frac{dr}{r^2} \frac{1}{\sqrt{(1 - \frac{2GM}{c^2 r})}} \frac{l}{[\frac{c^2 e^2}{(1 - \frac{2GM}{c^2 r})} - [c^2(1 + \frac{l^2}{c^2 r^2})]]^{\frac{1}{2}}}$$

Expanding denominators:

$$= 2 \int_{r_1}^{r_2} \frac{dr}{r^2} (1 + \frac{GM}{c^2 r}) \frac{l}{[c^2 e^2 (1 + \frac{2GM}{c^2 r} + (\frac{2GM}{c^2 r})^2) - [c^2(1 + \frac{l^2}{c^2 r^2})]]^{\frac{1}{2}}}$$

We can combine the $1/r^2$ terms and write:

$$= 2 \int_{r_1}^{r_2} \frac{dr}{r^2} (1 + \frac{GM}{c^2 r}) \frac{l}{[c^2 e^2 (1 + \frac{2GM}{c^2 r}) - [c^2(1 + \frac{(\bar{l})^2}{c^2 r^2})]]^{\frac{1}{2}}}$$

with $(\bar{l})^2 = l^2 - (\frac{2GM}{c})^2$.

$$= 2 \int_{r_1}^{r_2} \frac{dr}{r^2} (1 + \frac{GM}{c^2 r}) \frac{l}{[c^2(e^2 - 1) + \frac{2GM}{r} - \frac{(\bar{l})^2}{r^2})]]^{\frac{1}{2}}}$$

This separates into two terms:

$$= 2 \int_{r_1}^{r_2} \frac{dr}{r^2} \frac{l}{[c^2(e^2 - 1) + \frac{2GM}{r} - \frac{(\bar{l})^2}{r^2})]]^{\frac{1}{2}}} + 2 \int_{r_1}^{r_2} \frac{dr}{r^2} \frac{GM}{c^2 r} \frac{l}{[c^2(e^2 - 1) + \frac{2GM}{r}) - \frac{l^2}{r^2})]]^{\frac{1}{2}}}$$

Note that in the second integral we have replaced $\bar{l}$ by l to this order. The first integral is the Newtonian integral encountered earlier. Except that the numerator has l instead of $\bar{l}$. So we write it as $\bar{l}$ times a factor $\frac{l}{\bar{l}} \approx 1 + 2(\frac{GM}{lc})^2$ and thus gives $2\pi(1 + 2(\frac{GM}{lc})^2)$.

The second integral can be done by the same change of variables: $u = \frac{1}{r},\ v = \bar{l}u = (v_1 - v_2)x + v_2$.

$$-(\frac{GM}{lc^2}) \int_{v_1}^{v_2} dv\ \frac{2v}{[(v_1 - v)(v - v_2)]^{\frac{1}{2}}} = 2(\frac{GM}{lc^2}) \int_0^1 dx\ \frac{(v_1 x + v_2(1 - x))}{[x^{\frac{1}{2}}(1 - x)^{\frac{1}{2}}]}$$

$$= (\frac{GM}{lc^2})(v_1 + v_2)\pi$$

Using $v_1 + v_2 = \frac{2GM}{l}$ as before (again $\bar{l}$ is replaced by l) we get for the second integral

$$= 2\pi(\frac{GM}{lc})^2$$

Thus the final answer is

$$\Delta\phi = 2\pi(1 + 2(\frac{GM}{lc})^2) + (\frac{GM}{lc})^2 2\pi = 2\pi + (\frac{GM}{lc})^2 6\pi$$

Thus the precession of the orbit is

$$\delta\phi = 6\pi(\frac{GM}{lc})^2$$

6.3 Orbits of Photons

11. **Orbits of Light rays:**

$$e = -\tilde{\xi}.\tilde{U} = (1 - \frac{2M}{r})\frac{dt}{d\lambda} \quad ; \quad l = \tilde{\eta}.\tilde{U} = r^2 sin^2\ \theta\frac{d\phi}{d\lambda} \tag{62}$$

are conserved quantities. λ here is an affine parameter. Also

$$\tilde{U}.\tilde{U} = g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = 0$$

is fixed. Let us choose $\theta = \frac{\pi}{2}$ -equatorial plane, without any loss of generality, because of the spherical symmetry. Then this equation becomes

$$-(\frac{dt}{d\lambda})^2(1 - \frac{2M}{r}) + (\frac{dr}{d\lambda})^2\frac{1}{(1 - \frac{2M}{r})} + r^2(\frac{d\phi}{d\lambda}) = 0$$

Using (85) this becomes

$$-\frac{e^2}{(1 - \frac{2M}{r})} + (\frac{dr}{d\lambda})^2\frac{1}{(1 - \frac{2M}{r})} + \frac{l^2}{r^2} = 0$$

$$\frac{1}{l^2}(\frac{dr}{d\lambda})^2 + \frac{1}{r^2}(1 - \frac{2M}{r}) = \frac{e^2}{l^2} \equiv \frac{1}{b^2} \tag{63}$$

Note that the Newtonian $\frac{GM}{r}$ term is missing because the photon is massless.

The affine parameter λ is not observable. Hence if one scales $\lambda \to \frac{\lambda}{2}$ the equation should be invariant. In particular under this scaling $e \to 2e$ and $l \to 2l$. But the trajectory cannot be affected. One can check that the equation is indeed invariant under this.

significance of b: b is a parameter with dimensions of length. If the perpendicular distance of the photon trajectory from the star is d, then $\phi \approx \frac{d}{r}$.

$$\frac{d\phi}{dr} = -\frac{d}{r^2} \implies \frac{d\phi}{dt} = \frac{d\phi}{dr}\underbrace{\frac{dr}{dt}}_{\approx -1} = \frac{d\phi}{dr} = \frac{d}{r^2}$$

$$b = \frac{l}{e}|_{r\to\infty} \approx r^2\frac{\frac{d\phi}{d\lambda}}{\frac{dt}{d\lambda}} = r^2\frac{d\phi}{dt} \implies \frac{d\phi}{dt} = \frac{b}{r^2}$$

Thus we conclude that $b = d$ and gives the impact parameter.

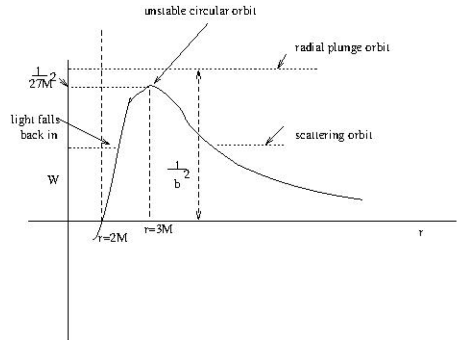


Figure 16: Effective potential and various types of photon orbits - they depend on b^2

(63) can be written as

$$\frac{1}{l^2}(\frac{dr}{d\lambda})^2 + W_{eff}(r) = \frac{1}{b^2}$$

where

$$W_{eff}(r) = \frac{1}{r^2}(1 - \frac{2M}{r})$$

is an "effective" potential.

It's extremum is at

$$\frac{dW}{dr} = 0 \implies r = 3M$$

and

$$W_{eff}(3M) = \frac{1}{27M^2}$$

Thus the allowed region is $\frac{1}{b^2} > W_{eff}$. **See figure 16.**

If $\frac{1}{b^2}$ is large we need large $\frac{e^2}{l^2}$ so small l. Photons with low ang mom penetrate further into the center. We will work in the equatorial plane $\theta = \frac{\pi}{2}$.

Thus if $\frac{1}{b^2} < Max(W_{eff}) = \frac{1}{27M^2}$ the light ray comes in from infinity, reaches a closest point and then goes off to infinity. The closest point is when $W_{eff}(r) = \frac{1}{b^2}$.

If $\frac{1}{b^2} > Max(W_{eff}) = \frac{1}{27M^2}$ light ray plunges into the center.

If $\frac{1}{b^2} = Max(W_{eff}) = \frac{1}{27M^2}$ then there is circular orbit possible at $r = 3M$. But it is unstable since this is a maximium of the potential.

At a given radius r one can calculate the direction of the light ray as a function of b. Let $\tilde{U}$ be 4-mom of light. Let

$$\frac{\tilde{U}.\hat{e}_\phi}{\tilde{U}.\hat{e}_r} = tan\ \psi$$

Then ψ is the angle from the radial as measured by a stationary observer at r. We have derived the orthonormal frame for this observer (57):

$$\hat{e}_r = (0, \sqrt{1 - \frac{2M}{r}}, 0, 0); \quad \hat{e}_t = (\frac{1}{\sqrt{1 - \frac{2M}{r}}}, 0, 0, 0); \quad \hat{e}_\phi = (0, 0, 0, \frac{1}{rsin\ \theta}); \hat{e}_\theta = (0, 0, \frac{1}{r}, 0)$$

$$U^\phi = \frac{d\phi}{d\lambda}; \quad l = \tilde{U}.\tilde{\eta} = g_{\phi\phi}U^\phi \implies U^\phi = \frac{l}{r^2}$$

$$\tilde{U}.\hat{e}_\phi = \frac{l}{r^2}\frac{1}{r}g_{\phi\phi} = \frac{l}{r}$$

$$(\frac{dr}{d\lambda})^2 = [\frac{1}{b^2} - W_{eff}]l^2 \implies \frac{dr}{d\lambda} = U^r = l\sqrt{[\frac{1}{b^2} - W_{eff}]} = l\sqrt{[\frac{1}{b^2} - \frac{1}{r^2}(1-\frac{2M}{r})]}$$

$$\therefore \tilde{U}.\hat{e}_r = U^r\sqrt{1-\frac{2M}{r}}g_{rr} = \frac{l}{\sqrt{1-\frac{2M}{r}}}\sqrt{[\frac{1}{b^2} - \frac{1}{r^2}(1-\frac{2M}{r})]}$$

$$\therefore tan\ \psi = \frac{\frac{l}{r}}{\frac{l}{\sqrt{1-\frac{2M}{r}}}\sqrt{[\frac{1}{b^2} - \frac{1}{r^2}(1-\frac{2M}{r})]}} = \frac{1}{r}\frac{\sqrt{1-\frac{2M}{r}}}{\sqrt{[\frac{1}{b^2} - \frac{1}{r^2}(1-\frac{2M}{r})]}} \tag{64}$$

If $\psi = 0$ the light ray is radial.

If $\psi = 90°$ the light ray is tangential.

Given b, at any r we can calculate ψ. One can ask what is the angle required at any radius for light to escape? This will be obtained if $\frac{1}{b^2} = \frac{1}{27M^2}$. Let ψ_{crit} be this critical angle - this will depend on r. Thus

$$tan\ \psi_{crit} = \frac{1}{r}\frac{\sqrt{1-\frac{2M}{r}}}{\sqrt{[\frac{1}{27M^2} - \frac{1}{r^2}(1-\frac{2M}{r})]}} \tag{65}$$

So at any given r, whenever $\psi < \psi_{crit}$ light will escape. One changes ψ by changing b. This is a useful quantity to calculate: A stationary observer seeing a photon go by him ca measure the angle and decide whether this photon is going to escape or not. It is an easy measurement and determines b. (How else would you determine b? Not easy.)

Let us calculate the critical angle at a radius $r = 3M$. At $r = 3M$ the denominator of (65) vanishes and $\psi = 90°$. So any change in ψ (by decreasing b) will increase the radial component and light will escape.

At $r = 2M$, on the other hand, $\psi_{crit} = 0$. Thus we cannot decrease ψ any further, so light cannot escape. Regardless of what b is i.e. making $\frac{1}{b^2} > \frac{1}{27M^2}$ will not help. So light cannot escape from $r = 2M$. This is the "horizon" in this space-time.

12. **Bending of light:**

See Figure 14

Same technique as was used for precession. Calculate $\frac{d\phi}{dr}$.

$$\frac{d\phi}{d\lambda} = \frac{l}{r^2}; \qquad \frac{dr}{d\lambda} = l\sqrt{[\frac{1}{b^2} - \frac{1}{r^2}(1-\frac{2M}{r})]}$$

$$\frac{d\phi}{dr} = \frac{1}{r^2}\frac{1}{\sqrt{[\frac{1}{b^2} - \frac{1}{r^2}(1-\frac{2M}{r})]}}$$

$$\Delta\phi = 2\int\frac{dr}{r^2}\frac{1}{\sqrt{[\frac{1}{b^2} - \frac{1}{r^2}(1-\frac{2M}{r})]}}$$

Let $r = \frac{b}{w}$. $dw = -\frac{bdr}{r^2}$. Thus

$$\Delta\phi = 2\int_0^{w_1} dw\ \frac{1}{\sqrt{1-w^2(1-\frac{2Mw}{b})}}$$

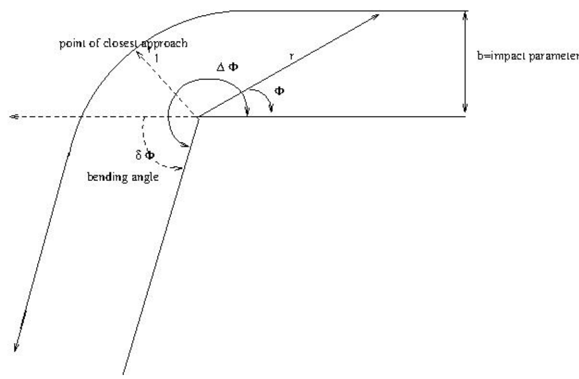


Figure 17: Bending of light angle $\delta\phi = \Delta\phi - \pi$

The limits correspond to $r = \infty$ and $r = r_1$ the closest point of approach. This is defined by the vanishing of the denominator. (Note that when there is no mass to deflect light, $\Delta\phi = \pi$.)

$$= 2\int_0^{w_1} dw \, \frac{1}{\sqrt{1 - \frac{2Mw}{b})}} \frac{1}{\sqrt{\frac{1}{1-\frac{2Mw}{b}} - w^2}} \approx 2\int_0^{w_1} dw \, \frac{1 + \frac{Mw}{b}}{\sqrt{1 + \frac{2Mw}{b} - w^2}}$$

This is the same kind of integral that we encountered in precession calculation. Answer: $\pi + \frac{4M}{b}$. Thus

$$\delta\phi = \Delta\phi - \pi = \frac{4M}{b} = \frac{4GM}{c^2 b}$$

Note that $\frac{GM}{c^2}$ is a length and so is b. If we take for b the radius of the sun and M its mass we get $\delta\phi = 1.7''$(arc second).

7 Gravitational Collapse, Black Holes, Thermodynamics

7.1 Gravitational Collapse

1. **End point of a star**

   A star/planet tries to collapse under its own gravity. What prevents it from collapse is the stiffness of matter - similar to why we don't fall thru the floor. So there is an outward pressure that is able to balance the inward gravitational force. For a small planet the ordinary elastic forces are sufficient. For a large enough star this is not what happens. Most of the star is hydrogen and helium anyway. If you think of a star as hot and gaseous it is the gas pressure - due to the kinetic energy. The heat energy comes from the thermo-nuclear reactions. What happens when the nuclear fuel is exhausted? It starts to collapse. When pressure increases all elements break up into a bunch of protons and electrons and neutrons (electrons and protons can combine reversibly to form neutrons.) The biggest contribution to pressure comes from the electrons because they are the lightest.

   There is no thermal energy, so the only source of pressure is the electron degeneracy (due to Pauli exclusion principle). As you squeeze electrons into a smaller volume, they spread out in momentum space, so their energy goes up. Standard calculation:

   No. of states in a volume element in momentum space is given by $2\frac{d^3p}{(2\pi\hbar)^3}V$ where V is the space volume. Factor of two is for electron spin. Number of particles is thus related to Fermi energy by

   $$N = \frac{4\pi}{3}p_F^3 \times 2 \times V \approx p_F^3 V \implies p_F \approx (\frac{N}{V})^{\frac{1}{3}} \approx \frac{N^{\frac{1}{3}}}{R}$$

   Energy is then given by an integral

   $$E = V8\pi \int^{p_F} p^2 dp \times p\ (Relativistic)\ or \frac{p^2}{2m}\ (Non\ relativistic)$$

   $$\approx p_F^4 V \approx N^{\frac{4}{3}} V^{-\frac{1}{3}}\ relativistic$$

   $$\approx p_F^5 V \approx N^{\frac{5}{3}} V^{-\frac{2}{3}}\ non-relativistic$$

   One can estimate the pressure in the two cases using $dE = PdV$:

   $$P_{Rel} = \frac{dE}{dV} \approx N^{\frac{4}{3}}(-\frac{1}{3})V^{-\frac{4}{3}} \implies \frac{d\ ln\ P_{Rel}}{d\ ln\ V} = -\frac{4}{3}$$

   $$P_{Non-Rel} = \frac{dE}{dV} \approx N^{\frac{5}{3}}(-\frac{2}{3})V^{-\frac{5}{3}} \implies \frac{d\ ln\ P_{Non-Rel}}{d\ ln\ V} = -\frac{5}{3}$$

   The logarithmic derivative gives us the fractional change (increase) in pressure for a fractional change (decrease) in volume. This is the stiffness. It is clear that the non relativistic electron gas is stiffer than the relativistic one.

   Thus as we squeeze the electron gas, the Fermi energy keeps increasing. At some point it starts becoming relativistic, and suddenly the outward pressure drops and it is unable to resist the gravitational pull and collapses.

   One can also see this in terms of energy: The kinetic energy of the electron is

   $$E_{Rel} \approx p_F^3 V \times p_F \approx N p_F c \approx \frac{N^{\frac{4}{3}}\hbar c}{R}$$

   We have inserted factors of $\hbar, c$ since we will use this expression to make a numerical estimate.

   $$E_{Non-Rel} \approx p_F^3 V \times \frac{p_F^2}{2m_{electron}} \approx N p_F^2 \approx \frac{N^{\frac{5}{3}}}{R^2}$$

The total energy in the non relativistic case is

$$E_{Tot} \approx \frac{N^{\frac{5}{3}}}{R^2} - \frac{GM^2}{R}$$

It is clear that there is a minimum energy at some finite R. Here M is the total mass of the star and is mainly protons. In the relativistic case

$$E_{Tot} \approx \frac{N^{\frac{4}{3}}\hbar c}{R} - \frac{GM^2}{R}$$

Here there is no minimum. If the potential dominates at some R then energy will be minimised at $R = 0$. The star collapses! This happens at some critical value, i.e. when $N_{crit}^{\frac{4}{3}}\hbar c \approx GM^2$. Note also that $M = Nm_{proton}$ where we assume that the star is made of (equal numbers of) electrons and protons. If one puts in numbers one finds that $N_{crit} \approx 10^{57}$. This gives for the critical mass $M_{crit} = m_{proton}N_{crit} \approx M_{sun}$. Exact value is $1.4M_{sun}$ and is called the Chandrasekhar mass.

In practices it collapses till the electrons and protons combine into neutrons. Then the collapse can stop under neutron degeneracy pressure. The neutron becomes relativistic at a much higher kinetic energy so it can resist collapse longer. So it is also much smaller - because p_F has to be about a 1000 times greater to be relativistic. So it is about a 1000 time smaller. And it requires greater gravitational pull to over come that - so the star is denser.

The star that depends on electron degeneracy is called a white dwarf. A solar mass white dwarf star ($M_{star} = 1.5km$) has a radius of about a few thousand kilometers. The corresponding "neutron star" has a radius of a few 10- kilometers. In both cases the radius is larger than the Schwarzschild radius.

If the neutron star cannot resist collapse, then in the absence of heavier fermions, we conclude that it just collapses all the way down into a "black hole".

Typically any (non rotating) star that has mass **twice the solar mass or greater** will end up as a black hole.

2. As long as we have spherical symmetry the metric is Schwarzschild - whether star or black hole.

   It is a theorem **Birkhoff** that the only spherically symmetric solution of $R_{\mu\nu} = 0$ (Einstein's eqn in vacuum) is the Schwarzschild solution. The matter can even be time dependent. As long as there is spherical symmetry, this is the metric. Intuitively there is no scalar (monopole) radiation because graviton is spin 2. Spherical symmetry means there is only a monopole radition possible.

   Thus before, after and during collapse it is the same metric. The region of validity of the metric increases. After it forms a black hole we have to worry about the metric even for $r \leq 2M$. At $r = 2M$ we have a coordinate singularity - some metric coefficients blow up. So now we need to understand this carefully. The way to deal with this coordinate singularity is (obviously) to introduce a non singular coordinate system.

   The singularity at $r = 0$ is genuine and cannot be gotten rid of. Curvature tensors ($R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma}(r = 0) = \infty$) are infinite. This is an unsolved problem in quantum gravity.

3. **Astrophysical Black Holes**

   There are three ways in which black holes might form:

   1) Have been detected in X-Ray binaries - Collapse of massive stars - supernova explosions - remnant is often a black hole. Typically a few solar masses. Matter forms an accretion disk and radiates as they lose angular momentum and fall into the black hole. The last stage of fall may be rapid and X rays can be emitted.

   2) Galaxy centres - Galaxies may undergo gravitational collapse in the centre - supermassive black holes. (The mass goes as R^3. So Schwarzschild radius increases as R^3. So it will soon become larger than R.)

3) Primordial black holes in the early universe. Density fluctuations could have produced small black holes (eg 10^{-19} solar masses).

4. **Eddington-Finkelstein Coordinates**

A new coordinate is motivated by the following considerations:

Consider a radial light ray. It follows a trajectory defined by $ds^2 = 0$. We follow it through $r = 2M$. If we can solve for the trajectory we expect that nothing singular is happening at $r = 2M$.

$$dt^2(1 - \frac{2M}{r}) = \frac{dr^2}{(1 - \frac{2M}{r})} \implies \int dt \ = \pm \int \frac{dr}{(1 - \frac{2M}{r})} \equiv \pm \int dr^*$$

r^* is often called the "tortoise coordinate". Doing the integral gives

$$t + const = \pm(r + 2Mln\ (r - 2M))$$

We choose a constant to make the argument of the log dimensionless:

$$t = \pm((r + 2Mln\ (\frac{r}{2M} - 1))$$

To make it well defined on either side of $r = 2M$, r^* is more precisely defined as

$$r^* = r + 2Mln\ |\frac{r}{2M} - 1|$$

Radial ingoing light rays obey

$$t + r + 2Mln\ |\frac{r}{2M} - 1| \equiv v = constant$$

The easiest way to see this is to use the coordinate v: by substituting

$$t = v - r^* = t - r - 2Mln\ |\frac{r}{2M} - 1|$$

Write $dt = dv - dr^*$. So the Schwarzschild metric becomes:

$$ds^2 = -(1 - \frac{2M}{r})(dv - dr^*)^2 + (1 - \frac{2M}{r})(dr^*)^2 + r^2 d\Omega_2^2$$

$$\implies ds^2 = -(1 - \frac{2M}{r})dv^2 + 2dvdr + r^2 d\Omega_2^2$$

Important point: Thus using these coordinates we can follow the light ray through $r = 2M$ and nothing singular happens. There is no singularity at $r = 2M$. This resolves the "coordinate singularity" - something that can be attributed to a bad choice of coordinates. The new non singular coordinates are called **"Ingoing Eddington-Finkelstein coordinates"**.

This method - of following null geodesics to define new coordinates - is a standard one.

It is also convenient to define a modified time coordinate

$$\tilde{t} = v - r = t + 2Mln\ |\frac{r}{2M} - 1|$$

5. **Light Cones**

We now study light cones in this new coordinate system. **See Figure**

They are given by $ds^2 = 0$ or

$$(1 - \frac{2M}{r})dv^2 = 2dvdr$$

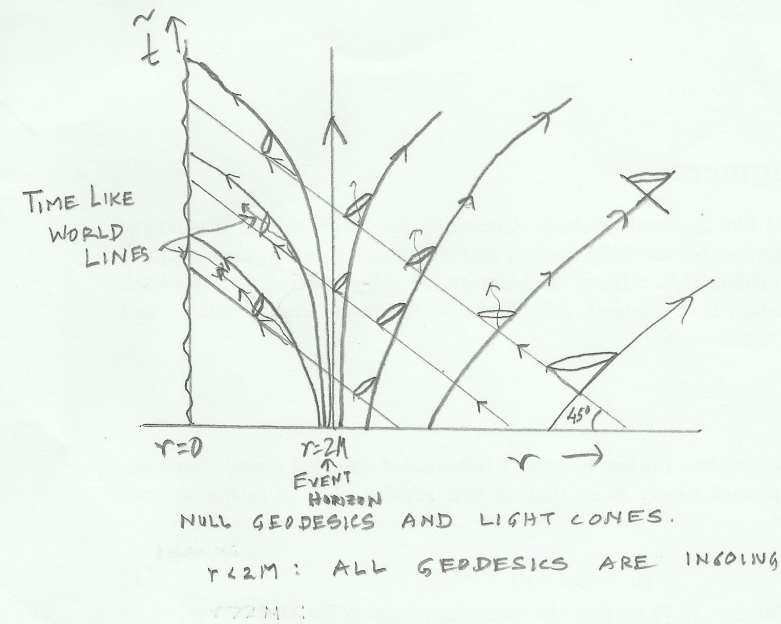


Figure 18: Light Cones in Schwarzschild geometry

There are several possibilities:

a) $dv = 0 \implies v = const \implies v = t + r + 2Mln\ |\frac{r}{2M} - 1| = t + r^* = const$

This describes ingoing light rays. This is particularly easy to see in $\tilde{t}$ coordinates since the lines are given by $\tilde{t} + r = const$. $\frac{d\tilde{t}}{dr} = -1$ everywhere. There is no complication at $r = 2M$ for this coordinate.

b)$(1 - \frac{2M}{r})dv = 2dr \implies v - 2r^* = 0 \implies t - r^* = 0$. This is clearly outgoing for $r > 2M$. One can calculate

$$\frac{d\tilde{t}}{dr} = 1 + \frac{4M}{r - 2M} \quad \to 1\ at\ r = \infty; \quad \to \infty \quad at\ r = 2M; \quad \to -1 \quad at\ r = 0$$

c) $r = 2M$ and $dr = 0$. This means light stays at the same radius $r = 2M$.

6. **Geometry of Horizon and Singularity**

   (a) Horizon is a null surface spanned by θ, ϕ, t i.e. $S^2 \times R^1$. It is null because $\frac{\vec{\partial}}{\partial t}.\frac{\vec{\partial}}{\partial t} = -(1 - \frac{2M}{r}) = 0$ at the horizon.

   (b) Light cones (and null surfaces in general) have a one way property. Once you cross it, you can't cross it again. Otherwise it would be like overtaking a light ray!

   The horizon is a null surface generated by light rays that stay at the horizon. See part c) above. Area of the horizon is $4\pi(2M)^2 = 16\pi M^2$.

   (c) $r = 0$ is a curvature singularity. But it is a moment in time - not a point in space! Because $\frac{\vec{\partial}}{\partial r}.\frac{\vec{\partial}}{\partial r} = \frac{1}{(1-\frac{2M}{r})} < 0$ when $r < 2M$. So inside the horizon, going towards the centre is like going forward in time. It is a space-like singularity.

   (d) Although the horizon is a coordinate singularity interesting things do happen. One such was mentioned above: For $r < 2M$ the radial direction becomes time like, so objects after crossing the horizon can only go towards smaller radius, otherwise it would be like going backwards in time.

   As you approach the horizon it becomes harder and harder to resist the gravitational pull. Estimate the acceleration of a stationary observer. His 4-velocity is $U_{obs} = (\frac{1}{\sqrt{1-\frac{2M}{r}}}, 0, 0, 0)$ in the Schwarzschild coordinate basis.

   The equation for geodesic motion can be interpreted as setting the 4-acceleration to zero. In flat space this is just $\frac{d^2x^\alpha}{d\tau^2} = 0$. In curved space it was calculated to be

$$\frac{d^2x^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = 0$$

   Thus if it is not zero we can use this to calculate the four acceleration:

$$a^\alpha = \frac{dU^\alpha}{d\tau} + \Gamma^\alpha_{\beta\gamma}U^\beta U^\gamma$$

   Since U_{obs} is fixed, the time derivative is zero. The second term is thus $\Gamma^\alpha_{00}U^0_{obs}U^0_{obs}$. $\Gamma^r_{00} = \frac{M}{r^2}(1 - \frac{2M}{r})$. Thus we get for the radial acceleration:

$$\vec{a} = \frac{M}{r^2}\frac{\vec{\partial}}{\partial r}$$

   This is in the Schwarzschild basis. To see what the observer himself feels one has to go to his orthonormal basis where $\hat{e}_r = \sqrt{1 - \frac{2M}{r}}\frac{\vec{\partial}}{\partial r}$. We get

$$\vec{a} = \frac{M}{r^2}\frac{1}{\sqrt{1 - \frac{2M}{r}}}\hat{e}_r$$

   This blows up as $r \to 2M$. In other words he needs a very powerful rocket to hover around the horizon.

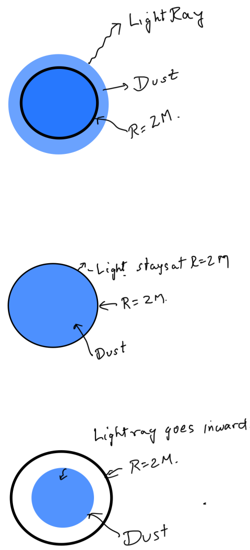


Figure 19: As the outer dust layer crosses $R = 2M$ the black hole horizon can be said to have formed. Light rays emitted by the dust particles cannot escape.

7. **Collapse**

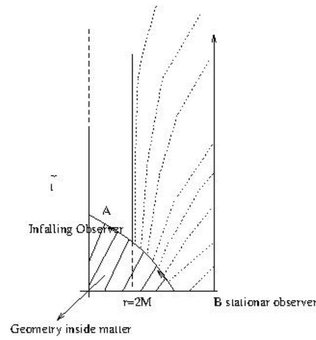
   Given the results we have till now, it is possible to study collapse into a black hole of "pressure less dust". Consider a swarm of non interacting particles. There is no pressure then. This is called "dust". We can consider a spherical shell falling into a black hole and going past the horizon and raising the mass of the black hole - the mass as observed from infinity. This is the process of collapse. See Figure 19 and 20.

   Since there is no pressure, each particle follows a radial geodesic. This is described by (55),(56).

$$r^{\frac{3}{2}} = \frac{3}{2}\sqrt{2M}(\tau_* - \tau)$$

$$t = t^* + 2M[-\frac{2}{3}(\frac{r}{2M})^{\frac{3}{2}} - 2(\frac{r}{2M})^{\frac{1}{2}} + ln\,|\frac{(\frac{r}{2M})^{\frac{1}{2}} + 1}{(\frac{r}{2M})^{\frac{1}{2}} - 1}|]$$

   The first equation gives the relation between r and τ and the second one between r and t. It is clear that the proper time taken for a particle to reach the horizon is finite, whereas the coordinate time is infinite - diverges logarithmically: $t \to \infty$ as $r \to 2M$.



Collapsing shell forms a black hole.

Figure 20: Collapsing matter forms a black hole. Signals emitted by an infalling observer, A, and received by an outside stationary observer, B. Vertical axis is $\tilde{t}$.

Let us study the other time coordinate that we have come across: $\tilde{t}$.

$$t = v - r - 2Mln\ |\frac{r}{2M} - 1|$$

$$\tilde{t} = v - r = t + 2Mln\ |\frac{r}{2M} - 1|$$

$$= t^* + 2M[-\frac{2}{3}(\frac{r}{2M})^{\frac{3}{2}} - 2(\frac{r}{2M})^{\frac{1}{2}} + ln\ |\frac{(\frac{r}{2M})^{\frac{1}{2}} + 1}{(\frac{r}{2M})^{\frac{1}{2}} - 1}|] + 2Mln\ |\frac{r}{2M} - 1|$$

$$\implies \tilde{t} = t^* + 2M[-\frac{2}{3}(\frac{r}{2M})^{\frac{3}{2}} - 2(\frac{r}{2M})^{\frac{1}{2}} + 2ln\ |(\frac{r}{2M})^{\frac{1}{2}} + 1|]$$

Let us define $\tilde{t} = 0$ at $r = 0$. Then t^* in the above expression is zero. Thus

$$\tilde{t} = 2M[-\frac{2}{3}(\frac{r}{2M})^{\frac{3}{2}} - 2(\frac{r}{2M})^{\frac{1}{2}} + 2ln\ |(\frac{r}{2M})^{\frac{1}{2}} + 1|]$$

In particular at $r = 2M$ it is finite. Also the time time it takes to reach $r = 2M$ from and finite $r > 2M$ is also finite. Finally the proper time the dust particle takes to reach $r = 0$ is given by (55).

$$(2M)^{3/2} = 3/2(2M)^{1/2}(\tau^* - \tau_1)$$

$\tau = \tau^*$ is when it reaches $r = 0$. Let $\tau = \tau_1$ be the time it reaches $r = 2M$. Then $\tau^* - \tau_1$ is the time for transit from $r = 2M$ to $r = 0$.

$$\tau^* - \tau_1 = \frac{4M}{3}$$

This can be taken as the time for the formation of the black hole. For a solar mass BH it is about 10^{-5} seconds.

8. **Redshift of emitted light - What do observers see?**

   The infalling observer falls into the centre of the BH in a finite proper time - which is the time of his clock. What about outside observers? Imagine the infalling observer sends out pulses of light at regular intervals - regular according to his proper time. What does the outside observer, stationary and far away, see?

   The outgoing light ray follows null geodesics:

$$v - 2r^* = v - 2(r + 2Mln\ |\frac{r}{2M} - 1|) = const$$

Let v_E, r_E be the coordinates of emission and v_R, r_R be the coordinates of the received signal. Then the light ray obeys

$$v_E - 2(r_E + 2Mln\,|\frac{r_E}{2M} - 1|) = v_R - 2(r_R + 2Mln\,|\frac{r_R}{2M} - 1|)$$

As $r_E \to 2M$ the logarithm dominates the LHS, which becomes $-4Mln\,|\frac{r_E}{2M} - 1|)$. For the RHS, if $r_R >> 2M$ we can neglect the logarithm and use $v_R = t_R + r_R + 2Mln\,|\frac{r_R}{2M} - 1| \approx t_R + r_R$. Thus RHS is $\approx v_R - 2r_R = t_R - r_R$. So we get

$$-4Mln\,|\frac{r_E}{2M} - 1|) = t_R - r_R$$

$$\implies \frac{r_E}{2M} - 1 = e^{-\frac{(t_R - r_R)}{4M}} \tag{66}$$

So as $r_E \to 2M$, $t_R \to \infty$. The observer at infinity almost never sees the last pulses of the infalling observer - he has to wait for an infinite amount of time. Equivalently, if the infalling observer moves to a new position $r_E + \Delta r_E$ after time $\Delta\tau$,

$$\frac{\Delta r_E}{2M} = -\frac{\Delta t_E}{4M} e^{-\frac{(t_R - r_R)}{4M}}$$

Thus for finite $\Delta\tau$, Δr_E is finite, but Δt_E clearly goes to ∞ as $r_E \to 2M$ since $e^{-\frac{(t_R - r_R)}{4M}} \to 0$ (by (66)). The frequency of the pulses decreases for the receiver. This can also be described as a manifestation of "redsfift". The redshift is infinite for rays emanating from the horizon. The time scales ($4M$) for a solar mass star is 10^{-5} sec (this is the same as 1.5 km). So in that much time the signals get hugely redshifted and for all practical purposes the observer sees a "black" hole. But in principle it takes an infinite amount of time to form a black hole from the point of view of the observer outside.

7.2 Black hole thermodynamics

9. What happens if you drop hot matter into a black hole? Does it's entropy disappear from the universe? Violates 2nd law of thermodynamics?

   Black hole area always increases. This sounds like a second law of thermodynamics. Beckenstein said maybe area is entropy! But if it has entropy $dE = TdS$ so it should be at a finite temperature. Hawking said, in that case it should radiate like a black body. All these seem to be true!

10. In units where $\hbar = c = 1$, Entropy = Area/ 4G (dimensionless: $\frac{G\hbar}{c^3} = l_P^2$). For a Schwarzschild BH ($c = 1$):

$$Area = 4\pi R_s^2 = 4\pi(2GM)^2 = 16\pi G^2 M^2 \implies S = 4\pi GM^2 = 4\pi M^2 \quad (geometric\ units)$$

11. Since $E = M,\ dE = dM$, we get

$$T\ dS = dE \implies T 8\pi M = 1 \implies T = \frac{1}{8\pi M}$$

   Putting back $\hbar, k_B, c, G$:

$$k_B T = \frac{\hbar c^3}{8\pi GM}$$

   So one should be able to associate a temperature with the BH! Black hole should then radiate like a black body. Hawking found that this is indeed true and he calculated the temp and found agreement with the thermodynamics above. In physical units

$$T = 6.2 \times 10^{-8} (\frac{M_\circ}{M}) K$$

12. **Hawking radiation:** So far everything has been classical. But in QM there are fluctuations and it is possible for an electron and positron to be virtually produced - i.e. for a short time during which they are not in energy momentum eigenstates. If they were energy conservation would rule out two positive energy particles being produced. So these are virtual and go back to the vacuum very soon.

    But suppose the elctron were produced outside the horizon and the positron inside. This is possible because in QM there is an uncertainty in the exact location - say, $O(\frac{\hbar}{m})$ - the Compton wave length. Then the conservation law is $\vec{\xi}.p_{el} + \vec{\xi}.p_{pos} = 0$. So if $\vec{\xi}.p_{el} > 0$ then $\vec{\xi}.p_{pos} < 0$. Normally this means that the positron with negative energy, cannot be real. However inside the horizon one can have a *real* positron, because $\vec{\xi}$ is space-like! So $\vec{\xi}.p_{pos}$ is like a *three momentum* - which can be of either sign and still be real. So in fact the process does take place -and real electrons escape to infinity as radiation. Hawking showed that in the semi classical approximation this radiation is thermal, with precisely the temperature given in the previous paragraph.

13. **BH thermodynamics and geometry:** There is an interesting and possibly deep interplay between geometry and black hole thermodynamics. Consider the Euclidean section of a BH - which means take the Schwarzschild metric and replace it by τ (Analytically continue: $t = -i\tau$ and consider τ real.)

    (Why? Because from Stat mech we know that in QFT, if we repace it by t_E and assume periodic bc for t_E and $t_E + \beta$ i.e. $\phi(t_E) = \pm\phi(t_E + \beta)$, with minus sign for fermions, then we get a description of *equilibrium* statistical mechanics of the system. Since it is in equilibrium we don't need the time coordinate.)

    Then the BH metric becomes:

$$ds^2 = (1 - \frac{2M}{r})dt_E^2 + \frac{dr^2}{(1 - \frac{2M}{r})} + r^2 d\Omega_2^2$$

    Let $r = 2M + \rho$ and assume $\rho << 2M$. Then

$$ds^2 = \frac{\rho}{2M}dt_E^2 + 2M\frac{d\rho^2}{\rho} + r^2 d\Omega_2^2$$

    Bring it into some standard form by replacing $\rho = \frac{y^2}{8M}$.

$$ds^2 = \frac{y^2}{16M^2}dt_E^2 + dy^2 + r^2 d\Omega_2^2$$

    Further replace $\frac{t_E}{4M} = \theta$.

$$ds^2 = dy^2 + y^2 d\theta^2 + r^2 d\Omega_2^2 \tag{67}$$

    This looks like a plane. But is it? Only if θ has a period of 2π. What is it otherwise? A cone:

    Digression on Cones:

    Consider polar coordinates on a plane:

$$ds^2 = dr^2 + r^2 d\theta^2$$

    Suppose we write instead

$$ds^2 = dr^2 + \frac{r^2}{4}d\theta^2$$

    This can also be made to look like a plane by writing $2\theta' = \theta$:

$$ds^2 = dr^2 + r^2 d\theta'^2 \tag{68}$$

    But the range of θ' is $0 - \pi$. Let the space have this periodicity. What does this correspond to? This space is periodic with π. Consider a semi circle - it has angle π at the centre. Join the two boundary radii and it becomes a cone. This has the required periodicity. So the metric (68) describes a cone. The half angle of this cone is 30 degrees. The deficit angle of the cone is π. This space is called a conical singularity. So the space described metric (68) is not a smooth plane but has a **conical singularity**. The periodicity of 2π is crucial if it is to describe a plane. **End of digression.**

So the space described by (67) has a conical singularity unless θ has periodicity 2π. But that means t_E has to have a periodicity of $8\pi M$. Thus this should describe an equilibrium system of inverse temperature $\beta = 8\pi M$. So the temperature is $T = \frac{1}{8\pi M}$. By requiring smoothness of the Euclidean section of the BH, we get its temperature - as obtained by different arguments!

7.3 Coordinate Singularities: Rindler Spacetime

To understand coordinate singularities it is useful to consider Rindler space time - which as we shall see later, is just Minkowski space in disguise.

14. Simple example (Wald):

$$ds^2 = -\frac{1}{t^4}dt^2 + dx^2, \quad -\infty < x < \infty, \quad 0 < t < \infty$$

There is a singularity at $t = 0$. Define $t' = \frac{1}{t}$.

$$ds^2 = -dt'^2 + dx^2$$

$t = 0$ corresponds to $t' = \infty$. Geodesics go all the way to $t' \to \infty, (t \to 0)$ and have arbitrarily large affine parameter. But as $t \to \infty, (t' \to 0)$ the geodesics stop. Add $t' < 0$ - extends t beyond $t = \infty$ and gives Minkowski space. This coord transf we got by guessing.

Geodesics are

$$\frac{dt}{t^2} = \pm dx \implies x = \pm\frac{1}{t}$$

Let $k^\alpha = (\frac{dt}{d\lambda}, \frac{dx}{d\lambda}) = (\frac{dt}{d\lambda}, -\frac{1}{t^2}\frac{dt}{d\lambda})$ be the null vector along the geodesic. $\xi = \partial_x$ is a Killing vector. So

$$\xi.k = E = -\frac{1}{t^2}\frac{dt}{d\lambda}$$

is a conserved quantity. Thus

$$\lambda = \frac{1}{Et} + const$$

So as $t \to 0, \ \lambda \to \infty$. However as $t \to \infty, \ \ \lambda \to 0$ and stops there. So we extend $\lambda < 0$, i.e. beyond $t = \infty$. This is achieved by going to t'.

More complicated is Rindler.

15. **Rindler Metric** is

$$ds^2 = -x^2dt^2 + dx^2 \tag{69}$$

There seems to be coordinate singularity at $x = 0$. If we make the change of variables: $x^2 = 4y, t' = 2t$, the metric becomes

$$ds^2 = -y \ dt'^2 + \frac{dy^2}{y} \tag{70}$$

If we let $r = 2M + \frac{\rho}{2M}, t' = \frac{t}{2M}$ the Schwarzschild metric becomes near the horizon,

$$ds^2 = -\rho \ dt'^2 + \frac{d\rho^2}{\rho} \tag{71}$$

Thus the Schwarzschild metric near the horizon looks a lot like Rindler spacetime metric and the same coordinate singularity.

16. A good way to analyse such space times is to study null geodesics - trajectory of a light ray - and see what happens at the suspected singularity. If it goes through without any problem then the chances are that a coordinate change will get rid of the singularity. The geodesic parameter is in fact a good coordinate in that case.

    If $\tilde{k}$ is the wave vector of the photon then $k^\alpha k_\alpha = 0$ gives an equation for null directions: $k^\alpha = (\frac{dt}{d\lambda}, \frac{dx}{d\lambda})$.

$$k^\alpha k_\alpha = 0 \implies -x^2(\frac{dt}{d\lambda})^2 + (\frac{dx}{d\lambda})^2 = 0$$

    So

$$(\frac{dx}{x}) = \pm dt \implies t = \pm ln\ x + const$$

    It is useful to introduce ingoing (moving in -x direction) and outgoing (moving in +x direction) coordinates:

$$v = t + ln\ x, \quad u = t - ln\ x, \quad t = \frac{1}{2}(u+v), \quad x = e^{\frac{v-u}{2}}$$

    Thus if v is held fixed and u is increasing one is going in the -x direction. Thus ∂_u is an ingoing vector. But the coordinates do not go beyond $x = 0$. x can never be negative.

$$dt = \frac{1}{2}(du + dv), \quad d\ ln\ x = \frac{1}{2}(dv - du), \quad dx = \frac{1}{2}(dv - du)e^{\frac{v-u}{2}}$$

    The Rindler metric becomes

$$ds^2 = -e^{v-u}\ du\ dv \tag{72}$$

    Clearly the range of u, v is $-\infty \le u, v \le \infty$. In these variables, the geodesics either satisfy $du = 0$ or $dv = 0$. The range of x is still $x > 0$. So we need to do some more work: Let us write the geodesic in these variables and find out what the affine parameter is in terms of u, v. The affine parameter should be allowed to range over its full course so that geodesics do not stop abruptly.

    $\xi = \partial_t$ is a Killing vector field. $k.\xi$ is conserved along the geodesic. Let us evaluate it:

$$\vec{\xi} = \vec{\partial}_t = \frac{\partial u}{\partial t}\vec{\partial}_u + \frac{\partial v}{\partial t}\vec{\partial}_v = \vec{\partial}_u + \vec{\partial}_v$$

    The geodesics are either outgoing (along lines of constant u) or ingoing (along lines of constant v).

$$\vec{k} = \frac{dv}{d\lambda}\vec{\partial}_v \quad or \quad \vec{k} = \frac{du}{d\lambda}\vec{\partial}_u$$

    Thus

$$E = -k.\xi = \frac{dv}{d\lambda}g_{uv} = -\frac{1}{2}\frac{dv}{d\lambda}(-e^{v-u}) = \frac{e^{v-u}}{2}\frac{dv}{d\lambda}$$

$$2E\int d\lambda = e^{-u}\int e^v dv \implies \lambda = \frac{1}{2E}e^{v-u} + const$$

    Thus $\lambda_{out} = e^v$ is an affine parameter for outgoing geodesics. Similarly we see that $\lambda_{in} = -e^{-u}$ is an affine parameter along ingoing geodesics. They should be allowed to range from $(-\infty, +\infty)$. So let

$$V = e^v, \quad U = -e^{-u}$$

    be the coordinates and let them range from $(-\infty, +\infty)$. Now we have extended the space time because in terms of u, v this is not possible.

    Metric in terms of U, V is

$$ds^2 = -dU\ dV \tag{73}$$

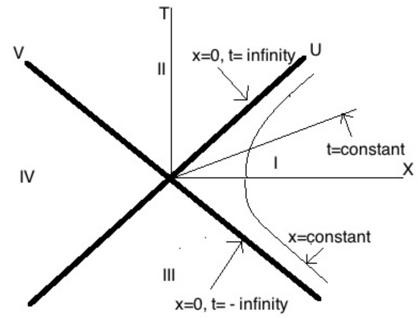


Figure 21: Rindler spacetime extended to Minkowski spacetime

where now we allow $-\infty < U, V < \infty$ thus extending space time. This is Minkowski space! Write $T = \frac{1}{2}(U + V)$ and $X = \frac{1}{2}(V - U)$ to get

$$ds^2 = -dT^2 + dX^2 \tag{74}$$

In terms of the original variables:

$$x = \sqrt{X^2 - T^2}, \quad t = tanh^{-1}(\frac{T}{X})$$

We can use $4y = x^2$ in region II. Here $4y = X^2 - T^2 < 0$. Thus $y = constant$ is a hyperbola that is space like.

$$X = \sqrt{4y}\ cosh\ t, \quad T = \sqrt{4y}\ sinh\ t$$

and in region II

$$X = \sqrt{-4y}\ sinh\ t, \quad T = \sqrt{-4y}\ cosh\ t$$

is a spacelike surface. So t is a spacelike coordinate.

Since $x > 0$ was real (and $x^2 > 0$), we were restricted to $X > |T|$ - Region I. When $x < 0$ and real, we get Region IV. But by changing variables to U (or y) we have effectively allowed ourselves $x^2 < 0$! This gives us the full Minkowski space time.

17. What does this metric correspond to physically? Consider an observer moving in flat Minkowski spacetime with constant acceleration a:

$$X(\tau) = \frac{1}{a} cosh\ (a\tau), \quad T(\tau) = \frac{1}{a} sinh\ (a\tau)$$

a is the magnitude of his 4-acceleration $\tilde{a}.\tilde{a}$. Let us label points in space time by using the world lines of these observers. Thus a point on this trajectory is labeled by a, τ. Let us set $x = \frac{1}{a}$ and $t = a\tau$. So

$$X = x\ cosh\ t, \quad T = x\ sinh\ t, \quad \implies X + T = xe^t, \quad X - T = xe^{-t}$$

$$ds^2 = (dX + dT)(dX - dT) = dx^2 - x^2 dt^2$$

The point $x = 0$ corresponds to infinite 4-acceleration. Similarly since the observer is accelerating, he reaches almost the speed of light soon, and asymptotes to $x = |t|$ and thus rays from any point $x < 0$ can never catch up with him. Hence he sees the null surface $x = |t|$ as a horizon. But it is observer dependent.

18. Consider the Euclidean section:

$$ds^2 = dx^2 + x^2 dt_E^2$$

This looks like a plane, provided t_E has peridicity 2π. $t_E = a\tau_E$. Thus τ_E has to have periodicity $\frac{2\pi}{a}$. So the observer sees a temperature of $\frac{a}{2\pi}$! This effect was demonstrated by Unruh - he showed that accelerating particles see a finite temperature, which is proportional to their acceleration in precisely this way. Further more the back hole temperature can also be understood as an Unruh effect - the surface acceleration we have calculated is $\frac{M}{r^2} = \frac{1}{4M}$. This corresponds to a temperature: $T = \frac{1}{8\pi M}$! (Note that this has the redshift factor $\frac{1}{\sqrt{1-\frac{2M}{r}}}$ taken out. So the actual temperature observed by a stationary observer at the horizon will be infinite - he feels an infinite gravitational attraction. The blackbody Hawking radiation will be redshifted so that the stationary observer at infinity sees a temperature $T = \frac{1}{8\pi M}$.)

7.4 Kruskal-Szekeres extension

19. We have seen (71) that the Schwarzschild metric near the horizon looks like Rindler. So use the same tricks: Null geodesics:

$$(1 - \frac{2M}{r})dt^2 = \frac{dr^2}{(1 - \frac{2M}{r})} \implies \int dt = \int \frac{dr}{(1 - \frac{2M}{r})} = \int dr^*$$

where r^* is the tortoise coordinate introduced earlier. Thus the geodesics satisfy $t = \pm r^* + const$. So as in the Rindler case introduce null coordinates

$$v = t + r^*, \quad u = t - r^*$$

The metric becomes

$$ds^2 = -(1 - \frac{2M}{r})\ du\ dv \tag{75}$$

Here r is understood as a function of u, v defined by the equation

$$r^* = r + 2M\ ln\ (\frac{r}{2M} - 1) = \frac{1}{2}(v - u)$$

Then

$$1 - \frac{2M}{r} = \frac{2M}{r} e^{-\frac{r}{2M}} e^{\frac{v-u}{2}}$$

We can also write

$$ds^2 = -\frac{2Me^{-\frac{r}{2M}}}{r}\ e^{\frac{(v-u)}{4M}}\ du\ dv \tag{76}$$

The r dependent piece is well behaved as $r \to 2M$, which corresponds to $u \to \infty$ or $v \to -\infty$. Now again following Rindler introduce U, V:

$$V = e^{\frac{v}{4M}}, \quad U = -e^{-\frac{u}{4M}}$$

Thus

$$1 - \frac{2M}{r} = \frac{2M}{r} e^{-\frac{r}{2M}} UV \tag{77}$$

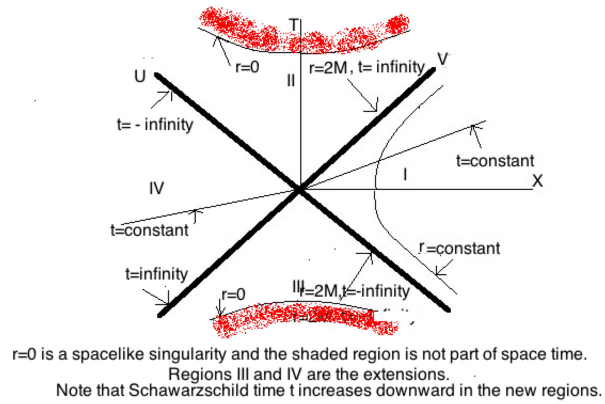


Figure 22: Kruskal Szekeres extension of Schwarzschild spacetime

Now

$$ds^2 = -\frac{32M^3 e^{-\frac{r}{2M}}}{r} \; dU \; dV \tag{78}$$

There are no singuarities in the metric coefficients except at $r = 0$. $r \to 2M$ corresponds to $U \to 0$ or $V \to 0$. But now we can extend U, V to negative values also - again as we did in Rindler. Looking at (77) we see that we can take $r < 2M$ by letting U, V be negative. We can also define

$$V = T + X, \quad U = T - X$$

and we get

$$ds^2 = \frac{32M^3 e^{-\frac{r}{2M}}}{r} \; (-dT^2 + dX^2) + r^2(d\theta^2 + sin^2\theta \; d\phi^2) \tag{79}$$

Region I is $V > 0, U < 0$. If we extend to $V > 0, U > 0$ we go inside the horizon to $r < 2M$ region II. This is all there is in the Schwarzschild coord. But now we also have regions III and IV corresponding to $V < 0$. This is the K-S extension.

Once again r has to be understood as a function of X, T. In any case, there is no trace of any singularity as $r \to 2M$.

$$(\frac{r}{2M} - 1)e^{\frac{r}{2M}} = X^2 - T^2, \quad \frac{t}{2M} = ln \; (\frac{T + X}{X - T}) = 2 \; tanh^{-1}(\frac{T}{X})$$

The condition $r > 0$ still holds. This means

$$X^2 - T^2 > -1$$

$r = 0$ is the space like hyperbola $X^2 - T^2 = -1$.

Important: Below we have used a different notation: X, T above are replaced by U, V below. U, V above are called $\tilde{U}, \tilde{V}$ below.

The following tables make the similarities between Schwarzschild and Rindler metrics clear. For the Rindler metric the coordinates U, V are the original Minkowski space-time coordinates. Region I is the usual Rindler wedge. Region II is the inside of the horizon. Region III, IV are the regions $\tilde{V} < 0$ of Minkowski space time. In the Schwarzschild these are the Kruskal extensions.

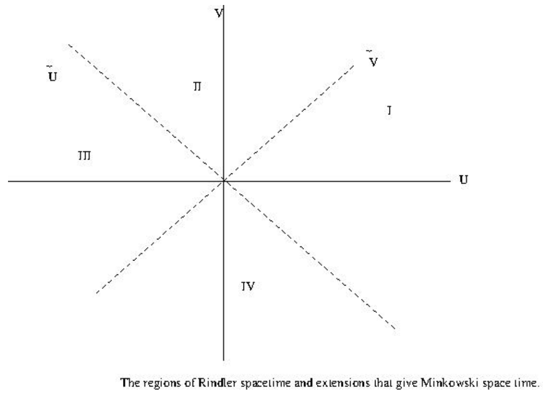


Figure 23: Extended Rindler Space-time in our new notation.

Region I Schwarzschild $r > 2M$	Region I Rindler $\rho > 0$
$ds^2 = -(1-\frac{2M}{r})dt^2 + \frac{dr^2}{(1-\frac{2M}{r})} = -(1-\frac{2M}{r})(dt^2 - dr^{*2})$	$ds^2 = -\rho dt^2 + \frac{d\rho^2}{4\rho} \quad (\rho = x^2 \implies ds^2 = -x^2dt^2 + dx^2)$
$v = t + r^* = t + (r + 2Mln\ \lvert 1-\frac{2M}{r}\rvert); \quad u = t - r^*$	$u = t - ln\ x = t - \frac{1}{2}ln\ \rho; \quad v = t + ln\ x = t + \frac{1}{2}ln\ \rho$
$\tilde{U} = -e^{\frac{-u}{4M}} = -e^{-\frac{t}{4M}}e^{\frac{r}{4M}}(\frac{r}{2M}-1)^{\frac{1}{2}}$	$\tilde{U} = -e^{-u} = -\sqrt{\rho}e^{-t}$
$\tilde{V} = e^{\frac{v}{4M}} = e^{\frac{t}{4M}}e^{\frac{r}{4M}}(\frac{r}{2M}-1)^{\frac{1}{2}}$	$\tilde{V} = e^{v} = \sqrt{\rho}e^{t}$
$\tilde{U} < 0, \tilde{V} > 0$	$\tilde{U} < 0, \tilde{V} > 0$
$\tilde{U} = V - U; \quad \tilde{V} = V + U; \quad \tilde{U}\tilde{V} = -e^{\frac{r}{2M}}(\frac{r}{2M}-1) < 0$	$\tilde{U} = V - U; \quad \tilde{V} = V + U; \quad \tilde{U}\tilde{V} = -\rho < 0$
$U = e^{\frac{r}{4M}}(\frac{r}{2M}-1)^{\frac{1}{2}}cosh(\frac{t}{4M})$	$U = \sqrt{\rho}\ cosh\ t$
$V = e^{\frac{r}{4M}}(\frac{r}{2M}-1)^{\frac{1}{2}}sinh(\frac{t}{4M})$	$V = \sqrt{\rho}\ sinh\ t$
$ds^2 = -\frac{32M^3}{r}e^{-\frac{r}{2M}}d\tilde{U}d\tilde{V}$	$ds^2 = -\rho dudv = -d\tilde{U}d\tilde{V}$
$ds^2 = \frac{32M^3}{r}e^{-\frac{r}{2M}}(dU^2 - dV^2)$	$ds^2 = dU^2 - dV^2$

Region II Schwarzschild $r < 2M$	Region II Rindler $\rho < 0$
$ds^2 = -(1-\frac{2M}{r})dt^2 + \frac{dr^2}{(1-\frac{2M}{r})} = -(1-\frac{2M}{r})(dt^2 - dr^{*2})$	$ds^2 = -\rho dt^2 + \frac{d\rho^2}{4\rho} \quad (\rho = -x^2) \implies ds^2 = x^2dt^2 - dx^2)$
$v = t + r^* = t + (r + 2Mln\ \lvert 1-\frac{2M}{r}\rvert); \quad u = r^* - t$	$u = ln\ x - t = \frac{1}{2}ln\ \rho - t; \quad v = t + ln\ x = t + \frac{1}{2}ln\ \rho$
$\tilde{U} = e^{\frac{u}{4M}} = e^{-\frac{t}{4M}}e^{\frac{r}{4M}}(1-\frac{r}{2M})^{\frac{1}{2}}$	$\tilde{U} = e^{u} = \sqrt{-\rho}e^{-t}$
$\tilde{V} = e^{\frac{v}{4M}} = e^{\frac{t}{4M}}e^{\frac{r}{4M}}(1-\frac{r}{2M})^{\frac{1}{2}}$	$\tilde{V} = e^{v} = \sqrt{-\rho}e^{t}$
$\tilde{U} = V - U; \quad \tilde{V} = V + U; \quad \tilde{U}\tilde{V} = e^{\frac{r}{2M}}(1-\frac{r}{2M}) > 0$	$\tilde{U} = V - U; \quad \tilde{V} = V + U; \quad \tilde{U}\tilde{V} = -\rho > 0$
$\tilde{U} > 0, \tilde{V} > 0$	$\tilde{U} > 0, \tilde{V} > 0$
$U = e^{\frac{r}{4M}}(1-\frac{r}{2M})^{\frac{1}{2}}sinh(\frac{t}{4M})$	$U = \sqrt{-\rho}\ sinh\ t$
$V = e^{\frac{r}{4M}}(1-\frac{r}{2M})^{\frac{1}{2}}cosh(\frac{t}{4M})$	$V = \sqrt{\rho}\ cosh\ t$
$ds^2 = -(1-\frac{2M}{r})(dt^2 - dr^{*2}) = (1-\frac{2M}{r})(dudv)$	$ds^2 = \rho dudv$
$ds^2 = \frac{32M^3}{r}e^{-\frac{r}{2M}}d\tilde{U}d\tilde{V}$	$ds^2 = -d\tilde{U}d\tilde{V}$
$ds^2 = \frac{32M^3}{r}e^{-\frac{r}{2M}}(dU^2 - dV^2)$	$ds^2 = dU^2 - dV^2$

In addition to these two regions one has two more regions with $\tilde{V} < 0$ (and either sign for $\tilde{U}$). In these two regions one simply changes the sign of U, V.

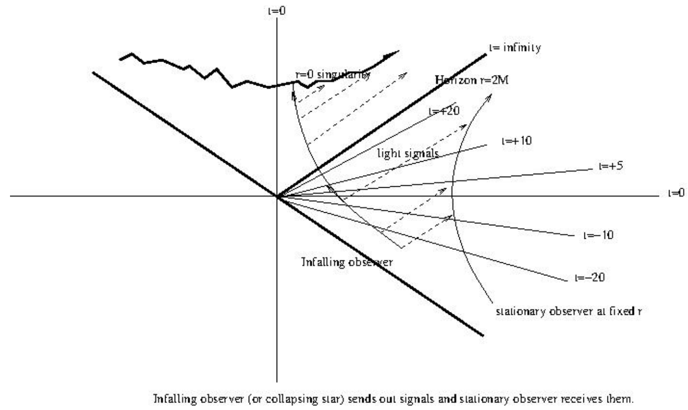


Figure 24: Collapsing star and stationary observer in the extended space-time.

Thus for Rindler, Region IV:

$$\begin{aligned} U &= -x\ cosh\ t = -\sqrt{\rho}\ cosh\ t, \\ V &= -x\ sinh\ t = -\sqrt{\rho}\ sinh\ t \end{aligned} \tag{80}$$

which gives $\tilde{U} = xe^{-t}, \tilde{V} = -xe^{t}$. This is just the negative of Region I.
For Region III,

$$\begin{aligned} U &= -x\ sinh\ t = -\sqrt{-\rho}\ sinh\ t, \\ V &= -x\ cosh\ t = -\sqrt{-\rho}\ cosh\ t \end{aligned} \tag{81}$$

which gives $\tilde{U} = -xe^{-t}, \tilde{V} = xe^{t}$, which is just the negative of region II.
For Schwarzschild Region IV

$$\begin{aligned} U &= -(\frac{r}{2M} - 1)^{\frac{1}{2}} e^{\frac{r}{4M}} cosh(\frac{t}{4M}) \\ V &= -(\frac{r}{2M} - 1)^{\frac{1}{2}} e^{\frac{r}{4M}} sinh(\frac{t}{4M}) \end{aligned} \tag{82}$$

For Schwarzschild Region III

$$\begin{aligned} U &= -(1 - \frac{r}{2M})^{\frac{1}{2}} e^{\frac{r}{4M}} sinh(\frac{t}{4M}) \\ V &= -(1 - \frac{r}{2M})^{\frac{1}{2}} e^{\frac{r}{4M}} cosh(\frac{t}{4M}) \end{aligned} \tag{83}$$

7.5 Penrose Diagram

20. Penrose invented a convenient way of representing these space-times, emphasizing the "conformal structure". This means bring infinity close by (since distance is not important we can rescale by an infinite amount) while preserving the light cones as 45 degree lines. So the causal structure is clear.

The basic idea is to work with light cone coordinates and perform coordinate transformations on them to make their range finite.

21. **Penrose Diagram for flat space:**

    (a) Step 1: Define light cone variables:

    $$v = t + r, \quad u = t - r$$

    This means that light trajectories are given by $u = const$(outgoing) or $v = const$(ingoing). One can draw 45 degree lines and let them stand for these trajectories.

    (b) Step 2: Make the range finite:

    $$u' = tan^{-1}\, u, \qquad v' = tan^{-1}\, v$$

    $$-\frac{\pi}{2} \leq u' \leq \frac{\pi}{2}, \qquad -\frac{\pi}{2} \leq v' \leq \frac{\pi}{2}$$

    (c) Step 3:
    Define new "space" and "time" coordinates by taking linear combinations:

    $$v' = t' + r', \quad u' = t' - r'$$

    In the $r' - t'$ graph, light cones continue to be at 45 degrees.

    (d) Step 4:
    Draw the diagram and identify various interesting regions and give them names.

8 Einstein's Equations

8.1 Einstein's Equation with Source

1. Einstein's equation is derived in "Mathematical Digressions III" (127).

$$(R_{ab} - \frac{1}{2}g_{ab}R) = 0$$

This is the equation without source. The source hs to be a conserved symmetric two tensor - which should reduce to "mass" in the Newtonian limit. The obvious candidate is the energy-momentum tensor (also called the stress-energy tensor) $T_{\mu\nu}$. To get the normalization we look at the Newtonian limit of $G_{\mu\nu}$: We start with the Schwarzschild metric (47):

$$ds^2 = -(1 - \frac{2M}{r})dt^2 + \frac{dr^2}{(1 - \frac{2M}{r})} + r^2(d\theta^2 + sin^2\, \theta d\phi^2)$$

If assume $r >> M$ we have

$$ds^2 = -(1 + 2\phi)dt^2 + (1 - 2\phi)dr^2 + r^2 d\Omega^2$$

with $\phi = -\frac{M}{r}$. Unfortunately the metric does not look isotropic, as it should in Newtonian mechanics. This is easily remedied by working with isotropic coordinates for the Schwarzschild metric:

$$ds^2 = (1 + \frac{M}{2\bar{r}})^4(d\bar{r}^2 + \bar{r}^2 d\Omega^2) - \frac{(1 - \frac{M}{2\bar{r}})^2}{(1 + \frac{M}{2\bar{r}})^2}dt^2$$

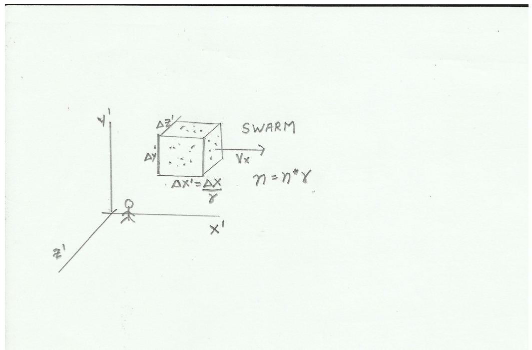


Figure 25: Fluid volume element in motion as see by an observer.

Here $r^2 = \bar{r}^2(1 + \frac{M}{2\bar{r}})^4$. For small $M << r$ we can approximate $r = \bar{r} + M$. Then to leading order in r the metric becomes

$$ds^2 = -(1+2\phi)dt^2 + (1-2\phi)(d\bar{r}^2 + \bar{r}^2 d\Omega^2) \approx -(1+2\phi)dt^2 + (1-2\phi)(dr^2 + r^2 d\Omega^2)$$

with $\phi = -\frac{M}{r}$. This has the isotropic form. We calculate G_{00} to find

$$G_{00} = 2\nabla^2\phi$$

On the other hand $T_{00} = \mu$ energy density. So if we let

$$G_{00} = 8\pi G T_{00}$$

this becomes the Newtonian equation

$$\nabla^2\phi = 4\pi G\mu$$

Thus we can obtain the normalization required for Einstein's equation.

$$G_{\mu\nu} = 8\pi T_{\mu\nu} \tag{84}$$

Also LHS satisfies $\nabla^\mu G_{\mu\nu} = 0$ and thus imposes energy momentum conservation $\nabla^\mu T_{\mu\nu} = 0$. Note that the conservation law is covariant.

8.2 Understanding the Sources for Einstein's eqn.

2. "Swarms of Particles": Fluid, gas, plasma etc. Consider a box of volume V with N particles. Value of N is frame independent. The value of V is frame dependent. Take "rest frame" of box. Call it V^*.

$$n^* = \frac{N}{V^*} = \#density\ in\ rest\ frame$$

Connection between $n = \frac{N}{V}$ and n^*. Let $V^* = \Delta x \Delta y \Delta z$. Consider an observer who sees the box move with velocity $\vec{v} = v_x \hat{i}$. Then in his (primed) frame $\Delta X' = \frac{\Delta X}{\gamma}$, $\Delta Y' = \Delta Y$, $\Delta Z' = \Delta Z$. See fig 21. Thus

$$V = \frac{V^*}{\gamma} \qquad \therefore n = \gamma n^*$$

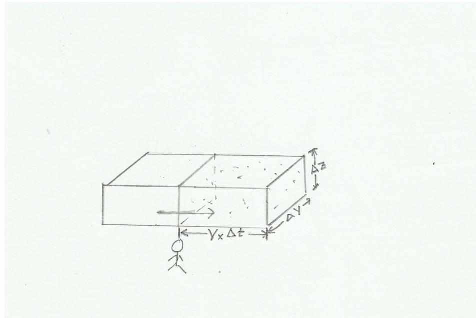


Figure 26: Flux of particles as seen by an observer.

3. Let $\vec{U}_{box}$ be 4-velocity of the box. Then $U^0_{box} = \gamma$ in the observer's frame and $U^0_{box} = 1$ in the rest frame. Suppose we let $\vec{n} = n^* \vec{U}_{box}$. Then, in the rest frame $n^0 = n^*$. In the primed frame $n^0 = \gamma n^*$. Thus the time component of this vector is the number density in any frame. Thus we can define number density n, in any observer's frame as

$$n = -\vec{n}.\vec{U}_{observer}$$

4. What are the other components of $\vec{n}$? $n^i = n^* \gamma v^i = n v^i$ where v^i is the three velocity of the box in any frame. Suppose it is moving in the x-direction. By the usual arguments: In time Δt a box of length $v^x \Delta t$ passes by the observer. Thus $\Delta X' = v^x \Delta t$. So the number of particles that pass by in a cross section $\Delta Y \Delta Z$, in time Δt, is $n \times volume \ = n v^x \Delta t \Delta Y \Delta Z$. So the **flux** he sees (flux =# particles per unit time per unit cross sectional area) is $n v^x = n^* \gamma v^x = n^x$. Thus the space components of $\vec{n}$ give the flux. (See figure 22)

   Thus

$$n^\mu = n^* U^\mu$$

   gives the four-current of particles. Here more generally we can think of U^μ as a 4-velocity field.

5. **4-divergence**

$$\partial_\mu n^\mu = \partial_t n^0 + \partial_i n^i$$

   Integrate over a volume V

$$\int_V d^3x \ (\partial_t n^0 + \partial_i n^i) = \underbrace{\frac{d}{dt} \int_V d^3x \ n^0}_{rate\ of\ change\ of\ total\ number} + \underbrace{\int_{\partial V} n^i dS_i}_{flux\ flowing\ out\ of\ volume}$$

   If particles are neither created nor destroyed, the two should add up to zero. Then

$$\partial_\mu n^\mu = 0$$

   and we say the number current is **conserved**.

6. Covariantization: In the above we used our usual notions of space and time and combined everything into 4-vectors. We can be more general and allow different space-like hypersurfaces to be our "space". This is important in curved space-time.

   Thus consider a general 3-volume in space-time - a hypersurface. It can be specified by a normal $\hat{n}$. How do you do that?

Let $f(x^\mu) = const$ be the hypersurface. $\underline{df}$ defines a normal *covector* (not vector). In a coordinate basis:

$$\underline{df} = \frac{\partial f}{\partial x^i}\underline{dx^i} \equiv \underline{n}$$

Then $n_i = \frac{\partial f}{\partial x^i}$ are the components of a normal covector, $\underline{n}$. One can however also define a normal vector $\vec{n} = n^i \frac{\vec{\partial}}{\partial x^i}$, from this by

$$n^i = g^{ij} n_j \quad or \quad \underline{n} = g(\vec{n}, \_)$$

Neither the normal covector, nor the vector defined above are normalized. This can always be done - provided it is not null.

Thus we let our space like volume be defined by a unit normal vector $\hat{n}$. Then we can write

$$d^3S^0 = \hat{n}^0 \Delta V$$

We assumed above that the normal vector has only a time component - this being normal to the usual space. However more generally

$$d^3S^\mu = \hat{n}^\mu \Delta V$$

This can point in any direction. In general ΔV is a three volume element in a 4-dim space-time, but it need not be purely spatial. We can also define using the normal covector (normalized to 1)

$$d^3S_\mu = \underline{\hat{n}}_\mu \Delta V$$

This is a three dimensional hypersurface volume element in a 4-volume, specified by a normal vector just as in the earlier example dS_i was the 2-surface in a three volume.

Let us define the number current to be $\vec{N}$ so as not to confuse with normal vectors:

$$N^\mu d^3S_\mu = \vec{N}.\hat{n}^\mu \Delta V$$

defines a generalized flux normal to the 3-hypersurface. (We have kept the upstairs μ index to show that it is a vector, not a covector. Thus if $\hat{n} = (1,0,0,0)$ then $\vec{N}.\hat{n}^\mu \Delta V = N^0 \Delta V = N^0 \Delta X \Delta Y \Delta Z$. Thus the number of particles in a volume element is a "flux" in the time direction in a 4 dim spacetime.

If we let $\hat{n} = (0,1,0,0)$ then $d^3S^\mu = \Delta V = \Delta t \Delta Y \Delta Z$ and

$$\vec{N}.\hat{n}^\mu \Delta V = n^x \Delta t \Delta Y \Delta Z = N^0 v^x \Delta t \Delta Y \Delta Z$$

is the number of particles crossing a surface $\Delta Y \Delta Z$ in time Δt in the x-direction.

In general we get a flux through the n^μ hypersurface.

7. There are other generalizations. We can define the electric current (density) $j^\mu = eN^\mu$ where e is the charge on each particle. Thus if we are talking about electrons, then each electron carries the same charge, and the number current and electric current are proportional.

$$\vec{j} = n^* e \vec{U}$$

Time component $j^0 = \rho = ne = \gamma n^* e$ is the charge density as seen by the observer.

One can also have non-Abelian currents -eg isospin currents where there are different kinds of charge. Thus for pions we have $\pi^\pm$ with $I_3 = \pm 1$ and π^0 with $I_3 = 0$. Thus we can define fluxes of $\pi^\pm, \pi^0$ particles - a multiplet of three currents, one for each type of particle: $(\vec{j}^+, \vec{j}^-, \vec{j}^0)$. If we call the isospin index a, then we have defined $j^{\mu a}$. Here μ is the usual space-time index and a is the iso-spin index. If isospin is conserved we can write

$$\partial_\mu j^{\mu a} = 0 : \quad a = +, -, 0$$

Note that quantities like electric charge e and isospin are scalars - they do not depend on the frame. The charge of the electron is fixed in any frame. Just as the number of particles is fixed in any frame.

8. Can we analogously define an energy current , say ϵU^μ where ϵ is energy of a particle. ϵ is not a scalar so this won't do. We can define $m_0 U^\mu$ where m_0 is the rest mass. This of course is the 4-momentum!

   If we still want to study flux of energy we can try to define ϵU^μ but we have to think of ϵ as the time component of a four vector. So we have to think of ϵU^μ as the time component of an object with *two* space -time indices. This is in fact the energy momentum tensor $T_{\mu\nu}$.

   Consider a hypersurface element $\hat{n}^\mu \Delta V$. Then $T^{\mu\nu}\hat{n}_\nu \Delta V$ is the flow of μ component of four momentum through that volume element. (Here the μ index is playing the role of a index for isospin - it is specifying the kind of charge that is flowing - in this case some component of momentum is that charge.)

   Consider various cases:

   a) $n^\mu = (1,0,0,0)$. In this case d^3S_μ is the usual volume element $\Delta X \Delta Y \Delta Z$. Then $T^{\mu 0}\Delta V$ is the amount of 4-momentum in this three volume spatial element. Thus $T^{00}\Delta V$ is the energy $\epsilon \Delta V$.

   $T^{i0}\Delta V = \pi^i \Delta V$ is the total amount of i-component of momentum, p^i, in that spatial volume. (π^i is then the momentum density.)

   Let us consider our swarm of particles. In the rest frame of the box energy density is $m_0 n^*$. Here m_0 is the rest mass. If the particles are moving, $m_0 \to \gamma m_0 = m,\ n^* \to \gamma n^* = n$. Thus we get $mn = m_0 n^* \gamma^2 = \epsilon = T^{00}$.

   Similarly momentum density is

$$mnv^i = m_0 n^* \gamma^2 v^i = \epsilon v^i = T^{i0}$$

   Clearly these are special cases of

$$T^{\mu\nu} = m_0 n^* U^\mu U^\nu = \mu U^\mu U^\nu$$

   - a nice covariant object. Here $\mu = m_0 n^*$ is a scalar and is equal to the energy density *in the rest frame of the box*. With this definition it is also clear that

$$T^{\mu\nu} = T^{\nu\mu}$$

   b) Consider $\underline{\hat{n}}_\mu = (0,1,0,0)$.

$$T^{\mu\nu}\hat{n}_\nu \Delta V = T^{\mu x}\Delta Y \Delta Z \Delta t = m_0 n^* U^\mu \gamma v^x \Delta t \Delta Z \Delta Y = p^\mu n v^x \Delta t \Delta Z \Delta Y$$

   Here $nv^x \Delta t \Delta Z \Delta Y$ is the number of particles in a volume element $v^x \Delta t \Delta Z \Delta Y$. This is thus a flux of particles that cross a cross section $\Delta Z \Delta Y$ in time Δt. Each particle carries 4-momentum p^μ so we get the total 4-momentum passing by the observer. So if we divide by $\Delta t \Delta Z \Delta Y$ we get flux i.e. per unit time per unit area of cross section. This is what $T^{\mu\nu}\hat{n}_\nu = T^{\mu x}$ is - the flux of 4-momentum p^μ, per unit time per unit area of cross section. gives.

   c) Let us consider T^{0x} specifically. By the above argument it is the flux of energy per unit time per unit area of cross section. The expression is

$$T^{0x} = mnv^x$$

   But this is also the total amount of p^x per unit volume in the observer's frame : n is the no.density of particles and mv^x is the momentum per particle. So the density of x-component of momentum is the same as the flux of energy in the x-direction. Energy × velocity is momentum. This is essentially what it means to say that $T^{0x} = T^{x0}$.

   d) Let us consider T^{ix}: This is the flux of p^i in the x-direction. But the rate of transfer of momentum is the force. So this is the i component of force exerted on a (unit area of) surface in the y-z plane (i.e. normal to x). It is a shear stress - when $i \neq x$.

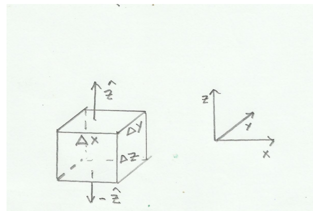


Figure 27: Stress on a volume element.

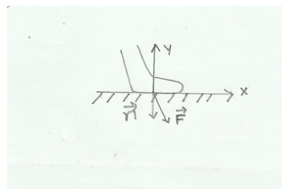


Figure 28: Force of the foot on the floor.

T^{xx} is the x-component of force acting on a unit area of the y-z plane - this is pressure.

We have to be careful about direction: In the expression $T^{ji}\hat{n}_i$, the normal $\hat{n}_i$ specifies the direction of momentum flow, and hence this expression gives the j'th component of force transmitted in the $\hat{n}_i$ direction. For eg if we have a unit normal $\hat{z}$ pointing upwards from a surface, then $T^{iz}n_z$ is the i'th comp of force by the part below the surface acting on the part above. ($n_z = +1$ is the component pointing upward and $n_z = -1$ is pointing downward.)

Similarly if we have a volume element with unit normals pointing outwards, then we get the force of the volume element on the region outside. (**see fig. 27**).

Example: (see fig.28) The normal $\hat{n}$ in the figure is pointing downward. So for eg $T^{xj}\hat{n}_j\Delta A = T^{xy}(-1)\Delta A = F_x = |F|sin\ \theta$ is the force in the x-direction of the foot acting on the floor. $T^{yj}\hat{n}_j\Delta A = T^{yy}(-1)\Delta A = F_y$ is the y-component of the force exerted by the foot on the floor. This is $-|F|cos\ \theta$ as shown in fig. The negative sign means it is acting downward.

9. The value of $T^{\mu\nu}$ depends on the frame. If an observer has 4-velocity $\vec{U}_{obs}$ then $T_{\mu\nu}U^{\mu}_{obs}U^{\nu}_{obs}$ gives the energy density in the observer's rest frame. Remember that in his frame $\vec{U}_{obs} = \hat{e}_0$, the normalized basis vector in the time direction. Similarly $T^{\mu\nu}\hat{e}_\mu\hat{e}_\nu$ gives the relevant components in his orthonormal frame.

10. Let us understand the conservation $\partial_\mu T^{\mu\nu} = 0$.

$$\frac{\partial T^{00}}{\partial t} + \frac{\partial T^{i0}}{\partial x^i} = 0 = \frac{\partial \epsilon}{\partial t} + \frac{\partial \Pi^i}{\partial x^i}$$

ϵ is the energy density and Π^i is the density of the i'th component of momentum (and also energy flux in the i'th direction). The three divergence of energy flow gives the net energy flowing out of a unit volume. Energy conservation says this must equal the rate of decrease of energy density.

Consider the i'th component of this equation:

$$\frac{\partial T^{i0}}{\partial t} + \frac{\partial T^{ij}}{\partial x^j} = 0 = \frac{\partial \Pi^i}{\partial t} + \frac{\partial T^{ij}}{\partial x^j} \tag{85}$$

Consider a volume element $\Delta X \Delta Y \Delta Z$ located at X, Y, Z. Consider the top surface with a unit normal $\hat{z}$ and the bottom surface with a unit normal $-\hat{z}$. (**See fig. 27**).

At the top $T^{xz}(Z+\Delta Z)(\hat{z})\Delta X \Delta Y$ is the x-component of force exerted by the volume on the surface area element $\Delta Y \Delta Z$ of the region outside, above it. The negative, $-T^{xz}(Z+\Delta Z)(\hat{z})\Delta X \Delta Y$ is the force exerted on the volume element by the outside.

At the bottom $T^{xz}(Z)(-\hat{z})\Delta X \Delta Y$ is the x-component of force exerted by the volume element on the region below at the surface element $\Delta X \Delta Y$. The negative of this $T^{xz}(Z)(\hat{z})\Delta X \Delta Y$ is the force exerted on the volume element by the outside.

The total x-force exerted by the outside on the volume element is $-T^{xz}(Z+\Delta Z)(\hat{z})\Delta Z \Delta Y + T^{xz}(Z)(\hat{z})\Delta X \Delta Y$. Taylor expanding, we get that the net force is $-\frac{\partial T^{xz}(Z)}{\partial z}(\hat{z})\Delta Z \Delta X \Delta Y$. This must equal (Newton's law) the net increase in x-component of momentum in that volume $= \frac{\partial \Pi^x}{\partial t}\Delta Z \Delta X \Delta Y$. This is the statement of conservation (85).

It is also true that

$$\int_V d^3x \ \frac{\partial T^{xj}(Z)}{\partial x^j} = \int_{\partial V} d^2A \ \hat{n}_{j \ out} T^{xj}$$

$n_{j \ out}$ is the outward pointing normal. The RHS gives the i'th component force on the outside by the volume element. Thus the integrated law becomes

$$\frac{\partial}{\partial t}\int_V d^3x \ \Pi^x = \int_{\partial V} d^2A \ \hat{n}_{j \ in} T^{xj}$$

where we have changed the sign and replaced the outward pointing normal by the inward pointing one.

11. **Perfect fluid:** In its rest frame a perfect fluid only has pressure, p, and energy density, ρ. There is no shear force because there is no viscosity. So in the rest frame of the fluid element it has the form:

$$T^{\mu\nu} = \begin{pmatrix} \rho & & & \\ & p & & \\ & & p & \\ & & & p \end{pmatrix}$$

Note: Energy density has the same dimensions as pressure: Energy density = Force× distance/Volume = Force/Area This can be written as a tensor:

$$T^{\mu\nu} = (\rho + p)U^\mu U^\nu + \eta^{\mu\nu} p \tag{86}$$

In curved space replace $\eta^{\mu\nu}$ by $g^{\mu\nu}$. Note that ρ, p are scalars and are defined by the rest frame. Also $\partial_\mu (n^* U^\mu) = 0$ is our particle number current conservation. So one can then see that $\partial_\mu T^{\mu\nu} = 0$ in flat space. In curved space this must become $\nabla_\mu T^{\mu\nu} = 0$. Note that for each component μ, $T^{\mu\nu}$ is a current. So this is the usual conservation of currents.

12. **Application to Cosmology:**

$$ds^2 = -dt^2 + a(t)^2(dx^2 + dy^2 + dz^2)$$

is the metric used in cosmology. The universe according to this is isotropic (rotationally symmetric) and homogeneous (translation invariance). Assume there is a perfect fluid filling all space. Let $\vec{U}$ be the 4-velocity of a fluid element. $\vec{U}.\vec{U} = -1$. In the above coordinates $U^\mu = (1,0,0,0)$. Because rotational and translation invariance means there is no special direction. So $\vec{U}$ cannot point in any direction. So

$$T^{00} = \rho(t); \quad T^{ij} = g^{ij} p(t) = \delta^{ij} \frac{p(t)}{a(t)^2}$$

Homeogeneity demands that there can only be a dependence on t. Write out the t component of the conservation equation:

$$\nabla_\beta T^{0\beta} = \frac{\partial T^{0\beta}}{\partial x^\beta} + \Gamma^0_{\beta\gamma} T^{\gamma\beta} + \Gamma^\beta_{\beta\gamma} T^{0\gamma} = 0$$

The non vanishing components of Γ are:

$$\Gamma^0_{ij} = a\dot{a}\delta_{ij}, \quad \Gamma^i_{jt} = \frac{\dot{a}}{a}\delta^i_j$$

So we get

$$\dot{\rho} + 3\frac{\dot{a}}{a}p + 3\frac{\dot{a}}{a}\rho = 0$$

or equivalently

$$\frac{d}{dt}(\rho a^3) = -p\frac{da^3}{dt}$$

Consider a fixed coordinate volume $\Delta V_{coord} = \Delta x \Delta y \Delta z$. Multiply the above equation by this volume:

$$\frac{d}{dt}(\rho a(t)^3 \Delta V_{coord}) = -p\frac{d(a(t)^3 \Delta V_{coord})}{dt}$$

$a(t)^3 \Delta V_{coord}$ is the physical volume defined by the coordinate volume. It changes with time because $a(t)$ depends on time. Thus LHS is the rate of change of energy in a physical volume element. RHS is the work done on the volume element. It is negative because the fluid is doing work against the external pressure as it expands. This equation expresses the conservation of energy.

8.3 Linearized Theory

13. If the deviation from flat space is small, then we can linearize Einstein's equation about flat space. These equations are similar to those of Maxwell electromagnetism, which is linear. A little more generally we can take any solution of Einstein's equation as our starting approximation, and study small deviations from that. The deviations are there because there are some extra source terms added as perturbation. If these are small, then the deviations are small

$$g_{\mu\nu} = g^0_{\mu\nu} + h_{\mu\nu}$$

Assume $h_{\mu\nu} << g^0_{\mu\nu}$. Suppose $g^0_{\mu\nu}$ is an exact solution. We then take the LHS of Einstein's equation (84) and keep terms upto linear order in $h_{\mu\nu}$.

Thus

$$\begin{aligned} \Gamma &= \Gamma^0 + \delta\Gamma^0 \\ R &= R^0 + \delta R \end{aligned}$$

For the vacuum solution $R^0 = 0$. Thus

$$\delta R_{\mu\nu} = 0$$

Take our exact solution to be flat space

$$g^0_{\mu\nu} = \eta_{\mu\nu} \implies \Gamma^0 = 0$$

$$\begin{aligned} \delta\Gamma^\rho_{\mu\nu} &= \frac{1}{2}g^{\rho\sigma}(g_{\sigma\mu,\nu} + g_{\sigma\nu,\mu} - g_{\mu\nu,\sigma}) \\ &= \frac{1}{2}\eta^{\rho\sigma}(h_{\sigma\mu,\nu} + h_{\sigma\nu,\mu} - h_{\mu\nu,\sigma}) \end{aligned} \tag{87}$$

Palatini's equation (125)

$$\delta R^{\alpha}{}_{\mu\rho\nu} = \delta\Gamma^{\alpha}{}_{\mu\nu,\rho} - \delta\Gamma^{\alpha}{}_{\mu\rho,\nu}$$

Contracting indices:

$$\delta R_{\mu\nu} = \delta\Gamma^{\rho}{}_{\mu\nu,\rho} - \delta\Gamma^{\rho}{}_{\mu\rho,\nu}$$

Substituting (87) one obtains the equation

$$\delta R_{\mu\nu} = \frac{1}{2}[-h_{\mu\nu}{}'^{\rho}{}_{\rho} + \partial_\mu V_\nu + \partial_\nu V_\mu] = 0$$

where

$$V_\mu = h^{\rho}{}_{\mu,\rho} - \frac{1}{2}h^{\rho}{}_{\rho,\mu} \tag{88}$$

14. **Gauge Invariance:**

These eqns must be invariant under general coordinate transformations:

$$g'_{\mu\nu}(x') = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta}(x)$$

Let $x^{\mu'} = x^\mu + \xi^\mu(x)$ or $x^\mu = x^{\mu'} - \xi^\mu(x')$ - to first order in ξ these are the same.

$$\frac{\partial x^\alpha}{\partial x'^\mu} = \delta^\alpha_\mu - \frac{\partial \xi^\alpha}{\partial x'^\mu} = \delta^\alpha_\mu - \frac{\partial \xi^\alpha}{\partial x^\mu}$$

$$\begin{aligned}\therefore g'_{\mu\nu} = \eta_{\mu\nu} + h'_{\mu\nu} &= (\delta^\alpha_\mu - \frac{\partial \xi^\alpha}{\partial x^\mu})(\delta^\beta_\nu - \frac{\partial \xi^\beta}{\partial x^\nu})(\eta_{\alpha\beta} + h_{\alpha\beta}) \\ &= \eta_{\mu\nu} + h_{\mu\nu} - \frac{\partial \xi_\nu}{\partial x^\mu} - \frac{\partial \xi_\mu}{\partial x^\nu} \\ \therefore h'_{\mu\nu} &= h_{\mu\nu} - \frac{\partial \xi_\nu}{\partial x^\mu} - \frac{\partial \xi_\mu}{\partial x^\nu}\end{aligned}$$

is the infinitesimal change in the metric due to coordinate changes - and has the form of a gauge transformation in electromagnetism.

15. **Counting physical states:**

One can use this to choose a coordinate system in which some components of $h_{\mu\nu}$ are set to zero. Since there are four arbitrary functions, one can set four components (out of 10) to zero. We can set $V_\mu = 0$ - "Lorentz Gauge" - and then the equations take the standard wave equation form:

$$-\frac{1}{2}\Box h_{\mu\nu} = 0$$

Thus we have 10-4=6 polarizations in a gravitational wave. (This is analogous to what happens in electromagnetism also. There choosing Lorentz gauge $\partial_\mu A^\mu = 0$ leaves us with 3 components obeying the wave equation.)

Again just as there, one can further make gauge transformations: If the gauge parameter, ξ, satisfies $\Box \xi_\mu = 0$, then V_μ does not change:

$$\delta V_\mu = \partial_\mu(\frac{\partial \xi_\nu}{\partial x^\mu} + \frac{\partial \xi_\mu}{\partial x^\nu}) - \frac{1}{2}\partial_\nu(2\frac{\partial \xi_\mu}{\partial x^\mu}) = \Box \xi_\mu$$

Thus even after setting $V_\mu = 0$, we still have some gauge freedom left. We have 4 parameters ξ_μ, obeying $\Box \xi_\mu = 0$. They can be used to set 4 more components of $h_{\mu\nu}$ - components *that obey the same wave equation* - to zero. Thus the plane wave - that had 6 components, now has 6-4=2 components.

These are the two polarizations that one expects for any massless spinning particle - spin ± 2 along the propagation direction. (In electromagnetism we have one further gauge transformation Λ obeying $\Box\Lambda = 0$ that preserves the Lorentz gauge and that reduces 3 polarizations to 2.)

So we have two polarizations and the wave solution looks like $h_{\mu\nu}(t,\vec{x}) = h_{\mu\nu}(\omega,\vec{k})e^{i(\vec{k}.\vec{x}-\omega t)}$.

16. We need to study the form of the matrix $h_{\mu\nu}(\omega,\vec{k})$. The Lorentz gauge condition is

$$h^{\rho}{}_{\mu,\rho} - \frac{1}{2}h^{\rho}{}_{\rho,\mu} = 0$$

In this gauge the wave equation is

$$\Box h_{\mu\nu} = 0$$

Let us choose $h^{\rho}{}_{\rho} = 0$ - traceless. This is one condition. Also choose $h_{ti} = 0$ - three conditions. The Lorentz condition becomes $h^{\rho}{}_{\mu,\rho} = 0$. These are:

$$h^{t}{}_{t,t} = 0 \quad ; \quad h^{j}{}_{i,j} = 0$$

It also has to satisfy the wave equation.

Let us take a wave propagating in the z direction for concreteness. Then these become

$$h^{z}{}_{i,z} = 0 = k_z h^{z}{}_{i} \implies h_{zi} = 0$$

and

$$\omega h_{tt} = 0 \implies h_{tt} = 0$$

This leaves us with a 2×2 matrix in the $x-y$ space. Thus tracelessness and symmetry of the remaining matrix implies that the gravitational wave looks like

$$h_{\mu\nu}(x) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & b & -a & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} e^{i(k_z z - \omega t)}$$

The matrix is transverse and traceless and this is called the "transverse-traceless gauge" or "tt gauge".

17. **Solving the wave equation:**

1. A solution of the homogeneous wave equation (i.e. without source) is $f(\omega t \pm \vec{k}.\vec{x})$ with $\frac{\omega}{|k|} = c = 1$.

2. Green function: Solve

$$\Box\phi = \delta(t)\delta^3(x)$$

Answer:

$$\phi(x,t) = \frac{f(t-r)}{r} + \frac{g(t+r)}{r}$$

We need to find f, g. The factor $\frac{1}{r}$ is required so that energy is conserved - energy flux through concentric spheres of area $4\pi r^2$, should be the same, and energy flux density is proportional to $\dot{\phi}^2$, which should therefore go as $\frac{1}{r^2}$. To see that this is the solution, one way is to integrate both sides of the equation over a small spherical volume of radius ϵ surrounding the origin.

$$(-\frac{\partial}{\partial t^2} + \vec{\nabla}^2)(\frac{f(t-r)}{r}) = \delta(t)\delta^3(x)$$

$$-\frac{\partial}{\partial t^2}\int_{Vol\ of\ radius\ \epsilon} d^3x\ \frac{f(t-r)}{r} \to 0\ as\ \ \epsilon \to 0$$

because the derivative acting on f gives something finite and the volume goes to zero as ϵ^3.

$$\int d^3x\ \nabla^2(\frac{f(t-r)}{r}) = \int_{r=\epsilon} d^2\vec{S}_r.\vec{\nabla}(\frac{f(t-r)}{r}) = \int_{r=\epsilon} d^2\vec{S}_r \frac{\partial}{\partial r}(\frac{f(t-r)}{r})$$

$$4\pi\epsilon^2[-\frac{f(t-\epsilon)}{\epsilon^2} - \frac{f'(t-\epsilon)}{\epsilon}] \to -4\pi f(t) \quad as \quad \epsilon \to 0$$

Thus

$$LHS = -4\pi f(t)$$

$$RHS = \int \delta(t)\delta^3(x) = \delta(t)$$

Thus

$$f(t) = -\frac{\delta(t)}{4\pi}$$

Thus the solution is

$$\frac{f(t-r)}{r} = -\frac{\delta(t-r)}{4\pi r}$$

The delta fn is non zero along a light cone. Not surprising. Massless particles travel at the speed of light. For a general source $J(t,r)$ which can always be written as a superposition of delta functions:

$$J(t,r) = \int d^3r'dt'\ J(t',r')\delta(t-t')\delta^3(r-r')$$

one obtains

$$\phi(t,r) = -\int d^3r'dt'\ \frac{\delta(t-t'-(r-r'))}{4\pi(r-r')} J(t',r')$$

The t' integral can be done and fixes t' to be the retarded time. $t' = t-(r-r') = t_{retarded}$. Thus the soln is

$$\phi(t,r) = -\int d^3r'\ \frac{J(t_{ret},r')}{4\pi(r-r')}$$

The soln to the em or grav wave eqn in the Lorentz gauge follows in the same way.

18. In the presence of sources the linearized Einstein's equation, in the Lorentz gauge, is

$$-\frac{1}{2}(\Box h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\Box h^{\rho}{}_{\rho}) = 8\pi G T_{\mu\nu}$$

The quantity $h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h^{\rho}{}_{\rho} \equiv \bar{h}_{\mu\nu}$ occurs often and is the trace reversed metric fluctuation. The Lorentz gauge condition is

$$V_\mu \equiv \partial_\rho \bar{h}^{\rho}{}_{\mu} = 0$$

Einstein's equation becomes

$$-\frac{1}{2}\Box\bar{h}_{\mu\nu} = 8\pi G T_{\mu\nu}$$

The solution is immediate:

$$\bar{h}^{\mu\nu} = \frac{16\pi G}{4\pi}\int d^3x'\ \frac{T^{\mu\nu}(t_{ret},r')}{(r-r')} \quad , \quad t_{ret} = t-(r-r')$$

Let us check if Lorentz gauge condn. is maintained.

$$\frac{\partial}{\partial x^\nu}\bar{h}^{\mu\nu} = 4G\int d^3x'\ \{[\frac{\partial}{\partial t}T^{\mu 0}(t_{ret},x')]\frac{1}{|x-x'|} + \frac{\partial}{\partial x^i}(\frac{T^{\mu i}(t_{ret},x')}{|x-x'|})\}$$

$$= 4G \int d^3x' \; \{[\frac{\partial}{\partial t_{ret}} T^{\mu 0}(t_{ret}, x')] \frac{1}{|x-x'|} + (\frac{\partial}{\partial x^i} T^{\mu i}(t_{ret}, x')) \frac{1}{|x-x'|} + T^{\mu i}(t_{ret}, x')(\frac{\partial}{\partial x^i} \frac{1}{|x-x'|})\}$$

$$= 4G \int d^3x' \; \{[\frac{\partial}{\partial t'} T^{\mu 0}(t', x')] \frac{1}{|x-x'|} + (\frac{\partial}{\partial x^i} T^{\mu i}(t_{ret}, x')) \frac{1}{|x-x'|} - T^{\mu i}(t_{ret}, x')(\frac{\partial}{\partial x'^i} \frac{1}{|x-x'|})\}$$

$$= 4G \int d^3x' \; \{[\frac{\partial}{\partial t'} T^{\mu 0}(t', x')] \frac{1}{|x-x'|} + (\frac{\partial}{\partial x^i} + \frac{\partial}{\partial x'^i}) T^{\mu i}(t - |x-x'|, x')) \frac{1}{|x-x'|}\}$$

$$= 4G \int d^3x' \; \{[\frac{\partial}{\partial t'} T^{\mu 0}(t', x')] \frac{1}{|x-x'|} + [\frac{\partial}{\partial x'^i} T^{\mu i}(t', x')] \frac{1}{|x-x'|}\}$$

$$= 4G \int d^3x' \; \{[\frac{\partial}{\partial x'^\nu} T^{\mu\nu}(t', x')] \frac{1}{|x-x'|}\} = 0$$

In the last step conservation of energy and momentum has been used. Thus the Lorentz gauge is maintained.

19. **Swarm of particles rotating like a rigid body:**

$$T^{\mu\nu} = m_0 n^* U^\mu U^\nu \equiv \mu U^\mu U^\nu$$

Here m_0 is the rest mass of the particle and n^* is the number density in the rest frame of a volume element.

$$T^{00} = \mu\gamma^2 \approx \mu \;\;,\;\; T^{0i} = \mu\gamma^2 v^i \approx \mu v^i$$

Let us assume that the rotation is steady, i.e. there is no time dependence. Also assume uniform density. The solution is

$$\bar{h}^{0i}(x) = 4G \int d^3x' \; \frac{\mu(x') v^i(x')}{|x-x'|}$$

Expand

$$\frac{1}{|x-x'|} = \frac{1}{r} + \frac{\vec{x}.\vec{x}'}{r^3} + ...$$

Also

$$v^x(x', y') = -\Omega y', \quad v^y(x', y') = \Omega x', \quad v^z = 0$$

Thus on the y-axis, where $\vec{x}.\vec{x}' = yy'$,

$$\bar{h}^{0x}(0, y) = -4G \int dx'dy'dz' \; [\frac{\mu\Omega y'}{r} + \frac{\mu\Omega y' \; yy'}{r^3} + ...]$$

The first term is zero - because of symmetry in y'. In the second term, also by symmetry,

$$\int dx'dy'dz' \mu y'^2 = \int dx'dy'dz' \mu x'^2 = \frac{1}{2} \int dx'dy'dz' \mu (x'^2 + y'^2) = \frac{1}{2} I_z$$

where I_z is the moment of inertia about the z-axis. Also $I_z \Omega = J_z$ - angular momentum about z-axis. Thus we get on the y- axis:

$$h^{0x} = -\frac{2GJy}{r^3}$$

h^{0x} is zero on the y-axis. It depends only on the “y” coordinate. Similarly on the x-axis

$$h^{0y} = \frac{2GJx}{r^3}$$

Thus in general the perturbed metric becomes (using $\frac{xdy-ydx}{r^2} = d\phi$)

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \underbrace{-\frac{4GJy}{r^3}dxdt + \frac{4GJx}{r^3}dydt}_{\frac{4GJ}{r}d\phi dt} \tag{89}$$

In azimuthal coordinates:

$$ds^2 = -dt^2 + dz^2 + dr^2 + r^2 d\phi^2 + \frac{4GJ}{r}d\phi dt \tag{90}$$

Suppose we start with

$$ds^2 = -dt^2 + dz^2 + dr^2 + r^2 d\phi^2$$

and go to a rotating (non inertial) coordinate system by writing $\phi = \phi' + \omega t$, then to first order in ω the metric acquires a cross term $2\omega r^2 d\phi' dt$. Thus we see that the effect of the rotating source is to make our stationary frame non inertial and rotating. In other words the inertial frame is one that is rotating w.r.t the stationary frame. The effective rotational velocity is

$$\omega = \frac{2GJ}{r^3}$$

This is called "frame-dragging" or "dragging of inertial frame" by a rotating star.

20. **Spin Precession: Lens-Thirring Effect**

Spins precess (even in the absence of any external torque) when viewed in an accelerated (non inertial) frame.

One can check this frame dragging effect by studying the motion of a gyroscope - spin - placed say on the z-axis, at a distance r from the centre. If the frame is rotating about the z- axis with an angular velocity $\omega = \frac{2GJ}{r^3}$, one should find the spin precesses at precisely this angular velocity.

The equation for a freely falling gyroscope is

$$\nabla_U \vec{S} = 0 \implies U^\alpha[\frac{\partial S^\gamma}{\partial x^\alpha} + \Gamma^\gamma_{\beta\alpha}S^\beta] = 0$$

$$\implies \frac{dS^\gamma}{d\tau} + U^\alpha \Gamma^\gamma_{\beta\alpha}S^\beta = 0$$

$$\frac{dS^0}{d\tau} + \Gamma^0_{\beta\alpha}U^\alpha S^\beta = 0, \qquad \frac{dS^x}{d\tau} + \Gamma^x_{\beta\alpha}U^\alpha S^\beta = 0$$

One can evaluate Γ in the metric (89).

$$\Gamma^\gamma_{\beta\alpha} = \frac{1}{2}\eta^{\gamma\delta}\Gamma_{\delta\beta\alpha}$$

$$= \frac{1}{2}\eta^\gamma(h_{\beta\delta,\alpha} + h_{\alpha\delta,\beta} - h_{\alpha\beta,\delta})$$

The only non vanishing derivatives (on the z axis) are

$$h_{0x,y} = \frac{2J}{r^3} = \frac{2J}{z^3}, \qquad h_{0y,x} = -\frac{2J}{z^3}$$

So

$$\Gamma^t_{xy} = 0, \quad \Gamma^x_{0y} = \frac{2J}{z^3}, \quad \Gamma^y_{0x} = -\frac{2J}{z^3}$$

Anything that involves 2 t indices, or z are zero. Thus

$$\frac{dS^0}{d\tau} = 0$$

$$\frac{dS^x}{d\tau} + \Gamma^x_{0y} U^0 S^y = 0 \implies \frac{dS^x}{d\tau} = -\frac{2J}{z^3} S^y$$

and

$$\frac{dS^y}{d\tau} = \frac{2J}{z^3} S^x$$

This is the expected precession.

21. **Quadrupole Formula**

In EM theory there is a well known dipole approximation for calculating em radiation for low frequency, long wavelengths. There is an analogous quadrupole formula for gravitational wave radiation.

Start with:

$$\bar{h}^{\alpha\beta} = 4 \int d^3x' \frac{T^{\alpha\beta}(t_{ret}, x')}{|x - x'|}$$

Here assume a compact source of size $\approx R_s$ and far away so that $x >> R_s$.

If $T(t_{ret}, x') \approx T(x') cos(\omega(t - |\vec{x} - \vec{x}'|))$ and if ω is small so that $\lambda >> R_s$, then replacing $t - |x - x'|$ by $t - |\vec{x}| = t - r$ will not make much difference. So using these approximations we get

$$\bar{h}^{\alpha\beta}(t, \vec{x}) = \frac{4}{r} \int d^3x' T^{\alpha\beta}(t - r, x') \tag{91}$$

We are interested in the space components $\bar{h}^{ij}$. We can rewrite this in a convenient form:

$$\frac{\partial T^{tt}}{\partial t} + \frac{\partial T^{kt}}{\partial x^k} = 0$$

$$\frac{\partial^2 T^{tt}}{\partial t^2} = -\frac{\partial}{\partial t}\frac{\partial T^{kt}}{\partial x^k} = -\frac{\partial}{\partial x^k}\frac{\partial T^{kt}}{\partial t} = +\frac{\partial^2}{\partial x^k \partial x^j} T^{kj}$$

Now consider

$$\int d^3x' x'^m x'^n \frac{\partial^2}{\partial x'^k \partial x'^j} T^{kj}(x') = \int d^3x' x'^m x'^n \frac{\partial^2 T^{tt}}{\partial t^2}$$

The LHS can be evaluated by an integration by parts to give

$$2 \int d^3x' T^{mn}$$

The RHS is (calling $T^{tt}(x') = \mu(x')$ the energy density:

$$\frac{\partial^2}{\partial t^2} \int d^3x' x'^m x'^n \mu(x') = \frac{\partial^2}{\partial t^2} I^{mn}$$

where I^{mn} is (the quadrupole) moment of inertia tensor.

Thus

$$\int d^3x' T^{mn}(x') = \frac{1}{2}\frac{\partial^2}{\partial t^2} I^{mn}$$

Plugging this into (91) we get

$$\bar{h}^{mn}(t, \vec{x}) = \frac{2}{r} \ddot{I}^{mn}(t - r)$$

This is the quadrupole formula for gravitational radiation.

9 Kerr Black Hole

9.1 Kerr Geometry

1. The Kerr metric is the following:

$$ds^2 = -(1-\frac{2Mr}{\rho^2})dt^2 - \frac{4Marsin^2\theta}{\rho^2}d\phi dt + \frac{\rho^2}{\Delta}dr^2 + \rho^2 d\theta^2 + (r^2+a^2+\frac{2Mra^2sin^2\theta}{\rho^2})sin^2\theta d\phi^2 \quad (92)$$

   where
$$a \equiv \frac{J}{M}, \quad \rho^2 \equiv r^2 + a^2cos^2\theta, \quad \Delta \equiv r^2 - 2Mr + a^2$$
   t, r, θ, ϕ are called "Boyer-Lindquist" coordinates.

2. It is asymptotically flat.

$$r >> M, a \implies \rho^2 \approx r^2, \quad \Delta \approx r^2 - 2Mr = r(1-\frac{2M}{r}), \qquad \implies \frac{\rho^2}{\Delta} \approx \frac{1}{(1-\frac{2M}{r})}$$

$$ds^2 = \underbrace{-(1-\frac{2M}{r})dt^2 + \frac{dr^2}{(1-\frac{2M}{r})} + r^2(d\theta^2 + sin^2\theta d\phi^2)}_{Schwarzschild} - \underbrace{\frac{4Ma}{r^2}sin^2\theta(rd\phi)dt}_{\frac{4J}{r}d\phi dt}$$

   It is clear from this that J is indeed the angular momentum.

3. Killing Vectors:

   Axisymmetric only (not spherically symmetric).

$$\vec{\xi} = \frac{\vec{\partial}}{\partial t}: \quad \xi^\alpha = (1,0,0,0); \qquad \vec{\eta} = \frac{\vec{\partial}}{\partial \phi}: \quad \eta^\alpha = (0,0,0,1)$$

   Existence of ξ means it is a "stationary" solution. The existence of diagonal $d\phi dt$ term in the metric means it is not "static". If a spacelike hypersurface can be found such that ξ is orthogonal to it, then if we label all the points on that by the time coordinate t then since ξ is perp to the surface the distance element will not have any cross term $dtdx$. (The neighbouring time slice $t + dt$ will use the same x coordinate associated with the starting point of ξ on the first time slice.)

4. As $a \to 0$ it becomes Schwarzschild.

5. There are two singularities: $\rho = 0$ and $\Delta = 0$.

$$\rho = 0 \implies r = 0 \quad \theta = \frac{\pi}{2}$$

   This is a real singularity.

$$\Delta = 0 \implies r = r_\pm = M \pm \sqrt{M^2 - a^2} \quad when\ a \leq M$$

   Both are coordinate singularities. r_+ is the horizon. If $M = a$ (and so $J = M^2$), then $r_+ = r_- = M$. This is called an extreme Kerr BH.

6. $a > M$ would mean that there is no horizon to cloak the singularity. This violates the cosmic censorship hypothesis, which is believed to be true.

   Astrophysically BH are often close to extreme, because particles falling in carry angular momentum even when they have radiated away most of the energy. So J can keep increasing. Detailed calculations show that $a < .998M$ seems to be the astrophysical bound.

   Energy release/unit mass in extreme Kerr BH is very large.

7. The horizon is a null surface (and one way - like all null sufaces). If $r < r_+$ light rays cannot come out. If $r = r_+$ they can hover around. These light rays generate the horizon.

   Let $\vec{t}$ be a tangent vector at the horizon. $t^\alpha = (t^t, 0, t^\theta, t^\phi)$ tangent to $r = r_+$. If one can find l^α of this type such that $l^\alpha l_\alpha = 0$ then this 3 surface is null.

$$\vec{l}.\vec{l} = g_{tt}(l^t)^2 + 2g_{t\phi}l^t l^\phi + g_{\phi\phi}(l^\phi)^2 + g_{\theta\theta}(l^\theta)^2 = 0$$

Details:

$$-(1 - \frac{2Mr_+}{\rho_+^2})(l^t)^2 - \frac{4Mar_+ sin^2\theta}{\rho_+^2} l^t l^\phi + (r_+^2 + a^2 + 2Mr_+ \frac{a^2 sin^2\theta}{\rho_+^2}) sin^2\theta (l^\phi)^2 + \rho_+^2 (l^\theta)^2 = 0 \quad (93)$$

Note that r_+ satisfies: $r_+^2 + a^2 = 2Mr_+$ thus

$$\rho_+^2 = r_+^2 + a^2 cos^2\theta = r_+^2 + a^2 - a^2 sin^2\theta = 2Mr_+ - a^2 sin^2\theta$$

and

$$(1 - \frac{2Mr_+}{\rho_+^2}) = (\frac{\rho_+^2 - 2Mr_+}{\rho_+^2}) = -\frac{a^2 sin^2\theta}{\rho_+^2}$$

and

$$(r_+^2 + a^2 + 2Mr_+ \frac{a^2 sin^2\theta}{\rho_+^2}) = (2Mr_+ + 2Mr_+ \frac{a^2 sin^2\theta}{\rho_+^2}) = (\frac{2Mr_+\rho_+^2 + 2Mr_+ a^2 sin^2\theta}{\rho_+^2})$$

$$= (\frac{2Mr_+(\rho_+^2 + a^2 sin^2\theta)}{\rho_+^2}) = \frac{(2Mr_+)^2}{\rho_+^2}$$

Plugging these in (93) we get

$$\frac{a^2 sin^2\theta}{\rho_+^2}(l^t)^2 - \frac{4Mar_+ sin^2\theta}{\rho_+^2} l^t l^\phi + \frac{(2Mr_+)^2}{\rho_+^2} sin^2\theta (l^\phi)^2 + \rho_+^2 (l^\theta)^2 = 0$$

$$\implies \frac{sin^2\theta}{\rho_+^2}[a^2(l^t)^2 - 4Mr_+ a l^t l^\phi + (2Mr_+)^2 (l^\phi)^2] + \rho_+^2 (l^\theta)^2 = 0$$

$$(\frac{2Mr_+ sin\theta}{\rho_+})^2 [(l^\phi - \frac{a}{2Mr_+} l^t)^2] + \rho_+^2 (l^\theta)^2 = 0$$

$$\implies l^\theta = 0 \quad \& \quad l^\phi = \frac{a}{2Mr_+} l^t$$

Thus

$$l^\alpha = (1, 0, 0, \Omega_H) \qquad \Omega_H = \frac{a}{2Mr_+}$$

For an extremal BH $r_+ = a = M$. Thus $\Omega_H = \frac{1}{2M}$.

Thus the horizon is generated by l^α above and $\frac{\partial}{\partial\theta} = (0, 0, 1, 0)$ and $\frac{\partial}{\partial\phi} = (0, 0, 0, 1)$

If $a = 0$ the null vector is just $\frac{\partial}{\partial t}$. When $a \neq 0$ the null vectors have an angular velocity $\frac{d\phi}{dt} = \Omega_H$. This can be called the angular velocity of the BH.

8. At $r = r_+$ the spatial slice has the metric

$$d\Sigma^2 = \rho_+^2 d\theta^2 + (\frac{2Mr_+}{\rho_+})^2 sin^2\theta d\phi^2$$

This is clearly not spherically symmetric. When $\theta = \frac{\pi}{2}$, $\rho_+^2 = 2Mr_+ - a^2 = r_+^2$. Thus

$$d\Sigma^2 = r_+^2 d\theta^2 + (2M)^2 d\phi^2$$

Thus the circumference around the equator is $4\pi M$. Whereas along a longitude the circumference is:

$$2\int_0^\pi d\theta\ \rho_+ = 2\int_0^\pi d\theta\ \sqrt{2Mr_+ - a^2 sin^2\theta}$$

We can choose the extremal case: $a = M = r_+$. Then we get

$$2\int_0^\pi M\sqrt{2 - sin^2\theta} \approx 7.64M < 4\pi M$$

Thus it is a squashed sphere.

9. The area of the horizon is

$$\int\int \rho_+ d\theta \frac{2Mr_+}{\rho_+} sin\theta d\phi = 2Mr_+ \int\int sin\theta d\theta d\phi = 8\pi M[M + \sqrt{M^2 + a^2}]$$

10. **Orbits of particles/light rays:**

We take the special case $\theta = \frac{\pi}{2}$. There is a symmetry $\theta \to \pi - \theta$, so there are orbits in the equatorial plane. When $\theta = \frac{\pi}{2}$, $\rho = r$.

$$ds^2 = -(1 - \frac{2M}{r})dt^2 - \frac{4aM}{r}d\phi dt + \frac{r^2}{\Delta}dr^2 + (r^2 + a^2 + \frac{2Ma^2}{r})d\phi^2$$

From this metric it is clear that aM is the angular momentum - from our earlier calculation in the linearized theory.

Killing Vectors, ξ, η as usual.

$$e = -\vec{\xi}.\vec{U}, \quad l = \vec{\eta}.\vec{U}, \quad \vec{U}.\vec{U} = -1$$

are the equations for $U^\alpha = (U^0, U^r, 0, U^\phi)$.

$$\begin{aligned} -e &= g_{tt}U^t + g_{t\phi}U^\phi \\ l &= g_{\phi t}U^t + g_{\phi\phi}U^\phi \end{aligned} \tag{94}$$

$$\begin{aligned} e &= (1 - \frac{2M}{r})U^t + \frac{2aM}{r}U^\phi \\ l &= -\frac{2aM}{r}U^t + (r^2 + a^2 + \frac{2Ma^2}{r})U^\phi \end{aligned} \tag{95}$$

Multiply e by $\frac{2aM}{r}$ and l by $(1 - \frac{2M}{r})$ and add:

$$l(1 - \frac{2M}{r}) + e(\frac{2aM}{r}) = (r^2 + a^2 - 2Mr)U^\phi = \Delta U^\phi$$

Thus

$$U^\phi = \frac{1}{\Delta}[(1 - \frac{2M}{r})l + (\frac{2aM}{r})e]$$

Similarly

$$U^t = \frac{1}{\Delta}[(r^2 + a^2 - 2Mr)e - (\frac{2aM}{r})l]$$

Write $\vec{U}.\vec{U} = -1$ to get

$$\frac{1}{2}(\frac{dr}{d\tau})^2 + V_{eff}(r;e,l) = \frac{e^2-1}{2}$$

where

$$V_{eff}(r;e,l) = -\frac{M}{r} + \frac{l^2 - a^2(e^2-1)}{2r^2} - \frac{M(l-ae)^2}{r^3} \tag{96}$$

Note that qualitatively the potential is similar to Schwarzschild with a negative $1/r^3$ term.

The corresponding equation for lightlike case impose $\vec{U}.\vec{U} = 0$ where $U^\mu = \frac{dx^\mu}{d\lambda}$. The equation is

$$\frac{1}{l^2}(\frac{dr}{d\lambda})^2 = \frac{1}{b^2} - W_{eff}(r;b,\sigma)$$

where

$$b = |\frac{l}{e}|, \quad \sigma = sign(l)$$

and

$$W_{eff}(r;b,\sigma) = \frac{1}{r^2}[1 - (\frac{a}{b})^2 - \frac{2M}{r}(1 - \sigma\frac{a}{b})^2] \tag{97}$$

11. **Special case of an orbit:**

Instead of studying general orbits, we consider a special case where $a = M$ (extremal BH) and a particle with $l = 2Me$. The reason for this choice is the following: If a particle of mass m falls into a BH then

$$\delta M = me, \quad \delta J = ml$$

$$\delta a = \delta(\frac{J}{M}) = \frac{\delta J}{M} - \frac{J\delta M}{M^2} = \frac{ml}{M} - \frac{ame}{M} = \frac{m}{M}(l - ae)$$

Violation of extremality requires $\delta a > \delta M$. Thus

$$\frac{m}{M}(l-ae) > me \quad \therefore l > (M+a)e = 2Me$$

So $l > 2Me$ would violate extremality and the $l = 2Me$ is the boundary.

So we study the orbit of this particle around an extremal BH.

$$V_{eff} = -\frac{M}{r} + \frac{(2Me)^2 - M^2e^2 + M^2}{2r^2} - \frac{M(Me)^2}{r^3}$$

Max is at:

$$\frac{dV}{dr} = \frac{M}{r^2} - \frac{3M^2e^2 + M^2}{r^3} + \frac{3M^2e^2}{r^4} = 0 = (1 - \frac{M}{r})(1 - \frac{3Me^2}{r})$$

$M = r$ is a solution. The other solution is the minimum.

$$V_{eff}|_{max} = -1 + \frac{3e^2+1}{2} - e^2 = \frac{e^2-1}{2}$$

This is just the total energy. Thus at the max $\frac{dr}{d\tau} = 0$ and the particle does not make it across the barrier. Thus there is no violation of extremality.

Consider the case $3e^2 = 1$. Then

$$\frac{dV}{dr} = (1 - \frac{M}{r})^2$$

Thus $\frac{d^2V}{dr^2} = 0$ at the max. This means the max and min merge and the stable orbit at the minimum is just about to become unstable. This a limiting case. The values are $e = \frac{1}{\sqrt{3}}$ and $l = \frac{2M}{\sqrt{3}}$. (e has to be less than this for stability.)

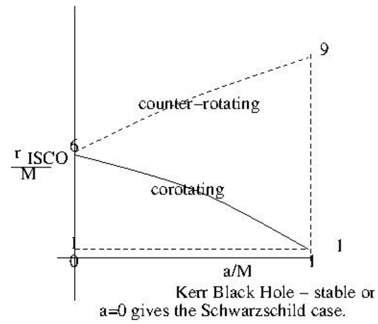


Figure 29: Innermost Stable Circular Orbits in a Kerr BH.

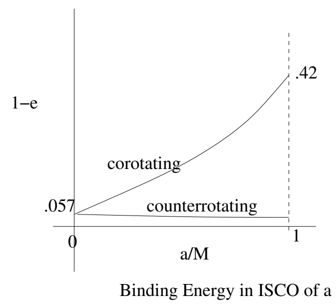


Figure 30: Binding energy in Innermost Stable Circular Orbits in a Kerr BH.

In general one has to satisfy three conditions for a stable orbit. The first one is the condition for it to be circular. The second one is says it should be at an extremum of the potential. The last one is the stability condition - the extremum should be a minimum.

$$
\begin{aligned}
V_{eff}(r;e,l) &= \frac{e^2-1}{2} \\
\frac{dV}{dr}|_{r=r_I} &= 0 \\
\frac{d^2V}{dr^2}_{r=r_I} &\geq 0
\end{aligned}
$$

If equality is satisfied in the last condition, then we have a limiting case, which gives the *innermost* stable circular orbit. We can plot the solutions as a function of $\frac{a}{M}$. $a=0$ is the Schwarzschild case $r_I = 6M$. The solution $r_I = M$ found above in the extremal case is the end point of one of the plots. (**See fig 25,26.**)

The second plot shows $1-e$ - energy released as the particle settles down in this orbit - the binding energy. For the Schwarzscild case this was 0.057 of the rest mass. In the extremal case it is $1-\frac{1}{\sqrt{3}} \approx 0.42$. This is a huge amount of energy. More realistically in astrophysical BH, when $a = .998M$ we get about 0.3. The solution given above is the "corotating" case. A counter rotating case is also plotted.

12. **Ergosphere**

$$ds^2 = -(1-\frac{2Mr}{\rho^2})dt^2 - \frac{4Marsin^2\theta}{\rho^2}d\phi dt + \frac{\rho^2}{\Delta}dr^2 + \rho^2 d\theta^2 + (r^2+a^2+\frac{2Mra^2sin^2\theta}{\rho^2})sin^2\theta d\phi^2$$

In the Schwarzschild case $\Delta = 0$ coincided with $g_{tt} = 0$, but not here. Here $\Delta = 0$ is the horizon: $r = r_+$ satisfying

$$r_+^2 + a^2 = 2Mr_+$$

$$g_{tt} = 0 \implies \rho^2 = r^2 + a^2 cos^2\theta = 2Mr$$

At $\theta = 0$ this is the same condition as $\Delta = 0$. However otherwise

$$r_e(\theta) = M + \sqrt{M^2 - a^2 cos^2\theta}$$

We pick the positive square root because we want $r_e > r_+$. At $\theta = \frac{\pi}{2}$, $r_e(\frac{\pi}{2}) = r_+$.

Note: For extremal BH $r_e(\theta) = M + Msin\theta$. This is M at the poles and $2M$ at the equator. The region between $r_e(\theta)$ and r_+ is the ergosphere. Here $g_{tt} \geq 0$.

13. What happens when $g_{tt} = 0$? If 4-velocity is $U^\mu = (U^0, 0, 0, 0)$ - i.e. a stationary observer, the $\vec{U}.\vec{U} = g_{tt}(U^0)^2 = 0$. Thus no stationary time like observer can exist. Inside the ergosphere, the stationary observer's 4-velocity is space-like. He has to be moving. Let his 4-velocity be (at fixed r, θ):

$$U_{obs} = (U^0, 0, 0, U^\phi) = U^0(1, 0, 0, \Omega_{obs})$$

Then

$$(U^\phi)^2 g_{\phi\phi} + 2g_{\phi t}U^\phi U^t = -1$$

At $r = r_e$ we have (Set $\theta = \frac{\pi}{2}$)

$$\rho^2 = r^2 + a^2 cos^2\theta = r^2 = 2Mr \implies r = 2M$$

and so

$$g_{\phi\phi} = r^2 + a^2 + \frac{2Mra^2 sin^2\theta}{\rho^2} = r^2 + 2a^2$$

$$g_{\phi t} = -2asin^2\theta = -2a$$

So

$$\Omega^2(r^2 + 2a^2) - \Omega 4a < 0 \implies 0 < \Omega < \frac{4a}{r^2 + 2a^2}$$

is the allowed range of values. On the other hand, on the horizon one finds

$$(U^\phi)^2 g_{\phi\phi} + 2g_{\phi t}U^\phi U^t + (U^t)^2 g_{tt} = (\Omega 2Mr - a)^2$$

Which can only be zero, never negative. So we get $\Omega = \frac{a}{2Mr_+} = \Omega_H$. Thus at the horizon only this value is allowed.

14. Extraction of energy is allowed from a Kerr BH - unlike Schwarzschild. This is the "Penrose process". Analogous to Hawking evaporation. Let a particle come in with $\vec{p}_{in}$ and leave with $\vec{p}_{out}$. Then

$$\vec{p}_{in} = \vec{p}_{out} + \vec{p}_{bh}$$

where p_{bh} is 4-momentum transferred to BH.

$$\vec{\xi}.\vec{p}_{in} = \vec{\xi}.\vec{p}_{out} + \vec{\xi}.\vec{p}_{bh} \implies E_{out} = E_{in} - E_{bh}$$

If $E_bh < 0$ then $E_{out} > E_{in}$. But inside ergosphere $\vec{\xi}.\vec{\xi} > 0$ - it is spacelike, so E_{bh} can be negative.

Consider Angular momentum: An observer inside ergosphere has $\vec{U}_{obs} = U^0_{obs}(\vec{\xi} + \Omega_{obs}\vec{\eta})$ with $\Omega_{obs} > 0$. Consider the same process.

$$-\vec{U}_{obs}.p_{bh} \geq 0$$

Energy observed in a local orthonormal frame of the observer must be positive.

$$-U^0_{obs}\underbrace{\vec{\xi}.p_{bh}}_{-E_{bh}}-U^0_{obs}\Omega_{obs}\underbrace{\vec{\eta}.p_{bh}}_{l_{bh}}\geq 0$$

$$E_{bh}\geq \Omega_{obs}l_{bh}$$

Since $E_{bh}<0$, we must have $l_{bh}<0$, i.e. it reduces the angular mom. of BH.
Thus

$$dM\geq \Omega_{obs}dJ$$

Since the max value of $\Omega_{obs}=\Omega_H$ we have

$$dM-\Omega_H dJ\geq 0\implies dA\geq 0$$

————–details

$$A=8\pi M(M+\sqrt{M^2-\frac{J^2}{M^2}})=8\pi(M^2+\sqrt{M^4-J^2})$$

$$\frac{\partial A}{\partial M}=8\pi(2M+\frac{2M^3}{\sqrt{M^4-J^2}})=16\pi M(1+\frac{M}{\sqrt{M^2-a^2}})=8\pi\frac{2Mr_+}{\sqrt{M^2-a^2}}$$

$$\frac{\partial A}{\partial J}=-\frac{8\pi J}{M\sqrt{M^2-\frac{J^2}{M^2}}}=-\frac{8\pi a}{\sqrt{M^2-a^2}}=-\frac{8\pi 2Mr_+\Omega_H}{\sqrt{M^2-a^2}}$$

Thus

$$dA=8\pi\frac{2Mr_+}{\sqrt{M^2-a^2}}(dM-\Omega_H dJ)$$

Finally, we have an identity: If $M_{irr}\equiv(\frac{A}{16\pi})^{\frac{1}{2}}$ then one can show using the expression for area that

$$M^2=M^2_{irr}+\frac{J^2}{4M^2_{irr}}$$

The above argument showed that $dM_{irr}\geq 0$. Thus the smallest possible value for M is M_{irr}, when $J=0$. This is just a Schwarzschild BH. So a Kerr BH can lose ang mom and become a Schw BH.

10 Mathematical Digression I

10.1 Manifolds

(These notes are based on the following books : Isham "Modern Differential Geometry for Physicist", Frankel "geometry of Physics" , Wald "Gravitation", Misner, Thorne and Wheeler "Gravitation".)

Let X be a manifold. First it is a topological space . Further a manifold - that means locally (i.e. any neighbourhood of any of its points) should be an *open* subset (open, because a neighbourhood is an open set) of $\mathbb{R}^n$. Manifold usually means differentiable manifold.

1. Collection of all open sets is the topology of a space X.

   Opens Sets:

   (a) Union of arbitrary collection of open sets is open.

   (b) Intersection of any finite collection of open sets is open

   (c) Φ and X are open sets.

   (d) Example: In a Metric Space, X $\exists$ a metric :

   $$d : X \times X \to \mathbb{R} : d(x, y) = d(y, x)$$

   $$d(x, y) \geq 0; \quad = 0 \ \ iff \quad x = y$$

   $$d(x, y) \leq d(x, z) + d(z, y) \quad \forall x, y, z \in \mathbb{Z}$$

   Define $B_\epsilon(x) = \{y \in X | d(x, y) < \epsilon\}$.
   Then for set $A \subset X$

   i. x is the interior of A if $\exists \epsilon > 0 : B_\epsilon(x) \subset A$
   ii. x is an exterior point of A if $\exists \epsilon > 0 : B_\epsilon(x) \cap A = \Phi$
   iii. x is a boundary point of A if $\forall \epsilon > 0, B_\epsilon(x) \cap A \neq \Phi, \quad B_\epsilon(x) \cap A^c \neq \Phi$

   A set A is open if it contains none of its boundary points.
   So for a metric space this defines a topology.

   (e) A set is closed if its complement is open. Also, a set is closed if it contains all its boundary points.

   (f) A neighbourhood of x is an open set that contains x.

   Once we have this much we have a topological space.

2. (a) $F : M \to N$ is continuous if $F^{-1}V = U$ where V is an open set in N and U is an open set in M.

   (b) If F, F^{-1} are continuous and F is $1:1$ and onto, then it is a homeomorphism i.e. "Preserves Topology".

3. (a) Topological Space X is compact if from *every* covering of X by open sets, one can pick a finite number of open sets that cover X.
   e.g $U_n : \{x | \frac{1}{n} < x < 1\}, n = 1, 2...$ defines a covering of the open interval $(0, 1)$. No *finite* number of U_n can cover the open interval. So this open interval is not compact.

   (b) A topological space is Hausdorff if for each pair of distinct points $p, q, \in X, \quad p \neq q$, one can find open sets O_p, O_q in the topology of X, such that $O_p \cap O_q = \Phi$. $\mathbb{R}^n$ with standard topology is Hausdorff.

   Examples of topological spaces:

   i. $M = \mathbb{R}^n$ covered by a single coordinate system.

ii. M = open ball in $\mathbb{R}^n$: $\{x \in \mathbb{R}^n | ||x-a|| < \epsilon\}$ centred around a. Covered by one coordinate patch.

iii. Closed ball $\mathbb{R}^n$: $\{x \in \mathbb{R}^n | ||x-a|| \leq \epsilon\}$ is NOT a manifold (see defn of manifold below). Because Nbd of a point on the boundary is not homeomorphic to an open subset of $\mathbb{R}^n$ - it only has points on one side - interior. Open subsets of R^n always have $||x-a|| < \epsilon$, which would require interior and exterior.

iv. $M = S^n$ unit sphere in $\mathbb{R}^{n+1}$. eg S^2 is locally $\mathbb{R}^2$. But one needs two coordinate patches.

v. Real Projective space $\mathbb{R}P^n$. Space of all *unoriented* lines L through origin of $\mathbb{R}^{n+1}$.
eg $\mathbb{R}P^2$ can be covered with 3 coordinate patches.
We need to specify a line. The lines go through the origin, so specifying any other point, or specifying the slope will do. So specify ratio of coordinates. eg in a coordinate patch U_x consisting of those lines not lying in the $y-z$ plane. ($x \neq 0$ everywhere on this line - it can be zero at one point).Then $U_1 = \frac{y}{x}$, $U_2 = \frac{z}{x}$ are well defined coordinates in U_x. They give the slope. We can take any point x, y, z on the line, and as long as $x \neq 0$ the point is good. They will all give the same value of U_1, U_2.
Doesn't work for lines in the yz plane because $x = 0$ everywhere! Which is why need more coordinate patches: U_y, U_z that are defined analogously.
There are also "homogeneous" coordinates. Take a pt. x, y, z on the line, not the origin. Then all points $tx, ty, tz, \forall t \in \mathbb{R}$, define the line. t can be positive or negative. If the lines had been directed, then we would have S^2. However, here antipodal points of S^2 define the same line, and so are identified.

(c) The closed interval $[0, 1]$ is **compact**. In fact for $\mathbb{R}^n$, X is compact iff :

i. X is a closed subset of $\mathbb{R}^n$

ii. X is a bounded subset i.e. $\| x \| < c, \quad \forall x \in X$

(d) Continuous image of a compact space is compact.

4. Differentiable Manifold M is a topological space that is locally $\mathbb{R}^n$. In more detail:

(a) We need coordinate charts: A pair (U, ϕ): U is an open set (in M) and $\phi : U \to \mathbb{R}^m$ is a homeomorphism into an open subset of $\mathbb{R}^m$.

(b) If (U_1, ϕ_1) and (U_2, ϕ_2) are two charts and $U_1 \cap U_2 \neq \Phi$, then the overlap function $\phi_2 \circ \phi_1^{-1}$ maps the open subset $\phi_1(U_1 \cap U_2) \subset \mathbb{R}^m \to \phi_2(U_1 \cap U_2) \subset \mathbb{R}^m$.

(c) An atlas is a family of charts (U_i, ϕ_i) with $M = \cup_{i \in I} U_i$ and $\phi_j \circ \phi_i^{-1}$ is a C_∞ map from $\phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$.
If we have such an atlas them M is a differentiable manifold. Thus if U has coordinates x_U^i and V has coordinates x_V^i, then $x_V^i = f_{VU}^i(x_U^1, x_U^2, ..., x_U^m)$ are differentiable functions.

(d) $(\phi^1(P), \phi^2(P), ..., \phi^m(P)) \in \mathbb{R}^m$ are the coordinates of pt. P in M. Call them $x^\mu(P)$, $\mu = 1, 2, .., m$. These are locally functions on M.

(e) If we replace $\mathbb{R}$ by $\mathbb{C}$ and make $\phi_j \circ \phi_i^{-1}$ holomorphic, then it is a complex manifold.

5. (a) A map f from M to N (see Fig. 1) can be realized by mapping $\phi(U) \subset \mathbb{R}^m$ (for $U \subset M$) to $\psi(V) \subset \mathbb{R}^n$ (for $V \subset N$) by the map: $\psi \circ f \circ \phi^{-1} : \phi(U) \subset \mathbb{R}^m \to \psi(V) \subset \mathbb{R}^n$. If $\psi \circ f \circ \phi^{-1}$ is C^r the we an say f is C^r. If C^∞ then "smooth". f is a C^r diffeomorphism if f, f^{-1} are C^r functions. If $\exists$ such a C^r map then the manifolds are C^r-diffeomorphic. Diffeomorphisms for diff manifolds are stronger versions of homeomorphisms for top spaces. More structure than top spaces.

(b) *Diff* (M) - diffeomorphisms of a manifold into itself.

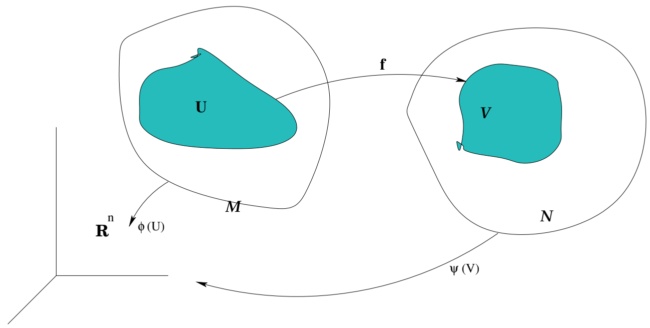


Figure 31: Mapping manifold M to N.

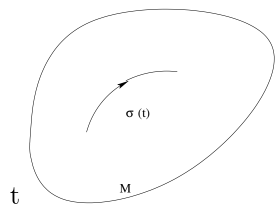


Figure 32: A Curve $\sigma(t)$

10.2 Tangent vectors, Vector fields

1. (a) Define a curve (See Fig. 2):
$$\sigma : (-\epsilon, \epsilon)(\in \mathbb{R}) \to M$$

   (b) σ_1 and σ_2 are tangent at $p \in M$ if
$$\sigma_1(0) = \sigma_2(0) = p$$
and
$$\frac{dx^i(\sigma_1(t))}{dt}|_{t=0} = \frac{dx^i(\sigma_2(t))}{dt}|_{t=0}, \quad i = 1, ..., m$$

   (c) Define the tangent as the equivalence class of curves related by tangency.

   (d) Tangent space T_pM is the set of all tangent vectors at p.

   (e) Tangent bundle is $\cup_{p\in M} T_pM = TM$.

2. So define, for a tangent vector, v, define the real number :
$$v(f) = \frac{df(\sigma(t))}{dt}|_{t=0}$$
This is what is usually called the **directional derivative** of f at p. Here σ is *any* curve with tangent v. Thus v *can be thought of as a differential operator*. Note that we can *define* the vector v by specifying its action on *all* functions f.

3. By chain rule
$$\frac{df}{dt}|_{t=0} = \frac{df(\sigma(t))}{dx^i}\frac{dx^i(\sigma(t))}{dt}|_{t=0}$$

Note: What does $\frac{df(\sigma(t))}{dx^i}$ mean? f is a function on the manifold - depends on p i.e. $f(p)$. So first consider the map $\phi : M \to \mathbb{R}^m$. Then $f \circ \phi^{-1}$ is a function on $\mathbb{R}^m$ - i.e of the coordinates $x^1, x^2 ... x^m$ assigned to the point p, and is being evaluated at the point $\phi(p) = (x^1(p)), x^2(p)...x^m(p))$. So $\frac{df(\sigma(t))}{dx^i} = \frac{\partial}{\partial x^i}[f \circ \phi^{-1}]|_{\phi(p)}$

Thus we can define the vector as an operator:

$$\vec{v} = v^i \frac{\partial}{\partial x^i}$$

We can think of v^i as the components in a basis defined by $\frac{\partial}{\partial x^i}$.

4. Having defined a (tangent) vector at a point we can define a vector field: assignment of a vector at every point on the manifold.

   Thus in the notation above, make v^i functions of $x^1, x^2, .., x^m$:

   $$\vec{v}(x^1, x^2, ..., x^m) = v^i(x^1, x^2, ..., x^m) \frac{\partial}{\partial x^i}$$

   Example: velocity field of water in streamlined flow.

5. Once a vector field is given one can construct the integral curves: curves whose tangent at any point gives the vector field at that point. So think of a curve $\sigma(t)$ that has the required tangent at the point $p = \sigma(t)$. Then the vector field is related to its integral curve by (given a coordinate chart) : $X^i(\sigma(t)) = X^i(p) = \frac{dx^i(\sigma(t))}{dt}$.

   That such an integral curve exists for a given vector field is an important result.

   Example: The streamlines of fluid flow are integral curves of the velocity vector field.

6. A vector field can be said to generate a one parameter group of local diffeomorphisms at a point p. (IMPORTANT for GR because of the fundamental role played by diffeomorphisms)

   (a) Given an open neighbourhood U of p
   (b) and an $\epsilon(\in \mathbb{R}) > 0$
   (c) and a family $\phi_t||t| < \epsilon$ of diffeomorphisms of U onto the open set $\phi_t(U) \subset M$;
   (d) with the property that the map should be smooth in t, q
   (e) and for all t, s, such that $|t|, |s|$ and $|t + s|$ are all less than ϵ,

   $$\phi_s(\phi_t(q)) = \phi_{s+t}(q)$$

   and also $\phi_0(q) = q, \quad \forall q$.
   For streamlines in fluid flow this is a consequence of time independence of the flow - where t, s represent time.
   If these are all true then ϕ is a one-parameter group of diffeomorphisms.

   So through each point on M there is a unique curve defined by this ϕ. The tangent to these curves defines a vector. This assignment of a vector to a point on M defines the vector field. We can call it X^ϕ_q.
   It is defined by its action on functions:

   $$X^\phi_q(f) := \frac{d}{dt} f(\phi_t(q))|_{t=0}$$

10.3 Cotangent Vectors

1. Cotangent vector at a point $p \in M$ is a real linear map from $T_pM \to \mathbb{R}$.

$$k : v \in T_pM \to \mathbb{R}$$

   denoted by $\langle k, v\rangle_p \in \mathbb{R}$.

2. Cotangent space T_p^*M is the set of all such linear maps. Dual vector space to T_pM.

3. Cotangent bundle is the set of all cotangent spaces: $\cup_{p\in M} T_p^* = T^*M$.

4. One form ω on M is a smooth assignment of a cotangent vector ω_p to each point $p \in M$. (or smooth cross section of cotangent bundle). This is counterpart of a vector field.

5. $(\frac{\partial}{\partial x^\mu})_p$ is a basis for T_pM.
   $(dx^\mu)_p$ is a basis for T_p^*M. Can be called a "coordinate basis".

$$\langle dx^\mu, \frac{\partial}{\partial x^\nu}\rangle_p = \delta^\mu_\nu$$

   Sp $k \in T_p^*M = k_\mu(dx^\mu)_p$. k_μ are the components of the cotangent vactor. If we make it a fn of x, $k_\mu(x)$ we get a one form or a "differential one-form".

6. Example: Let $f : M^n \to \mathbb{R}^n$. Then we can define a "differential" df as a map from $T_pM \to \mathbb{R}$ at p by $\langle df, v\rangle_p = df(v_p) = v_p(f)$. v_p is a derivative operator at p - commonly called the directional derivative - the direction being specified by the components v^i. So if $v = v^i \frac{\partial}{\partial x^i}$, then

$$df(v)_p = v^i(p)\frac{\partial f}{\partial x^i}|_p$$

   .
   Special Case: Take $f = x^j$. Then $df(v)_p = v^j_p$.

10.4 Tensors

1. Covariant tensor: multilinear real valued function of vectors. generalization of cotangent.

$$Q : \underbrace{E \times E \times E....... \times E}_{r-times} \to \mathbb{R}$$

   Covariant tensor of rank r. $Q(v_1, v_2, ..., v_r) \in \mathbb{R}$ is linear in each vector v_i. The value does *not* depend on the basis chosen to express v.
   If $V_i = V_i^j \frac{\partial}{\partial x^j}$, then

$$Q(V_1, V_2, ..., V_m) = \sum_{i_1,i_2,..i_r} V_1^{i_1} V_2^{i_2} ....V_r^{i_r} Q(\frac{\partial}{\partial x^{i_1}}, \frac{\partial}{\partial x^{i_2}}, ..., \frac{\partial}{\partial x^{i_r}})$$

$$= \sum_{i_1,i_2,..i_r} V_1^{i_1} V_2^{i_2} ....V_r^{i_r} Q_{i_1 i_2 ... i_r}$$

   As before the components of Q can be made fns of x and we get a tensor field.
   Note: $Q \in \underbrace{E^* \otimes E^* \otimes ..... \otimes E^*}_{r-times}$

2. Contravariant tensors: Acts on cotangent vectors.

$$T : \underbrace{E^* \times E^* \times E^* .... \times E^*}_{r-times} \to \mathbb{R}$$

If k are cotangent vectors, then

$$T(k_1, k_2, ..k_r) = k_{1i_1} k_{2i_2} ... k_{ri_r} \underbrace{T(dx^{i_1}, dx^{i_2}, ..., dx^{i_r})}_{T^{i_1 i_2 .... i_r}}$$

Note: $T \in \underbrace{E \otimes E \otimes E .... \otimes E}_{r-times}$.

3. Mixed tensor

$$W : \underbrace{E \times E \times E .... \times E}_{r-times} \times \underbrace{E^* \times E^* \times E^* .... \times E^*}_{s-times} \to \mathbb{R}$$

Acts on r vectors and s cotangent vectors.

$$\underbrace{W_{i_1 i_2 .... i_r}^{\quad j_1 j_2 .... j_s}}_{W(\frac{\partial}{\partial x^{i_1}}, \frac{\partial}{\partial x^{i_2}}, ..., \frac{\partial}{\partial x^{i_r}}, dx^{j_1}, dx^{j_2}, ..., dx^{j_s})} \qquad V_1^{i_1} V_2^{i_2} .... V_r^{i_r} k_{1j_1} k_{2j_2} ... k_{sj_s}$$

$W \in \underbrace{E \otimes E \otimes E .... \otimes E}_{r-times} \otimes \underbrace{E^* \otimes E^* \otimes E^* .... \otimes E^*}_{s-times}$

4. Example: Metric tensor: $g(u, v) = g_{ij} u^i v^j$ a number. $g(\partial_i, \partial_j) = g_{ij}$

5. Transformation Properties:

$$V = v^i \frac{\partial}{\partial x^i} \to V^i \frac{\partial x'^j}{\partial x^i} \frac{\partial}{\partial x'^j} = V'^j \frac{\partial}{\partial x'^j}$$

$$\omega = \omega_i dx^i \to \omega_i \frac{\partial x^i}{\partial x'^j} dx'^j = \omega'_j dx'^j$$

6. $\omega(V) = \omega_i V^i$ does not depend on coordinate system and is an invariant.

7. Infinitesimal transformations of scalar,vector,form....$x'^i = x^i - \epsilon^i(x)$. **Scalar Field:**

$$\phi'(x') = \phi(x) \implies \phi'(x - \epsilon) = \phi(x) \implies \phi'(x) - \phi(x) = \delta\phi(x) = \epsilon^k \phi(x)_{,k} \tag{98}$$

Vector Field

$$V'^j(x') = V'^j(x - \epsilon) = V^i(x) \frac{\partial x'^j}{\partial x^i} = V^j(x)[\delta_i^j - \epsilon^j_{\ ,i}]$$

So

$$\delta V^j(x) = \epsilon^k V^j_{\ ,k} - \epsilon^j_{\ ,k} V^k \tag{99}$$

Form

$$\omega'_j(x') = \omega_i \frac{\partial x^i}{\partial x'^i} = \omega^i[\delta_i^j + \epsilon^j_{\ ,i}]$$

$$\delta\omega_j(x) = \epsilon^i \omega_{j,i} + \epsilon^j_{\ ,i} \omega^i \tag{100}$$

11 Mathematical Digression II

11.1 Push Forward and Pull Back:

1. **Push Forward of a Vector:** The concept of pushing forward is clear from figure 3. Since we know what it means to map points of a manifold M to points of a manifold N, by a map h, we know what it means to map a curve $\sigma(t)$ to it's image $h(\sigma(t))$. A point P on the curve is mapped to its image point Q on the image curve in N. The curve $h(\sigma(t)$ in N defines a tangent vector X_Q at Q. We define this to be the image under h of the vector X_P. It is called the push forward.

$$h_* : X_P \to X_Q$$

We can do all this in terms of derivatives. Let v be the tangent vector at P in M on curve $\sigma(t)$. More precisely, there is an equivalence class of curves $[\sigma]$ that have the same tangent as our curve and this equivalence class defines v. Then by definition $v(g) = \frac{dg(\sigma(t))}{dt}$ where g is any function on M.

Let f be any function on N. The h_*v, the image of v in N under the map h, acts on f. We will define it by its actions on the arbitrary function f (on N).

$$h_*v(f) \equiv v(f \circ h)$$

(Note: (i) Both RHS and LHS are real numbers. (ii) $f \circ h$ acts on points in M.) This defines the push forward map h_* on a vector.

2. **Push Forward of a Vector Field** A vector field can be pushed forward in the same way. But it *may not be well defined* as Figure 3 shows.

So if $\exists$ a vector field Y on N : $h_*(X_p) = Y_{h(p)} \forall p \in M$ then $Y = h_*(X)$.

3. Example: Consider a manifold M with coordinates x^1, x^2 and another manifold N with coordinates y^1, y^2. Let h be the map $h : M \to N, \quad h(x^1, x^2) = (y^1(x^1, x^2), y^2(x^1, x^2))$. Consider a function $f(y)$ on N. We can pull it back to M as $f \circ h$. This function acts on M. In practical terms $f(y(x) \equiv g(x)$ is what we are calling $f \circ h$. So functions are pulled back.

Vectors are defined by action on functions and are pushed forward. Forms are defined by their action on vectors and are pulled back.

So by definition

$$h_*\vec{v}(f) = (h_*v)^1 \frac{\partial f}{\partial y^1} + (h_*v)^2 \frac{\partial f}{\partial y^2} \tag{101}$$

Consider $\vec{v}(f) = v^1 \frac{\partial f}{\partial x^1} + v^2 \frac{\partial f}{\partial x^2}$. Using chain rule this is

$$v^1(\frac{\partial f}{\partial y^1}\frac{\partial y^1}{\partial x^1} + \frac{\partial f}{\partial y^2}\frac{\partial y^2}{\partial x^1}) + v^2(\frac{\partial f}{\partial y^1}\frac{\partial y^1}{\partial x^2} + \frac{\partial f}{\partial y^2}\frac{\partial y^2}{\partial x^2})$$
$$= (v^1\frac{\partial y^1}{\partial x^1} + v^2\frac{\partial y^1}{\partial x^2})\frac{\partial f}{\partial y^1} + (v^1\frac{\partial y^2}{\partial x^1} + v^2\frac{\partial y^2}{\partial x^2})\frac{\partial f}{\partial y^2}$$

Comparing with (101) we get

$$(h_*v)^1 = (v^1\frac{\partial y^1}{\partial x^1} + v^2\frac{\partial y^1}{\partial x^2}), \qquad (h_*v)^2 = (v^1\frac{\partial y^2}{\partial x^1} + v^2\frac{\partial y^2}{\partial x^2})$$

This is the same rule we got for change of coordinates.

Consider say $y^1 = (x^1)^2, \quad y^2 = (x^2)^2$. This is a 2 to 1 map. Will find that there is a contradiction because x and $-x$ are mapped to the same point and in general the mapped vectors don't match - eg take a constant vector field in M and try.

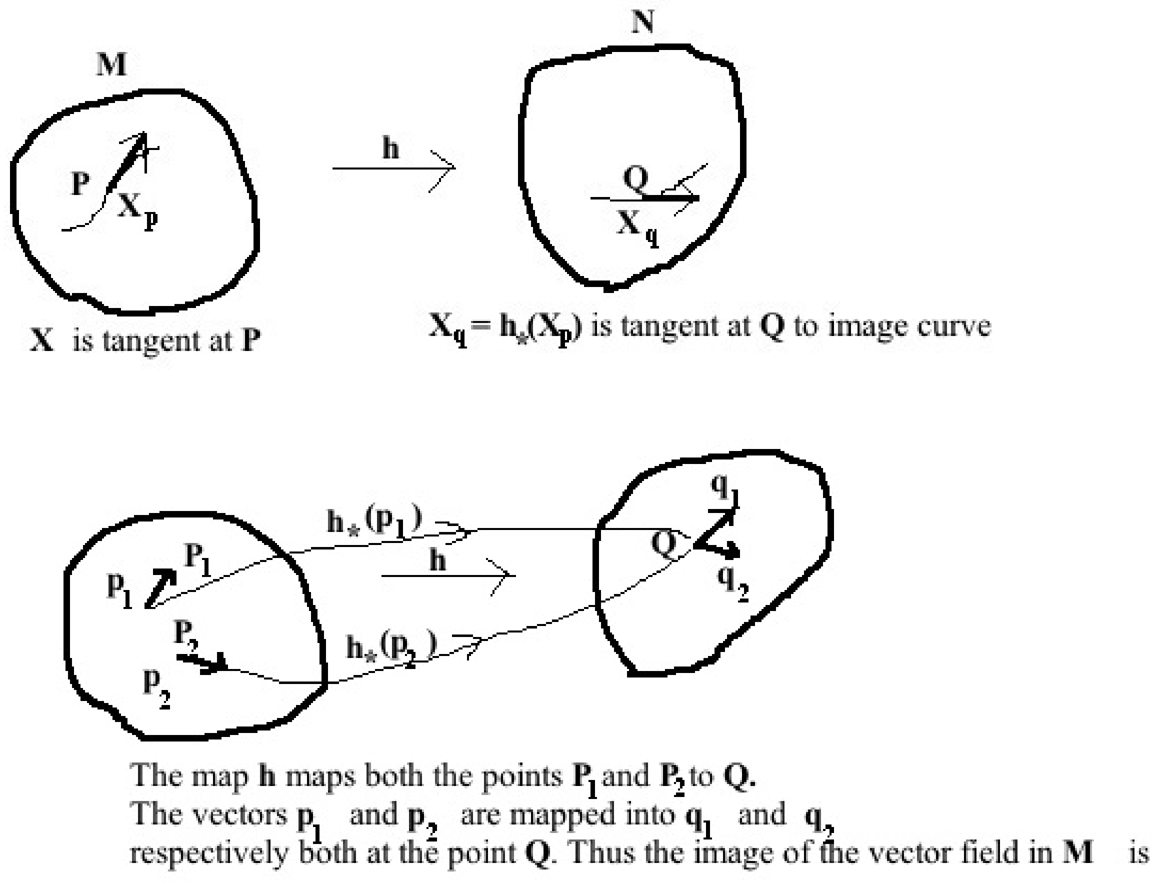


Figure 33: Mapping a vector from M to N.

4. **Pull Back of a Cotangent Vector:**

   Cotangent Vector is defined by action on vectors. So let

$$k : v \in T_pM \to \mathbb{R}, \quad k(v) = \langle k, v\rangle \in \mathbb{R}$$

   Instead of defining the action of the map on T^*_pM we define it on $k \in T^*_qN$ - which is why it is a "pull back". Let $h(p) = q$. Let $h^* : T^*_q(N) \to T^*_p(M)$. Then define $h^*(k)$ which is in $T^*_p(M)$ by ($v \in T_pM$)

$$\langle h^*(k), v\rangle_p = \langle k, h_*v\rangle_{q=h(p)}$$

5. **Pull Back of a form:** This map can directly be extended to a form $\omega \in T^*N$.:

$$\langle h^*\omega, v\rangle_p = \langle \omega, h_*v\rangle_{h(p)}$$

   This is necessarily well defined because two points in N cannot ever be the image of one point in M - if the map h from M to N is to be well defined.

11.2 Lie Derivative

1. Lie Derivative of a one form: Given X a vector field on M, and $\Phi^X_t : M \to M$ is a diffeomorphism of M generated by the vector field, construct the pull back $\Phi^{X*}_t(\omega)$ of a one form $\omega \in T^*_pM$. This is a function of t. So define

$$\mathcal{L}_X\omega = \frac{d}{dt}\Phi^{X*}_t(\omega)|_{t=0}$$

   Note that $\Phi^{X*}_t(\omega)$ is a pull back from the point $\Phi_t(p)$ to $\Phi_0(p) = p$.

2. For a vector field, we have to use a push forward. So we need to push forward from $\Phi_t(p)$ to p - but this requires the map Φ_{-t}. So for a vector field we use the diff generated by Φ^X_{-t}, which gives the push forward $\Phi^X_{-t*}(v_{\Phi_t(p)})$. So

$$\mathcal{L}_Xv = \frac{d}{dt}\Phi^X_{-t*}(v_{\Phi_t(p)})|_{t=0}$$

3. More generally we need to compare geometric "objects" (vectors, forms, tensors...) at $\Phi^X_t(x)$ with those at x. If we need to push forward from $\Phi^X_t(x)$ to x we need to act on the object (eg vector fields) with Φ^X_{-t*}. Whereas if we need to pull back (eg for forms), from $\Phi^X_t(x)$ to x, the action on the form has to be by $\Phi^{X*}_t(x)$. See Figure 4.

$$\mathcal{L}_XO \equiv Lim_{t\to 0}\Big(\frac{\overbrace{O(\Phi^X_t(x))}^{move\ to\ x} - O(x)}{t}\Big)$$

   The operation "move to x" is accomplished by Φ^X_{-t*} or Φ^{X*}_t depending on whether it is a vector or form. Note that $\Phi^X_{-t*} = ((\Phi^X)^{-1})_{t*}$. This will be used in the next para.

4. We can generalize this notion for tensors if we adopt the following definition: Let Φ^* be the "pull back/push forward" operation on a general tensor, for the diffeomorphism Φ. Thus for vectors $\Phi^* = (\Phi^{-1})_*$ whereas for forms it is Φ^*.

   For a general tensor T we can define the map in terms of the action of tensors on vectors and forms:

$$[(\Phi^*T)^{b_1b_2...b_k}_{a_1a_2...a_l}(\mu_1)_{b_1}(\mu_2)_{b_2}...(\mu_k)_{b_k}(v_1)^{a_1}(v_2)^{a_2}...(v_l)^{a_l}]_{\Phi(p)}$$

$$\equiv [T^{b_1b_2...b_k}_{a_1a_2...a_l}(\Phi^*\mu_1)_{b_1}(\Phi^*\mu_2)_{b_2}...(\Phi^*\mu_k)_{b_k}(\Phi^{-1}_*v_1)^{a_1}(\Phi^{-1}_*v_2)^{a_2}...(\Phi^{-1}_*v_l)^{a_l}]_p$$

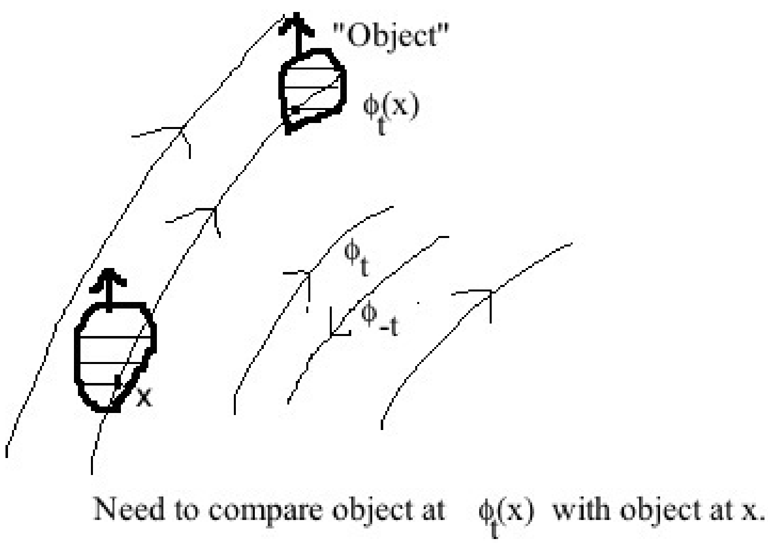


Figure 34: Mapping objects from $\phi(x)$ to x.

The notation used is that of a push forward, i.e. this operation moves the tensor to $\Phi(p)$ from p. Thus when we define the Lie Derivative we have to define it with a "$-t$" as for vectors:

$$\mathcal{L}_X T(x) \equiv Lim_{t\to 0}\Big(\frac{\Phi_{-t}^{X*} T(\Phi_t^X(x)) - T(x)}{t}\Big)$$

5. We can use these definitions and work out some cases explicitly:

   (a) Function:

   $$\mathcal{L}_X f(x) = X(f) = X^i \frac{\partial f(x)}{\partial x^i}$$

   (b) Vector:
   Since the Lie Derivative is a derivative we can use Leibniz rule.

   $$\mathcal{L}_X[v(f)] \equiv Xv(f) = (\mathcal{L}_X[v])(f) + v(\mathcal{L}_X f) = (\mathcal{L}_X[v])(f) + vX(f)$$

   $$\implies (\mathcal{L}_X[v])(f) = [Xv - vX](f) = [X,v](f)$$

   Thus

   $$\mathcal{L}_X[v] = [X,v] \implies \mathcal{L}_X[v]^c = X^b v^c_{,b} - X^c_{,b} v^b$$

   (c) One form: Consider $\langle \mu, v\rangle$, which is a function on the manifold. Then

   $$\mathcal{L}_X\langle \mu, v\rangle = \langle (\mathcal{L}_X[\mu]), v\rangle + \langle \mu, (\mathcal{L}_X[v])\rangle$$

   Choose a basis $\frac{\partial}{\partial x^a}, dx^b$ with $\langle dx^b, \frac{\partial}{\partial x^a}\rangle = \delta^b_a$. Let $\mu = \mu_a dx^a$ and $v = v^a \frac{\partial}{\partial x^a}$. Then we have :

   $$\mathcal{L}_X(\mu_a v^a) = X(\mu_a v^a)$$

   $$\mathcal{L}_X(\mu_a v^a) = \mathcal{L}_X[\mu]_a v^a + \mu_a \mathcal{L}_X[v]^a] = \mathcal{L}_X[\mu]_a v^a + \mu_a[X,v]^a$$

   $$\implies \mathcal{L}_X[\mu]_a v^a = \mathcal{L}_X(\mu_a v^a) - \mu_a[X,v]^a$$

   Let us choose v^a to have only one component $= \delta_{ac}$ in the c direction. It is a constant vector field. (Even if this may not exist globally, this always exists in a neighbourhood of any point, which is all we need to define a derivative at that point.) Then LHS is $\mathcal{L}_X[\mu]_c$. RHS is

   $$\mathcal{L}_X(\mu_c) - \mu_a(-\frac{\partial X^a}{\partial x^c}) = X^b(\mu_c)_{,b} + \frac{\partial X^a}{\partial x^c}\mu_a$$

   Thus

   $$\mathcal{L}_X[\mu]_c = X^b(\mu_c)_{,b} + X^a_{,c}\mu_a$$

6. The rules derived above for vectors and one forms generalizes readily for tensors, with each upper index behaving as a vector and a lower index as a form:

$$\mathcal{L}_X[T]^{a_1 a_2 \dots a_k}_{b_1 b_2 \dots b_l} = X^c (T^{a_1 a_2 \dots a_k}_{b_1 b_2 \dots b_l})_{,c} - X^{a_i}{}_{,c} T^{a_1 a_2 \dots a_{i-1} c a_{i+1} .. a_k}_{b_1 b_2 \dots b_l} + X^c{}_{,b_j} T^{a_1 a_2 \dots a_k}_{b_1 b_2 .. b_{j-1} c b_{j+1} \dots b_l}$$

11.3 Exterior Algebra or Grassmann Algebra

1. Tensor Product: $\alpha \in E^*$, $\beta \in E^*$ (α, β are covectors, which $\implies$ $\alpha, \beta : E \to \mathbb{R}$). Then $\alpha \otimes \beta \in E^* \otimes E^*$. Similarly, if $\alpha \in \underbrace{E^* \otimes E^* \otimes .....E^*}_{p}$ and $\beta \in \underbrace{E^* \otimes E^* \otimes .....E^*}_{q}$ are covariant tensors, then $\alpha \otimes \beta \in \underbrace{E^* \otimes E^* \otimes .....E^*}_{p+q}$

2. If the covariant tensor is completely antisymmetric: i.e. $\alpha(\underbrace{..., v_r, ...v_s, ...}_{p}) = -\alpha(...., v_s, ..., v_r, ...)$, $\forall r, s$, then α is a **p-form**.

   Collection of p-forms is a vector space $\wedge^p E^* = \underbrace{E^* \wedge E^* ..... \wedge E^*}_{p} \subset \otimes^p E^*$. If we define the components of α by $\alpha(\partial_{i_1}, ..., \partial_{i_p}) = \alpha_{i_1.....i_p}$ then

   $$\alpha_{i_1...i_r...i_s....i_p} = -\alpha_{i_1...i_s...i_r....i_p}$$

   This then implies that on an n-dimensional manifold $p \leq n$.

3. wedge product: Defn: $\alpha \wedge \beta \equiv \alpha \otimes \beta - \beta \otimes \alpha$. Thus

   $$\alpha \wedge \beta \ (v, w) = \alpha(v)\beta(w) - \alpha(w)\beta(v)$$

   Similarly we can define wedge product of p-forms and q-forms: Define a permutation P on a form by : $\omega^P(v_1, ...v_n) = \omega(v_{P(i)}, ..., v_{P(n)})$. Then if ω_1, ω_2 are n_1-forms and n_2-forms,

   $$\omega_1 \wedge \omega_2 = \frac{1}{n_1! n_2!} \sum_{perm \ P} (-1)^{deg P} (\omega_1 \otimes \omega_2)^P$$

   Here $deg\ P = +1$ if it is an even permutation, and -1 if odd permutation.

4. Pull back of a wedge product: $h^*(\alpha \wedge \beta) = h^*(\alpha) \wedge h^*(\beta)$.

5. Geometrical Meaning in $\mathbb{R}^n$

   $$dx^i \wedge dx^j \ (v, w) = dx^i(v) dx^j(w) - dx^i(w) dx^j(v)$$

   $dx^i(v) = v^i$. Thus we get

   $$\begin{vmatrix} v^i & w^i \\ v^j & w^j \end{vmatrix}$$

   which is the area of the parallelogram spanned by the projections of v, w onto the $i - j$ plane.

6. Exterior Derivative:

   $$\begin{aligned} d : \wedge^p M^n &\to \wedge^{p+1} M^n : \\ 1) \quad d(\alpha + \beta) &= d\alpha + d\beta && (102) \\ 2) \ d(\alpha^p \wedge \beta^q) &= d\alpha^p \wedge \beta^q + (-1)^p \alpha^p \wedge d\beta^q && (103) \\ 3) \qquad d^2\alpha &= d(d\alpha) = 0 && (104) \\ 4) \quad d\alpha^0 \ is \quad & the \quad usual\ diff\ operator\ acting\ on\ \alpha^0 && (105) \end{aligned}$$

   e.g. $d(a(x)dx^i) = da \wedge dx^i = a_{,j} dx^j \wedge dx^i$.

   Other properties:

(a) $\underbrace{d \underbrace{\omega}_{one-form}}_{2-form} (X,Y) = \underbrace{X(\langle\omega, Y\rangle)}_{fn} - Y(\langle\omega, X\rangle) - \langle\omega, [X,Y]\rangle$.

(b) $d(h^*\omega) = h^*(d\omega)$

(c) $\underbrace{d\omega(X_1, X_2, ...fX_i, ...X_{n+1})}_{n+1-form} = fd\omega(X_1, X_2, ...X_i, ...X_{n+1})$ i.e. depends only on the value of f at that point and not on the gradient of f.

7. Connection to ordinary vector analysis: $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = \frac{\partial f}{\partial x^i}dx^i$. $\vec{\nabla} f$ is a vector with components $(\vec{\nabla} f)^i$ - so it cannot be $\frac{\partial f}{\partial x^i}$. In fact $(\vec{\nabla} f)^i = g^{ij}\frac{\partial f}{\partial x^j}$.

$$df = g_{ij}(\vec{\nabla} f)^j dx^i$$

Uses the metric structure. If the metric is δ^{ij} then the two are equal.

In Cartesian coordinates in $\mathbb{R}^3$ one can associate a vector with a form by:

(a)

$$\alpha = \alpha_i dx^i \leftrightarrow \vec{A} = a^i \frac{\partial}{\partial x^i}$$

$a^j = g^{ji}\alpha_i$ or $\alpha_i = g_{ij}a^j$. Since $g_{ij} = \delta_{ij}$, $a^i = \alpha_i$.

(b) One can calculate $d\alpha$ and $\vec{\nabla} \times \vec{A}$:

$$d\alpha = (\frac{\partial \alpha_1}{\partial y} - \frac{\partial \alpha_2}{\partial x})dy \wedge dx + ... = \frac{\partial \alpha_i}{\partial x^j}dx^j \wedge dx^i$$

$$\vec{\nabla} \times \vec{A} = (\frac{\partial a^1}{\partial y} - \frac{\partial a^2}{\partial x})(-\hat{k}) + ...$$

If $\hat{k}$ is normal to $dx \wedge dy$ then we can associate $\vec{\nabla} \times \vec{A} \leftrightarrow d\alpha$.

(c) Similarly if we have a 2-form $\beta = b_{ij}dx^i \wedge dx^j = \vec{B}.\vec{d}S$. So $b_{12}dx \wedge dy = B^3 dS\hat{k}$. $\therefore b_{12} \leftrightarrow B^3$.

$$\begin{aligned} d\beta &= b_{ij,k}dx^k \wedge dx^i \wedge dx^j \\ &= b_{12,3}dx \wedge dy \wedge dz + ... \end{aligned}$$

In general ${B^k}_{,k} = \vec{\nabla}.\vec{B}$ so $d\beta \leftrightarrow \vec{\nabla}.\vec{B}$. (eg in EM : $F = dA$. Thus $dF = 0$ identically. This is the statement that $divB = 0$ identically.)

$$b_{ij} = \epsilon_{ijk}B^k$$

$$b_{ij,k} = \epsilon_{ijl}{B^l}_{,k} \quad (\therefore k = l)$$

8. Volume form:

Just as $dx \wedge dy(V,W)$ gives area spanned by parallelgm $\vec{V}, \vec{W}$ in xy plane, similarly $dx \wedge dy \wedge dz(\vec{U}, \vec{V}, \vec{W})$ will give a volume of parallelpiped $\vec{U}, \vec{V}, \vec{W}$. Note $dx \wedge dy \wedge dz(\vec{V}, \vec{U}, \vec{W})$ will be of opposite sign.

$$dx \wedge dy \wedge dz(\partial_x, \partial_y, \partial_z) = +1, \quad dx \wedge dy \wedge dz(\partial_y, \partial_x, \partial_z) = -1$$

A change of coordinates gives

$$vol \quad = dx^1 \wedge dx^2 ... \wedge dx^n = \left|\frac{\partial x}{\partial y}\right| dy^1 \wedge dy^2 ... \wedge dy^n$$

Also $g_{ij} = \frac{\partial x^r}{\partial y^i}\frac{\partial x^s}{\partial y^i}\delta_{rs} = M^r{}_i M^r{}_j = (M^T M)_{ij}$. So $\sqrt{Det\ g} = Det\ M$. Thus Volume form $= \underbrace{\sqrt{det\ g}\ dy^1 \wedge dy^2 ... \wedge dy^n}_{vol\ form\ as\ a\ tensor}$. Thus

$$Vol\ form(\partial_{y^{i_1}}, \partial_{y^{i_2}}, ..., \partial_{y^{i_n}}) = o(y)\sqrt{g}\epsilon_{i_1 i_2 ... i_n}$$

Thus RHS is a component of a tensor. ϵ has values ± 1 or 0. More generally specifying a vol form also involves specifying an **orientation,** $o(y)$ of the coordinate system.

11.4 Interior Product

1. i: Given $\vec{V}$ and α^p p-form: $i_V \alpha \sim V^i \alpha_{ijk...}$. Thus

$$\begin{aligned} i_V \alpha^0 &= 0 \\ i_V \alpha^1 &= \alpha(V) \\ i_V \alpha^p(W_2, W_3, ...W_p) &= \alpha^p(V, W_2, W_3, ...W_p) \end{aligned} \quad (106)$$

It is an antiderivation like d:

$$i_V(\alpha^p \wedge \beta^q) = i_V \alpha^p \wedge \beta^q + (-1)^p \alpha^p \wedge i_V \beta^q$$

i_V (Vol form) = n-1 form.

For $n = 3$, $i_B(vol) = \beta^2$; We have seen $d\beta^2 = (\nabla.B)vol^3$

$$\begin{aligned} d(i_B Vol^3) &= d[\sqrt{g}b^1 du^2 \wedge du^3 + \sqrt{g}b^2 du^3 \wedge du^1 + \sqrt{g}b^3 du^1 \wedge du^3] \\ &= (\sqrt{g}b^1)_{,1} du^1 \wedge du^2 \wedge du^3 + ... \\ &= \underbrace{\frac{1}{\sqrt{g}}(\sqrt{g}b^i)_{,i}}_{defn\ of\ div\ \vec{B}}\ vol^3 \end{aligned}$$

2. More manipulations with vector calculus and forms:

   (a) Associate one forms with vectors: $\vec{V}$ and ν^1 by $(\nu^1)_i = g_{ij}V^j$ and similarly $\vec{W}$ and ω^1. Thus we can write:

   $$i_{\vec{V}}\omega^1 = \langle \omega^1_i dx^i, V^j \frac{\partial}{\partial x^j} \rangle = \omega^1_i V^i = g_{ij} W^i V^j = \vec{V}.\vec{W}$$

   (b) Associate $n-1$ forms and vector: $\vec{W}$ and $i_{\vec{W}} vol\ form\ = \omega^{n-1} = W^i \sqrt{g}\epsilon_{ijk..} dx^j \wedge dx^k ...$

   $$\nu^1 \wedge \omega^2 = (\vec{V}.\vec{W})\ vol\ form$$

   (c) Cross Product:

   $$i_{\vec{V}\times\vec{W}}\ vol^3 = \nu^1 \wedge \omega^1$$

   $$-i_{\vec{V}}\omega^2 \leftrightarrow \vec{V} \times \vec{W}$$

   (d) $df \leftrightarrow \nabla f$

(e) If $\alpha^1 \leftrightarrow \vec{A}$ then $d\alpha^1 \leftrightarrow Curl\ \vec{A}$ by the 2-form vector map

$$d\alpha^1 = i_{Curl\ \vec{A}} vol^3$$

And as already seen $d\beta^2 = (div\vec{B})vol^3$.

(f)

$$d(i_{\nabla f}(vol)) = \nabla^2 f(vol)$$

3. **Formulae involving Lie Derivative**

(a) Lie derivative $\mathcal{L}$ is a derivation. So is d. i is an anti derivation. If A, B are both derivations or anti derivations, then to prove $A\alpha^p = B\alpha^p$, you need to show it only on fns and 1-forms. Because any form is of the form $a(x)dx^i \wedge dx^j .... \wedge dx^l$. So you need its action on $a(x)$ and on dx^i.

(b)

$$\mathcal{L}_X \circ d = d \circ \mathcal{L}_X$$

Most commonly used formula. Prove from first principles defn.

(c)

$$\mathcal{L}_X = i_X \circ d + d \circ i_X$$

Proof: On fns: LHS: $\mathcal{L}_X f = X(f)$. RHS: $i_X f = 0$. $i_X df = X(f)$ - by defn.
On forms df: RHS $ddf = 0$. $d \circ [i_X df] = d[X(f)] = d\mathcal{L}_X f = \mathcal{L}_X df$.

(d) On forms:

$$\mathcal{L}_X \circ i_Y - i_Y \circ \mathcal{L}_X = i_{[X,Y]}$$

(e) α^1 is a 1-form and X_x, Y_x are vectors at x. Extend to smooth vector fields (arbitrarily!). Then

$$d\alpha^1(X_x, Y_x) = X_x[\alpha^1(Y)] - Y_x[\alpha^1(X)] - \alpha^1([X, Y])$$

Note some of the operations on the RHS involve derivatives of the vector field, but the LHS only depends on the vector fields at the one point x! Generalization to p-forms:

$$d\alpha^p(Y_0, ..., Y_p) = \sum_r (-1)^r Y_r[\alpha^p(Y_0, ..., \hat{Y}_r, ..., Y_p)] + \sum_{r<s} (-1)^{r+s} \alpha^p([Y_r, Y_s], ...\hat{Y}_r, ..., \hat{Y}_s, ..., Y_p)$$

The hat means drop that vector fld.

(f) Given $p-$ form α^p and vector flds $X, Y_1, ...Y_p$: Leibniz Rule:

$$X[\alpha^p(Y_1, ..., Y_p)] = \mathcal{L}_X[\alpha^p](Y_1, ..., Y_p) + \sum_r \alpha^p(Y_1, ..., (\mathcal{L}_X Y_r), ...Y_p)$$

4. Integration: of a p-form over a p-dimensional submanifold of M^n - we have explicitly indicated the dimension of M to be n- is done in two steps.

Step 1:

First we define integration over over a region $U \subset R^p$: $\alpha^p = a(u)du^1 \wedge du^2 \wedge ... \wedge du^p$.

$$\int_U \alpha^p = \int_U a(u)du^1 \wedge du^2 \wedge ... \wedge du^p \equiv \int_U a(u)du^1 du^2 ... du^p$$

Step 2:

If $F : U \to M^n$ then integration over a p-dim submanifold (if the rank of F is p) is defined by pulling back the form to a p-dimensional submanifold, $U \subset R^p$. Thus $U \subset R^p$ "parametrizes" (in the sense we use in Physics) a p-dim submanifold of M^n:

$$\int_{F(U)\subset M^n} \alpha^p = \int_U F^* \alpha^p = \int_U [F^* \alpha^p](\frac{\partial}{\partial u^1}, \frac{\partial}{\partial u^2}, ..., \frac{\partial}{\partial u^p}) du^1 du^2 ... du^p$$

$$= \int_U \alpha^p (F_* \frac{\partial}{\partial u^1}, F_* \frac{\partial}{\partial u^2} ..... F_* \frac{\partial}{\partial u^p}) du^1 du^2 ... du^p$$

Now $F_* \frac{\partial}{\partial u^1} = \frac{\partial}{\partial x^i}(\frac{\partial x^i}{\partial u^1})$. So $(\alpha^p_{i_1, i_2, ... i_p} \equiv \alpha^p(\frac{\partial}{\partial x^{i_1}}, \frac{\partial}{\partial x^{i_2}}, ..., \frac{\partial}{\partial x^{i_p}}))$

$$= \int_U \alpha^p_{i_1, i_2, ... i_p} \frac{\partial x^{i_1}}{\partial u^1} \frac{\partial x^{i_2}}{\partial u^2} ... \frac{\partial x^{i_p}}{\partial u^p} du^1 du^2 ... du^p$$

which is just the usual chain rule.

eg. If we have a curve $C : \vec{x} = F(t)$ (here t is like u) then

$$\int_C \alpha^1 = \int_C \alpha [F_* \frac{\partial}{\partial t}] dt = \int_C \alpha [\frac{\partial}{\partial x^i} \frac{dx^i}{dt}] dt$$

$$= \int_C \alpha_i \frac{dx^i}{dt} dt$$

usual parametrized line integral.

12 Mathematical Digression III

12.1 Parallel Transport and Covariant Derivative

1. We can define Parallel Transport geometrically using geodesics and define Covariant Derivatives in terms of that. This is the Schild's ladder construction - see MTW. Or we can define covariant derivatives and then define parallel transport in terms of that - see Wald.

2. In flat space parallel transport is very easy. Take a vector. To transport it parallel to itself we complete a parallelogram. Thus let AB be a vector of given length. Suppose we construct a parallelogram by defining the lengths of its diagonals. Then since diagonals bisect we can find the point of intersection, P, using a compass. Since AP=PD and BP=PC, extend the two diagonals to their full length to obtain the points C and D. Join the ends. We have a vector parallel to AB and moved to another point C. Note that once the point C is specified, BC and hence the point P is known. This immediately fixes D.

   Thus for instance a constant vector field is such that the vector at any point is the same as the vector at any other point pararllel transported to this point.

3. In flat space the derivative of a vector field $v^\mu(x)$ in some direction given by vector t^ν is given by $t^\nu \partial_\nu v^\mu(x)$. If the vector field is constant then $t^\nu \partial_\nu v^\mu(x) = 0$ This is true in Cartesian coordinates but not in polar coordinates because the basis vectors change direction. Thus we need to take into account that also.

   As an example let us take a constant vector field $\vec{v}(x) = 2\hat{i}$. It has components $v^1 = 2, v^2 = 0$ and thus satisfies $\partial_a v^b(x) = 0$. Let us describe the same vector field in polar coordinates. let us use the basis vectors
   $$\vec{e}_r = \vec{\partial}_r = cos\theta\hat{i} + sin\theta\hat{j}, \quad \vec{e}_\theta = \vec{\partial}_\theta = -rsin\theta\hat{i} + rcos\theta\hat{j}$$
   They can be inverted to get
   $$\hat{i} = cos\theta\vec{e}_r - \frac{1}{r}sin\theta\vec{e}_\theta, \quad \hat{j} = sin\theta\vec{e}_r + \frac{1}{r}cos\theta e_\theta$$
   Notice that at $\theta = 0$ $\hat{i} = \vec{e}_r$. But at $\theta = 90$, $\hat{i} = -\frac{1}{r}\vec{e}_\theta$.

   So our constant vector has components
   $$v^r = cos\theta, \quad v^\theta = -\frac{1}{r}sin\theta$$
   Thus $\partial_a v^b \neq 0$ and it doesn't look constant at all! We need a definition of constant that works in all coordinates. This motivates the idea of a covariant derivative. We should be able to define a "covariant" derivative $\vec{\nabla}$ and write
   $$\nabla_a v^b = 0$$
   in all coordinates.

   Thus let the vector be written geometrically as $\vec{v}$. Then if is constant then $t^\nu \partial_\nu \vec{v} = 0$. This must be true in any coordinate system. Thus, if $\vec{e}_\mu$ are the basis vectors,
   $$t^\nu(\partial_\nu v^\mu \vec{e}_\mu) = t^\nu[(\partial_\nu v^\mu)\vec{e}_\mu + v^\mu(\partial_\nu \vec{e}_\mu)] = 0$$
   Now let us write
   $$\partial_\nu \vec{e}_\mu = \tilde{\Gamma}^\rho_{\nu\mu}\vec{e}_\rho$$
   to describe how the basis vectors change as we move around. Then our equation becomes
   $$t^\nu[(\partial_\nu v^\rho + \tilde{\Gamma}^\rho_{\nu\mu}v^\mu)]\vec{e}_\rho = 0 \equiv t^\nu(\nabla_\nu v^\rho)\vec{e}_\rho \tag{107}$$
   and this can be used in any coordinate system. We just have to evaluate $\tilde{\Gamma}$. This is called a covariant derivative and now this equation is valid in all coordinate systems in flat space.

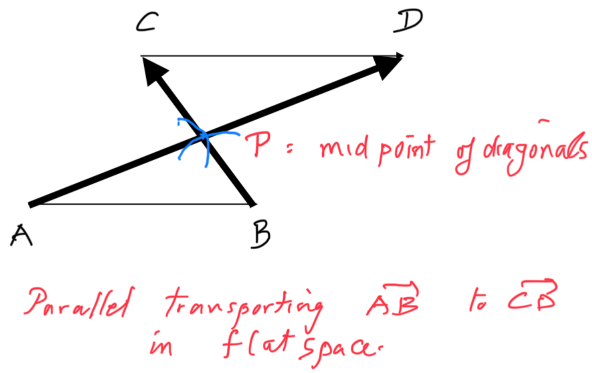


Figure 35: Parallel Transporting a Vector in flat space.

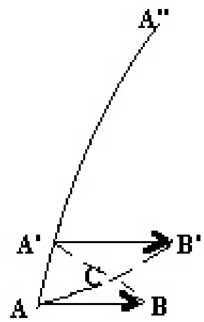


Figure 36: Parallel Transporting a Vector.

4. First do it geometrically: AB is an infinitesimal vector at A that needs to be parallel transported to A'. Draw a geodesic from B to A' and find it's midpoint C (in the sense that the affine parameter of C $\lambda_C = \frac{1}{2}(\lambda_{A'} + \lambda_B)$). Draw a geodesic from A to C. Extend it further by an equal amount to B' (equal amount of affine geodesic). The vector A'B' is parallel transport of AB to the point A'. See fig. 1

   Now that we have defined parallel transporting along a curve, we can define covariant derivative along a curve. So if we have a vector $\vec{V}(t_0)$ such as AB at A, and $\vec{U}$ is a tangent vector to the curve AA" (parametrised by t) at A (which is $t = t_0$), along which to transport the vector, then

$$\nabla_{\vec{U}}\vec{V}(t_0) = \lim_{\epsilon \to 0} \frac{(\vec{V}(t_0 + \epsilon)|_{parallel\ transported\ to\ t_o} - \vec{V}(t_0)}{\epsilon}$$

5. The above defn of covariant derivative gives a very special covariant derivative - the one used in GR. Axiomatically one can define more general covariant derivatives. Imposing an extra requirement narrows it down to this one.

6. If the manifold is embedded in $\mathbb{R}^n$ (for eg M^2 in $\mathbb{R}^3$) then one can give a simple definition of covariant

derivative: it is the **intrinsic derivative**. i.e. Transport the vector along a curve in M^2 such that it is parallel to itself in $\mathbb{R}^3$. Then at the new point project onto the surface (i.e. discard the normal component of the vector). This gives a definition of parallel transport on the surface. It uses the notion of parallel transport in flat $\mathbb{R}^n$ in which the manifold is embedded. This gives the same connection as the Christoffel connection or Levi-Civita connection used in GR.

We can try to do this in curved space. Find an equation that we know is right in a particular coordinate system (i.e. a local inertial frame)then try to define it in a more general frame by finding $\tilde{\Gamma}$.

One equation that we know how to write in curved space in general is the geodesic equation. Thus in RNC the geodesic equation is

$$\frac{d^2y^\mu}{d\tau^2} = 0 = \frac{d}{d\tau}U^\mu = U^\nu \partial_\nu U^\mu$$

This is also true in the local inertial frame of a freely falling observer - clearly in his frame $\frac{d\vec{U}}{d\tau} = 0$ - and he is moving along a geodesic.

What about a general coordinate system?

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\rho}U^\nu U^\rho = 0 = U^\nu \partial_\nu U^\mu + \Gamma^\mu_{\nu\rho}U^\nu U^\rho = U^\nu \nabla_\nu U^\mu \tag{108}$$

The last equality shows that, if we take $\vec{v}$ in (107) to be $\vec{U}$, the required $\tilde{\Gamma}$ that we used to define a covariant derivative is just the Christoffel connection defined earlier. (Technical Point: Christoffel connection, Γ is always defined in a coordinate basis. If $\vec{e}_\mu$ is not a coordinate basis vector (i.e. $\frac{\vec{\partial}}{\partial x^\mu}$), then the $\tilde{\Gamma}$ defined are not given by the same expression.)

The rate of change of U along a curve gives 4-acceleration. In free fall 4-acc is zero. So we can say that the covariant derivative of U along its trajectory is zero. (108) can be written as

$$\nabla_{\vec{U}}\vec{U} = 0$$

Thus we can define the **covariant derivative**

$$\nabla_\rho U^\mu \equiv \frac{\partial U^\mu}{\partial x^\rho} + \Gamma^\mu_{\rho\sigma}U^\sigma \tag{109}$$

An equation written using covariant derivatives is valid in all frames. We will see below that Γ is what is called a connection and has a general axiomatic definition. We have defined a special connection that is used in GR and depends closely on the metric.

7. Axiomatic Definition: Let ∇ be a covariant derivative.

   (a) Linearity:

   $$\nabla(\alpha\mathbf{A} + \beta\mathbf{B}) = \alpha\nabla\mathbf{A} + \beta\nabla\mathbf{B}$$

   $\mathbf{A}, \mathbf{B}$ are tensors and α, β are constants.

   (b) Leibnitz's rule:

   $$\nabla(\mathbf{AB}) = (\nabla\mathbf{A})\mathbf{B} + \mathbf{A}(\nabla\mathbf{B})$$

   (c) Commutativity with contractions:

   $$\nabla(\mathbf{A}^{\cdots\cdot\mathbf{b}}{}_{\cdots\cdots\mathbf{b}}) = (\nabla\mathbf{A})^{\cdots\cdot\mathbf{b}}{}_{\cdots\cdots\mathbf{b}}$$

   (d) Action of vectors on scalars:

   $$t(f) = t^a \nabla_a f$$

(e) Torsion Free: Property that when acting on a function

$$\nabla_a \nabla_b f - \nabla_b \nabla_a f = -T^c{}_{ab} \nabla_c f$$

where $T^a{}_{bc}$ is the torsion tensor, has to be zero in GR. Not necessarily in all theories of gravitation.

8. Condition (d) and (e) can be used to evaluate the commutator of two vector fields

$$[V,W](f) = VW(f) - WV(f) = V^a \nabla_a (W^b \nabla_b f) - W^a \nabla_a (V^b \nabla_b f)$$

Note: f is a scalar function so $W(f) = W^b \partial_b f = W^b \nabla_b f$. $W(f)$ is also a scalar function. So

$$VW(f) = V(W(f)) = V^a \partial_a W(f) = V^a \nabla_a (W^b \nabla_b f) = (V^a \nabla_a W^b) \nabla_b f + V^a W^b \nabla_a \nabla_b f$$

So using Torsion free property

$$[V,W](f) = (V^a \nabla_a W^b - W^a \nabla_a V^b) \nabla_b f$$

So

$$[V,W]^b = (V^a \partial_a W^b - W^a \partial_a V^b) = (V^a \nabla_a W^b - W^a \nabla_a V^b) = (\nabla_V W - \nabla_W V)^b$$

This is also sometimes imposed as a "symmetry" of covariant derivative property. (If $C^a_{bc} = C^a_{cb}$, where C^a_{bc} is defined below, this holds.)

9. Ordinary Derivative satisfies all the axioms and is an example of a Covariant Derivative: Choose a coordinate patch of the manifold with coordinates x^i: so the basis vectors are $\frac{\partial}{\partial x^i}$ and basis one forms dx^i. If T is a tensor with coordinates $T^{a_1 a_2 \ldots a_k}{}_{b_1 b_2 \ldots b_l}$ then the ordinary derivative $\partial_c T$ is the tensor whose components are $\frac{\partial}{\partial x^j} T^{a_1 a_2 \ldots a_k}{}_{b_1 b_2 \ldots b_l}$ in this coordinate basis. This satisfies all the five axioms trivially.

10. A Covariant derivative is not a linear operator: so it is not a tensor: Thus consider $\nabla_{\vec{U}}(f\vec{V})$:

$$\nabla_{\vec{U}}(f\vec{V}) = f\nabla_{\vec{U}}(\vec{V}) + \nabla_{\vec{U}} f \vec{V} = f\nabla_{\vec{U}}(\vec{V}) + U[f]\vec{V}$$

Had it been a linear operator, acting on a tensor fV, it should have been proportional to f without any derivative of f.

It is nevertheless a geometrical object in the sense that $\nabla[\hat{\sigma}, \vec{V}(x), \vec{U}] \equiv \langle \hat{\sigma}, \nabla_{\vec{U}} \vec{V} \rangle$ is a number independent of the coordinate system. Here $\hat{\sigma}$ is a covector (one form). Equivalently $\nabla_{\vec{U}} \vec{V}$ is a vector and is a geometric object.

11. The difference of two cov derivatives is a tensor :

$$(\nabla_{\vec{U}} - \tilde{\nabla}_{\vec{U}})(f\vec{V}) = f(\nabla_{\vec{U}} - \tilde{\nabla}_{\vec{U}})(\vec{V})$$

Thus it is linear. In fact if f is one at the point where the derivative is being evaluated and varies at other points, then

$$(\nabla - \tilde{\nabla})_a (fV^b) = (\nabla - \tilde{\nabla})_a (V^b)$$

Similarly on forms:

$$(\nabla - \tilde{\nabla})_a (f w_b) = (\nabla - \tilde{\nabla})_a (w_b)$$

and this shows that the difference clearly depends only on the value of the vector field V or the form w at that point, and not on its values elsewhere. Thus we can write the linear operator as tensor C^c_{ab}

$$(\tilde{\nabla}_a - \nabla_a) w_b = C^c_{ab} w_c$$

By considering the special case $w_b = (\tilde{\nabla}_b - \nabla_b)f$, since the derivative operator is symmetric on scalar fields, (no torsion property), we see that C^c_{ab} is symmetric in a, b. Thus we can write

$$\nabla_a w_b = \tilde{\nabla}_a w_b - C^c_{ab} w_c \tag{110}$$

Also using $(\tilde{\nabla}_a - \nabla_a)f = 0$, we get $(\tilde{\nabla}_a - \nabla_a)(w_b V^b) = 0$. Thus

$$V^b(\tilde{\nabla}_a - \nabla_a)(w_b) + w_b(\tilde{\nabla}_a - \nabla_a)V^b = V^b C^c_{ab} w_c + w_b(\tilde{\nabla}_a - \nabla_a)V^b$$

Thus

$$(\tilde{\nabla}_a - \nabla_a)V^c + V^b C^c_{ab} = 0$$
$$\nabla_a V^c = \tilde{\nabla}_a V^c + C^c_{ab} V^b \tag{111}$$

Thus by combining (110) and (111) one can get the action on any tensor. If we choose $\tilde{\nabla}_a = \partial_a$ then in GR the C^c_{ab} are denoted by Γ^c_{ab} the Christoffel connection. This is a special choice of covariant derivative as we will see below.

12. We can define parallel transport along t^a by defining that $t^a \nabla_a \vec{V} = 0$. Thus we have a notion of transporting vectors. Thus the tangent vector spaces at two points along a curve can be Identified. This is the role of a **"connection"**. Thus Connections $\leftrightarrow$ Covariant Derivative.

13. Thus there are many covariant derivatives. We pick one that leaves the dot product invariant along a curve when the vectors are parallel transported. i.e. If t^c is the tangent vector along some curve, $t^c \nabla_c (g_{ab} V^a W^b) = 0$ *if* the vectors obey $t^c \nabla_c V^a = 0 = t^c \nabla_c W^a$. Then Leibnitz rule gives $t^c V^a W^b \nabla_c g_{ab} = 0$. This has to hold for all vectors and all curves. This immediately gives $\nabla_c g_{ab} = 0$. We can solve this equation to determine C's.

$$\nabla_a g_{bc} = \tilde{\nabla}_a g_{bc} - C^d_{ab} g_{dc} - C^d_{ac} g_{bd} = 0$$
$$C_{cab} + C_{bac} = \tilde{\nabla}_a g_{bc} \tag{112}$$

Similarly

$$\begin{aligned} C_{bca} + C_{acb} &= \tilde{\nabla}_c g_{ab} \\ C_{abc} + C_{cba} &= \tilde{\nabla}_b g_{ac} \end{aligned} \tag{113}$$

Adding the two (113) and subtracting (112) we get

$$2C_{acb} = \tilde{\nabla}_b g_{ac} + \tilde{\nabla}_c g_{ab} - \tilde{\nabla}_a g_{bc}$$

and so

$$C^a_{bc} = \frac{1}{2} g^{ad} [\tilde{\nabla}_b g_{dc} + \tilde{\nabla}_c g_{db} - \tilde{\nabla}_d g_{bc}]$$

This is the connection that is "compatible with the metric". The special choice $\tilde{\nabla}_a = \partial_a$ gives us the usual GR expression for the Christoffel symbol

$$\Gamma^a_{bc} = \frac{1}{2} g^{ad} [\partial_b g_{dc} + \partial_c g_{db} - \partial_d g_{bc}] \tag{114}$$

14. Change of basis: $(\nabla_\mu - \tilde{\nabla}_\mu) f^\alpha$ is a tensor. So

$$(\nabla_\mu - \tilde{\nabla}_\mu) f^\alpha = \frac{\partial x^\alpha}{\partial x'^\beta} \frac{\partial x'^\lambda}{\partial x^\mu} [(\nabla'_\lambda - \tilde{\nabla}'_\lambda)] f'^\beta$$

So let us say that $\tilde{\nabla}'_\lambda = \frac{\partial}{\partial x'^\lambda}$. The question is what is $\tilde{\nabla}_\mu$? RHS is

$$\frac{\partial x^\alpha}{\partial x'^\beta}\frac{\partial x'^\lambda}{\partial x^\mu}[\nabla'_\lambda - \frac{\partial}{\partial x'^\lambda}][\frac{\partial x'^\beta}{\partial x^\gamma}f^\gamma]$$

$$= \frac{\partial x^\alpha}{\partial x'^\beta}\frac{\partial x'^\lambda}{\partial x^\mu}[\nabla'_\lambda f'^\beta] - \frac{\partial x^\alpha}{\partial x'^\beta}\frac{\partial x'^\lambda}{\partial x^\mu}\frac{\partial}{\partial x'^\lambda}[\frac{\partial x'^\beta}{\partial x^\gamma}f^\gamma]$$

$$= \frac{\partial x^\alpha}{\partial x'^\beta}\frac{\partial x'^\lambda}{\partial x^\mu}[\nabla'_\lambda f'^\beta] - \frac{\partial x^\alpha}{\partial x'^\beta}\frac{\partial}{\partial x^\mu}[\frac{\partial x'^\beta}{\partial x^\gamma}f^\gamma]$$

$$= [\nabla_\mu f^\alpha] - \frac{\partial x^\alpha}{\partial x'^\beta}\frac{\partial^2 x'^\beta}{\partial x^\mu \partial x^\gamma}f^\gamma - \frac{\partial f^\alpha}{\partial x^\mu}$$

Thus we conclude that

$$\tilde{\nabla}_\mu f^\alpha = \frac{\partial x^\alpha}{\partial x'^\beta}\frac{\partial^2 x'^\beta}{\partial x^\mu \partial x^\gamma}f^\gamma + \frac{\partial f^\alpha}{\partial x^\mu}$$

Thus the derivative is not mapped to a derivative.

The extra term is usually absorbed into the connection:

$$= (\nabla_\mu f^\alpha - \tilde{\nabla}_\mu f^\alpha) = C^\alpha_{\mu\gamma} f^\gamma = \frac{\partial x^\alpha}{\partial x'^\beta}\frac{\partial x'^\lambda}{\partial x^\mu} C'^{\beta'}_{\lambda'\delta'}\frac{\partial x'^{\delta'}}{\partial x^\gamma}f^\gamma$$

$$\therefore \nabla_\mu f^\alpha = C^\alpha_{\mu\gamma} f^\gamma + \frac{\partial x^\alpha}{\partial x'^{\beta'}}\frac{\partial^2 x'^\beta}{\partial x^\mu \partial x^\gamma}f^\gamma + \frac{\partial f^\alpha}{\partial x^\mu}$$

Thus

$$\nabla_\mu f^\alpha = \Gamma^\alpha_{\mu\gamma} f^\gamma + \frac{\partial f^\alpha}{\partial x^\mu} = \frac{\partial x^\alpha}{\partial x'^\beta}\frac{\partial x'^\lambda}{\partial x^\mu} C'^{\beta'}_{\lambda'\delta'}\frac{\partial x'^{\delta'}}{\partial x^\gamma}f^\gamma + \frac{\partial x^\alpha}{\partial x'^{\beta'}}\frac{\partial^2 x'^\beta}{\partial x^\mu \partial x^\gamma}f^\gamma + \frac{\partial f^\alpha}{\partial x^\mu} \tag{115}$$

We have used Γ to denote the conventional Christoffel connection, and the notation C to denote the tensor definition. In the conventional definition, since we always separate out the ordinary derivative the Christoffel connection is forced to have non tensorial pieces in its transformation. This eqn defines the conventional transformation of the Christoffel connection. Symbolically and as a useful mnemonic one can write

$$\Gamma(new) = Tensor \;\; rotate \;\; \Gamma(old) + rotate \; \frac{\partial^2 Y \;(old)}{\partial Y(new)\partial Y(new)}$$

or one can derive by using $\partial M^{-1} = -M^{-1}\partial M M^{-1}$ where $M^\mu_\nu = \frac{\partial Y^\mu \;(new)}{\partial Y^\nu \;(old)}$,

$$\Gamma(new) = Tensor \;\; rotate \;\; \Gamma(old) - rotate \; \frac{\partial^2 Y \;(new)}{\partial Y(old)\partial Y(old)}$$

12.2 Covariant Exterior Derivative and the Forms

We define covariant derivatives and connections using the language of forms and Cartan's Covariant Exterior Derivatives.

1. Let $\vec{e}_\alpha$ be a basis for the tangent vector space. (Note that the vector index is suppressed - it is shown as an arrow. α is a *label* for the vector. Thus $\vec{e}_1, \vec{e}_2$.. are different vectors and may be basis vectors. In a coordinate basis $\vec{e}_\mu = (\frac{\vec{\partial}}{\partial x^\mu})$.) $\nabla_\mu \vec{e}_\alpha$ is also a vector and can be expanded as

$$\nabla_\mu \vec{e}_\alpha = C^\beta_{\mu\alpha}\vec{e}_\beta \tag{116}$$

This is true in any basis. Then acting on a vector $\mathbf{f} = f^\alpha \vec{e}_\alpha$,

$$\nabla_\mu [f^\alpha \vec{e}_\alpha] = \partial_\mu f^\alpha \vec{e}_\alpha + f^\alpha \nabla_\mu \vec{e}_\alpha$$

$$= \partial_\mu f^\alpha \vec{e}_\alpha + f^\alpha C^\beta_{\mu\alpha} \vec{e}_\beta$$

$$= \partial_\mu f^\beta \vec{e}_\beta + f^\alpha C^\beta_{\mu\alpha} \vec{e}_\beta = (\nabla_\mu f^\beta) \vec{e}_\beta$$

The expression in brackets is the covariant derivative described in the last subsection if we take a coordinate basis set for vectors $\vec{e}_\mu = (\frac{\vec{\partial}}{\partial x^\mu})$. (In this case $C \equiv \Gamma$. We reserve the symbol $\Gamma^\mu_{\nu\rho}$ for the usual Christoffel connection.) The present definition is more general. The index μ refers to a coordinate, but α, β need not.

2. We can also define

$$\nabla_{\vec{f}} \vec{e}_\alpha = f^\gamma \nabla_\gamma \vec{e}_\alpha = f^\gamma C^\beta_{\gamma\alpha} \vec{e}_\beta$$

Here α, β, γ can all be a general basis - not necessarily coordinate bases. Thus

$$\nabla_{\vec{e}_\gamma} \vec{e}_\alpha \equiv \nabla_\gamma \vec{e}_\alpha = C^\beta_{\gamma\alpha} \vec{e}_\beta$$

3. Change of basis vectors:

$$\nabla_\beta \vec{e}_\alpha = C^\gamma_{\beta\alpha} \vec{e}_\gamma$$

Let $\vec{e}_\alpha = e^a_\alpha \vec{e}_a$. (For instance if $\vec{e}_a$ refers to an orthonormal basis, that e^α_a is a "vierbein".) Then

$$\nabla_\beta \vec{e}_\alpha = e^a_\beta \nabla_a (e^b_\alpha \vec{e}_b)$$

$$= e^a_\beta [(\partial_a e^c_\alpha) \vec{e}_c + e^b_\alpha C'^c_{ab} \vec{e}_c]$$

$$\therefore [(\partial_a e^c_\alpha) \vec{e}_c + e^b_\alpha C'^c_{ab} \vec{e}_c] = [e^\beta_a C^\gamma_{\beta\alpha} e^c_\gamma] \vec{e}_c$$

$$[(e^\alpha_b \partial_a e^c_\alpha) + C'^c_{ab}] = e^\alpha_b e^\beta_a C^\gamma_{\beta\alpha} e^c_\gamma$$

$$\therefore C'^c_{ab} = e^\alpha_b e^\beta_a C^\gamma_{\beta\alpha} e^c_\gamma - (e^\alpha_b \partial_a e^c_\alpha) \tag{117}$$

(119) has the form of the gauge transformation of a gauge field. The second term on the RHS is of the form $g^{-1}\partial_a g$. The first term is a rotation of each index by the matrix g or g^{-1} depending on its location and is a tensor transformation.

4. In the case that both are coordinate bases $e^a_\beta = \frac{\partial x^a}{\partial x^\beta}$. In that case the expression $\partial_a e^c_\alpha$ can be written as (Using $\partial_a M^{-1} = -M^{-1} \partial_a M M^{-1}$ for a matrix M)

$$\partial_a [(\frac{\partial x^\alpha}{\partial x^c})^{-1}]^c_\alpha = -[\frac{\partial x^c}{\partial x^\gamma} \partial_a (\frac{\partial x^\gamma}{\partial x^d}) \frac{\partial x^d}{\partial x^\alpha}]$$

So

$$e^\alpha_b \partial_a e^c_\alpha = -(\frac{\partial x^\alpha}{\partial x^b})[\frac{\partial x^c}{\partial x^\gamma} \partial_a (\frac{\partial x^\gamma}{\partial x^d}) \frac{\partial x^d}{\partial x^\alpha}]$$

$$= -\frac{\partial x^c}{\partial x^\gamma} \frac{\partial^2 x^\gamma}{\partial x^a \partial x^b}$$

Thus (setting $C = \Gamma$, since we are in a coordinate basis)

$$\Gamma'^c_{ab} = \frac{\partial x^c}{\partial x^\gamma} \frac{\partial^2 x^\gamma}{\partial x^a \partial x^b} + \frac{\partial x^c}{\partial x^\gamma} \frac{\partial x^\alpha}{\partial x^a} \frac{\partial x^\beta}{\partial x^b} \Gamma^\gamma_{\alpha\beta}$$

This agrees with (115).

5. Since

$$[\vec{e}_\alpha, \vec{e}_\beta] = \nabla_\alpha e_\beta - \nabla_\beta e_\alpha = (C^\gamma_{\alpha\beta} - C^\gamma_{\beta_\alpha})\vec{e}_\gamma$$

We see that $C^\gamma_{\alpha\beta}$ is not necessarily symmetric in α, β, unless $[\vec{e}_\alpha, \vec{e}_\beta] = 0$. This is necessarily true in a coordinate basis, where $\vec{e}_\alpha = \frac{\partial}{\partial x^\alpha}$.

6. **Postpone 6-10 till later.**

We now introduce Cartan's **Covariant exterior derivative d**. We think of everything as forms. Thus a vector is a vector valued zero form. So $\vec{e}_\alpha$ is a vector valued zero form. Vectors are dual to forms. So if $\vec{e}_\alpha, \alpha = 1, 2, 3...$ is a basis of vectors, then the dual basis of forms is $\underline{\omega}^\beta, \beta = 1, 2, 3....$ Also their inner product by definition is

$$\langle \underline{\omega}^\beta, \vec{e}_\alpha \rangle = \delta^\beta_\alpha \tag{118}$$

While $\underline{\omega}^\beta$ is an ordinary form, $\mathbf{d}\vec{e}_\alpha$ is a vector valued one form. Thus the inner product with a vector, instead of giving a number as in (118), will give a vector. To emphasize the similarity with Yang-Mills, we will use $\vec{e}_a$, instead of $\vec{e}_\alpha$ for the vector basis and $\underline{\omega}^\beta$ for the form basis - not necessary the dual basis to the vector basis. They could refer to the same basis, but need not in general. So if $\vec{e}_\alpha$ is the basis dual to the form, $\langle \mathbf{d}\vec{e}_a, \vec{e}_\beta \rangle$ is a vector and can be expanded as

$$\langle \mathbf{d}\vec{e}_a, \vec{e}_\beta \rangle = C^c_{\beta a}\vec{e}_c$$

We can also write $\mathbf{d}\vec{e}_a \equiv \underline{\vec{\omega}}_a$ where $\underline{\vec{\omega}}_a$ is a vector valued one form. The index a is a label for this vector valued one form. Thus $\underline{\vec{\omega}}_a = \underline{\omega}^c{}_a \vec{e}_c$. ($\underline{\omega}^c{}_a$ can be called a $\binom{1}{1}$ tensor in GR. Here the index c is the vector index and β is the form index, and in GR they would refer to the same space. The additional index a is just a label here. Again in GR all these belong to the same space.) Thus

$$\mathbf{d}\vec{e}_a = \underline{\omega}^c{}_a \vec{e}_c$$

$$\underline{\omega}^c{}_a = C^c_{\beta a}\underline{dx}^\beta = \omega_\beta{}^c{}_a \underline{\omega}^\beta$$

(Note that if we think of a, c as internal indices then $\omega_\beta{}^c{}_a$ is like a Yang-Mills gauge field.) and thus

$$\mathbf{d}\vec{e}_a = \underline{\vec{\omega}}_a = \omega^c_{\beta a}\underline{\omega}^\beta \otimes \vec{e}_c = C^c_{\beta a}\underline{dx}^\beta \otimes \vec{e}_c \tag{119}$$

where $\underline{\omega}^\beta$ is a basis of one forms and $\underline{dx}^\beta$ is a coordinate basis of one forms. This is the form version of (116).

$\mathbf{d} = \underline{\omega}^\beta \nabla_\beta$. If we combine this with (116) we get (119).

7. **Compatibility with metric:** Just as we specified a cov derivative by requiring that inner products remain invariant, here we require that:

$$\mathbf{d}(\vec{e}_\mu.\vec{e}_\nu) = \mathbf{d}\vec{e}_\mu.\vec{e}_\nu + \vec{e}_\mu.\mathbf{d}\vec{e}_\nu$$

Note that LHS is a scalar and so it is an ordinary (exterior) derivative. On the RHS we have it acting on a vector so it is a covariant (exterior) derivative. This equation is not automatically true. It imposes a condition on the connection.

$$LHS = \mathbf{d}g_{\mu\nu} = RHS = \underline{\omega}^\rho{}_\mu \vec{e}_\rho.\vec{e}_\nu + \vec{e}_\mu.\underline{\omega}^\rho{}_\nu \vec{e}_\rho = \underline{\omega}^\rho{}_\mu g_{\rho\nu} + \underline{\omega}^\rho{}_\nu g_{\mu\rho} = \underline{\omega}_{\mu\nu} + \underline{\omega}_{\nu\mu}$$

(Note that in $\underline{\omega}^\rho{}_\mu$, ρ is the index associated with a vector because it is a vector valued one form, whereas μ is just a label for the form. But in this case they both refer to the same vector space.)

8. **Cartan's Structural equations:**

   The object $\vec{e}_\mu \otimes \underline{\omega}^\mu$ is $\binom{1}{1}$ tensor. But it is always equal to the identity:

$$\vec{e}_\mu \otimes \underline{\omega}^\mu(\underline{\omega}^\beta, \vec{e}_\alpha) = \delta^\beta_\alpha$$

   So $\mathbf{d}(\vec{e}_\mu \otimes \underline{\omega}^\mu) = 0$ must give some constraint:

$$\underbrace{\mathbf{d}\vec{e}_\mu \wedge \underline{\omega}^\mu}_{2-form} + \vec{e}_\mu \underbrace{\mathbf{d}\underline{\omega}^\mu}_{2-form} = \vec{e}_\rho \underbrace{\underline{\omega}^\rho{}_\mu}_{1-form} \wedge \underline{\omega}^\mu + \mathbf{d}\underline{\omega}^\mu \vec{e}_\mu = 0$$

   Thus

$$\mathbf{d}\underline{\omega}^\mu + \underline{\omega}^\mu{}_\rho \wedge \underline{\omega}^\rho = d\underline{\omega}^\mu + \underline{\omega}^\mu{}_\rho \wedge \underline{\omega}^\rho = 0$$

   Since $\underline{\omega}^\mu$ is a one form (not vector valued) we have an ordinary exterior derivative. These are **Cartan's Structural equations**. They can be used to calculate $\underline{\omega}^\mu{}_\nu$.

9. How does one calculate $\underline{\omega}^a{}_b$?

   One method is by trial and error. This is often the simplest.

   The systematic method is to work in an orthonormal basis and use the following construction:

$$\langle d\underline{\omega}^c, \vec{e}_a, \vec{e}_b \rangle = \underbrace{\partial_{\vec{e}_\alpha} \underbrace{\langle \underline{\omega}^c, \vec{e}_b \rangle}_{\delta^c_b}}_{0} - \partial_{\vec{e}_b} \langle \underline{\omega}^c, \vec{e}_a \rangle - \langle \underline{\omega}^c, [\vec{e}_a, \vec{e}_b] \rangle$$

$$= -\langle \underline{\omega}^c, [\vec{e}_a, \vec{e}_b] \rangle = -2A^c_{ab}$$

   A^c_{ab} is the antisymmetric part of the Christoffel connection in a non coordinate basis.

$$\therefore d\underline{\omega}^c = -A^c_{ab}\underline{\omega}^a \wedge \underline{\omega}^b = -\underline{\omega}^c{}_b \underline{\omega}^b$$

$$\implies A_{cab}\underline{\omega}^a \wedge \underline{\omega}^b = \underline{\omega}_{cb}\underline{\omega}^b$$

   Note that since $d(g_{ab}) = \underline{\omega}_{ab} + \underline{\omega}_{ba}$, if $g_{ab} = \eta_{ab}$ then $\underline{\omega}_{ab}$is antisymmetric in a, b. Thus let us assume that our basis is orthonormal.

$$\underline{\omega}_{cb}\underline{\omega}^b = \omega_{acb}\underline{\omega}^a \wedge \underline{\omega}^b$$

$$(A_{cab} - A_{bac} - A_{abc})\underline{\omega}^a \wedge \underline{\omega}^b = \omega_{acb}\underline{\omega}^a \wedge \underline{\omega}^b$$

   where we have added the last two terms. Being symmetric in a, b it is automatically zero due to the wedge product. The result is that the tensor in brackets is antisymmetric in b, c and therefore can be equated to the RHS. Thus

$$A_{cab} - A_{bac} - A_{abc} = \omega_{acb}$$

   This construction requires that we work in an orthonormal basis.

10. Example of a calculation:

$$ds^2 = R^2(d\theta^2 + sin^2\theta d\phi^2) = \omega^{\hat{\theta}} \otimes \omega^{\hat{\theta}} + \omega^{\hat{\phi}} \otimes \omega^{\hat{\phi}}$$

   is the metric on a 2-sphere S^2.We have defined $\underline{\omega}^{\hat{\theta}} = Rd\theta$ and $\underline{\omega}^{\hat{\phi}} = R\ sin\theta\ d\phi$ as orthonormal basis forms and $\vec{e}_{\hat{\theta}} = \frac{1}{R}\partial_\theta$ and $\vec{e}_{\hat{\phi}} = \frac{1}{R\ sin\theta}\partial_\phi$ as dual basis vectors.

   Since the metric is orthonormal $\underline{\omega}_{\hat{\theta}\hat{\phi}} = -\underline{\omega}_{\hat{\phi}\hat{\theta}}$ is the only non zero connection one form.

   Let us write the Cartan Structure equations:

$$d\omega^{\hat{\theta}} + \omega^{\hat{\theta}}{}_{\hat{\phi}} \wedge \omega^{\hat{\phi}} = 0$$

$$d\omega^{\hat{\phi}} + \omega^{\hat{\phi}}{}_{\hat{\theta}} \wedge \omega^{\hat{\theta}} = 0$$

$$d\omega^{\hat{\theta}} = d(Rd\theta) = 0$$

and

$$d\omega^{\hat{\phi}} = d(R\ sin\theta\ d\phi) = R\ cos\theta\ d\theta \wedge d\phi$$

$$= \frac{1}{R} cot\theta\ \omega^{\hat{\theta}} \wedge \omega^{\hat{\phi}} = -\frac{1}{R} cot\theta\ \omega^{\hat{\phi}} \wedge \omega^{\hat{\theta}}$$

Thus

$$\underline{\omega}^{\hat{\phi}}{}_{\hat{\theta}} = \frac{1}{R} cot\theta\ \underline{\omega}^{\hat{\phi}} = -\underline{\omega}^{\hat{\theta}}{}_{\hat{\phi}} \tag{120}$$

We also get $\omega_{\hat{\phi}}{}^{\hat{\phi}}{}_{\hat{\theta}} = -\omega_{\hat{\phi}}{}^{\hat{\theta}}{}_{\hat{\phi}} = \frac{1}{R} cot\theta$. Also $\omega_{\hat{\theta}}{}^{\hat{\theta}}{}_{\hat{\phi}} = -\omega_{\hat{\theta}}{}^{\hat{\phi}}{}_{\hat{\theta}} = 0$.

12.3 Curvature

1. When the gravitational field is non-uniform (i.e.g is position dependent) then finite size objects are squeezed or stretched - tidal forces. This is curvature. In Newtonian gravity we can measure this by observing how neighbouring geodesics deviate from each other. If $\phi(x)$ is the Newtonian gravitational potential (eg $\frac{GM}{r}$):

$$\frac{d^2x^i}{dt^2} = -\delta^{ij}\frac{\partial\phi(x)}{\partial x^j}$$

$$\frac{d^2(x^i+n^i)}{dt^2} = -\delta^{ij}\frac{\partial\phi(x+n)}{\partial x^j} = -\delta^{ij}[\frac{\partial\phi(x)}{\partial x^j} + n^k\frac{\partial^2\phi(x)}{\partial x^j\partial x^k}]$$

Thus

$$\frac{d^2n^i}{dt^2} = -\delta^{ij} \underbrace{[\frac{\partial^2\phi(x)}{\partial x^j\partial x^k}]}_{curvature\ in\ Newtonian\ gravity} n^k = \text{``}R^i_k\text{''}n^k$$

Then EOM for gravitational field is

$$R^i_i = \nabla^2\phi = 4\pi GM \tag{121}$$

2. Note that $\frac{dn^i}{d\lambda}$ being non zero is not enough. Two particles in fat space going in a straight line, but not parallel, will have $\frac{dn^i}{d\lambda} \neq 0$.

3. Consider a family of geodesics $\gamma_s(\lambda)$ where the geodesics are labeled by the parameter s. The map $(s,\lambda) \to \gamma_s(\lambda)$ is smooth and one-one and has an inverse. Σ is the two dimensional submanifold spanned by the curves $\gamma_s(\lambda)$. Then s, λ can be chosen as coordinates on Σ.

   $\vec{U} \equiv \frac{\vec{\partial}}{\partial\lambda}$ is a vector field of tangents to geodesics. It satisfies $\nabla_{\vec{U}}\vec{U} = U^a\nabla_a\vec{U} = 0$.

   $\vec{n} \equiv \frac{\vec{\partial}}{\partial s}$ is a vector field that gives the displacement to a nearby geodesic. Note that if we change $(\lambda, s) \to (\lambda, s) = (\lambda' + c(s'), s')$ then

$$\vec{n}' = \frac{\partial}{\partial s'} = \frac{\partial\lambda}{\partial s'}\frac{\partial}{\partial\lambda} + \frac{\partial s}{\partial s'}\frac{\partial}{\partial s} = c'(s')\frac{\partial}{\partial\lambda} + \frac{\partial}{\partial s} = c'(s')\vec{U} + \vec{n}$$

   Thus by changing the parametrisation we can add to n an amount proportional to U. So $\vec{n}.\vec{U} \to c'(s')\vec{U}.\vec{U} + \vec{n}.\vec{U}$. Thus if we consider a curve $C(s)$ which is the set of points satisfying $\lambda = 0$, all along this curve we can choose $\vec{n}.\vec{U} = 0$ by choosing $c(s)$ suitably. Furthermore if we choose the Christoffel connection, $\vec{U}.\vec{n}$ is preserved along a geodesic, and thus remains zero everywhere in Σ.

   So if we have a collection of geodesics then $\nabla_{\vec{U}}\vec{n}$ is the rate of change of geodesic separation along the particle path. What we need is $\nabla_{\vec{U}}(\nabla_{\vec{U}}\vec{n})$. This is a measure of tidal forces - curvature and a hence a physical gravitational field. See fig 12.

   $\nabla_U U = 0$ along a geodesic. So $\nabla_n\nabla_U U = 0$. Thus

$$\nabla_U\nabla_n U + [\nabla_n, \nabla_U]U = 0$$

   $\frac{\vec{\partial}}{\partial s} \sim \vec{n}$ and $\frac{\vec{\partial}}{\partial\lambda} \sim \vec{U}$, so $[\vec{n}, \vec{U}] = 0$. So $\nabla_n U = \nabla_U n$.

$$\nabla_U\nabla_U n + [\nabla_n, \nabla_U]U = 0$$

$$\mathbb{R}(\_, \vec{U}, \vec{n}, \vec{U}) \equiv \mathbb{R}(n, U)U \equiv [\nabla_n, \nabla_U]U = -\nabla_U\nabla_U n$$

   is the equation of geodesic deviation. We will see below that the curvature tensor so defined is a linear operator i.e. a tensor.

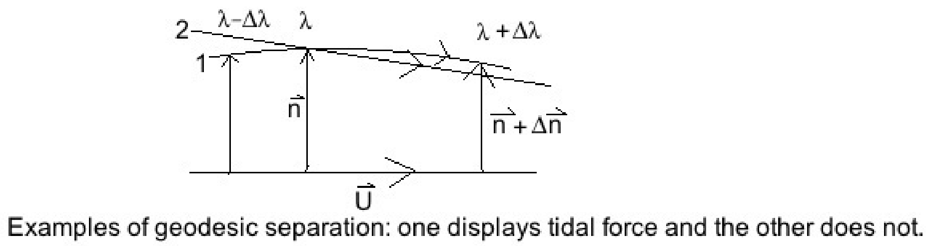


Figure 37: Tidal force and geodesic deviation

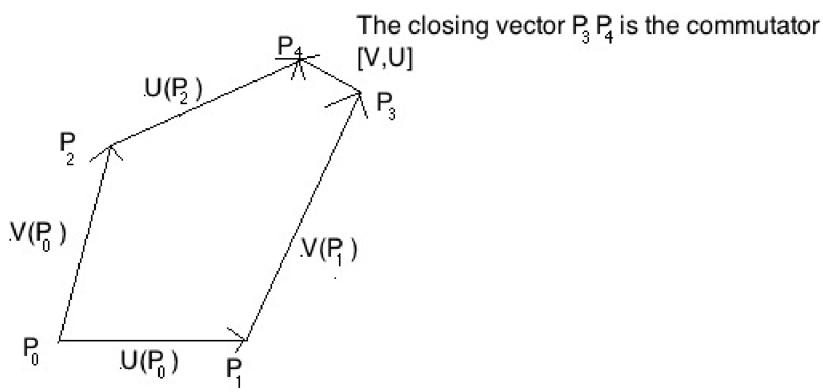


Figure 38: Commutator of Two Vector Fields $[V, U]$

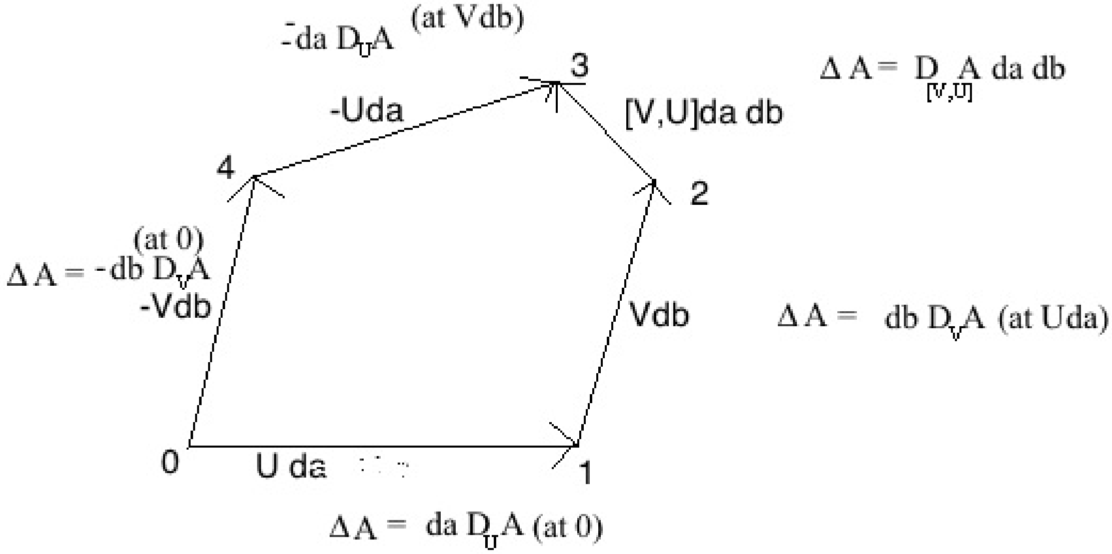


Figure 39: Curvature

4. When you parallel transport a vector around a closed loop, in general it comes back rotated. This is a measure of curvature.

5. As a preliminary step we show that the closure of a quadrilateral spanned by two vector fields (see figure 31) is the commutator.

   $U(P_2) + V(P_0) = \vec{P_0P_4}$, $U(P_0) + V(P_1) = \vec{P_0P_3}$. So

$$U(P_2) + V(P_0) - (U(P_0) + V(P_1)) = \vec{P_3P_4}$$

   $U(P_2) - U(P_0) = V^\alpha U^\beta_{,\alpha} \vec{e}_\beta|_{P_0}$ and $V(P_1) - V(P_0) = U^\alpha V^\beta_{,\alpha} \vec{e}^\beta|_{P_0}$.

$$\vec{P_0P_3} = [V^\alpha U^\beta_{,\alpha} - U^\alpha V^\beta_{,\alpha}]\vec{e}_\beta = [V, U]$$

6. Transport a vector $\vec{A}$ around this quadrilateral and see how much it rotates. It is useful to define an arbitrary vector field $\vec{A}_{field}(x)$ so that one can measure the change in our vector w.r.t this vector field at each stage of the journey i.e. at each leg. Choose $A_f(0) = A$ for convenience. ∇A_f along a leg gives how much $A(0)$ is rotated w.r.t $A(1)$ if it is parallel transported. Similarly do it for each leg. Add up the changes. To this approx the sum of these changes is the total change experience by one vector as it goes around. (The vector that is being transported from 1 to 2 is not quite the same as the one that arrived from 1, but this difference is $O(da)$ so the error due to this is higher order.)

   Define A_m as a mobile field which is defined at the vertices of the quadrilateral and is obtained by parallel transport. Thus $A_m(1)$ is obtained fro $A_m(0)$ by parallel transport. $A_m(1)$ is the original vector A. $A_m(2)$ is obtained by parallel transporting $A_m(1)$ and so on.

   Then $A_m(2)-A_m(1)+A_f(1)$ is a good approximation to the vector at 2 obtained by parallel transporting $A_m(1)$. It would be exact if $A_m(1) = A_f(1)$. The error is second order in the size of the quadrilateral because $A_m(1) - A_f(1)$ is already first order in da, db. Thus $A_f(2) - (A_m(2) - A_m(1) + A_f(1)) = db\nabla_V A_f(1)$. If we add up these quantities along all the sides one obtains $A_m(0') - A_m(0)$ because everything else cancels. Here $A_m(0')$ is the value of the mobile field at 0 after going around. While $A_f(0)$ is a fixed value of the field A_m does not define a single valued field. Thus adding up covariant derivatives as shown in the figure gives the net rotation.

   So as the figure shows, the change in step 4 to 0 and 1 to 2 is

$$db[\nabla_V A_f(0) - \nabla_V A_f(1)] = dadb\nabla_U\nabla_V A_f(0)$$

   Similarly 0 to 1 and 3 to 4 is

$$da[\nabla_U A_f(0) - \nabla_U A_f(4)] = -dadb\nabla_V\nabla_U A_f(0)$$

   Finally from 2 to 3

$$dadb\nabla_{[V,U]}A_f$$

   Total change (to this order of accuracy all derivatives can be evaluated at 0):

$$dadb[\nabla_U\nabla_V - \nabla_V\nabla_U + \nabla_{[V,U]}]A = dadb[\nabla_U\nabla_V - \nabla_V\nabla_U - \nabla_{[U,V]}]A \qquad (122)$$

$$\equiv dadb\ \mathbb{R}(\vec{U}, \vec{V})A$$

   The output is a vector so $\mathbb{R}(U, V)$ is a linear operator. Thus

$$\delta A^\alpha = R^\alpha_{\ \beta\gamma\delta} A^\beta U^\gamma V^\delta$$

   Alternatively, one can think of $R(...., U, V, A)$ as a tensor with 4 slots. The last slot is to be filled with a one form and the result is a number.

   In a coordinate basis with $U = \partial_j$ and $V = \partial_i$, the commutator $[U, V] = 0$. Then the curvature tensor is just $(\nabla_j\nabla_i - \nabla_i\nabla_j)A^k = R^k_{\ lji}A^l$.

7. Let us show that $(\nabla_j\nabla_i - \nabla_i\nabla_j)A^k$ is a linear operator. Consider $(\nabla_j\nabla_i - \nabla_i\nabla_j)fA^k$. Use Leibnitz rule

$$\nabla_j\nabla_i(fA^k) = \nabla_j[(\nabla_i f)A^k + f\nabla_i A^k] = (\nabla_j\nabla_i f)A^k + (\nabla_i f)\nabla_j A^k + (\nabla_j f)\nabla_i A^k + f\nabla_j\nabla_i A^k$$

The first three terms are symmetric in i, j and cancel when we antisymmetrize. So we get

$$[\nabla_j\nabla_i - \nabla_i\nabla_j](fA^k) = f[\nabla_j\nabla_i - \nabla_i\nabla_j](A^k)$$

This shows that it is a linear operator. It does not depend on the values of A^k at neighbouring points. For if $f(0) = 1$, but $f(\epsilon) \neq 1$, we have a different vector field, but the answer we get for the commutator depends only on $f(0) = 1$.

8. Let us use (122) to evaluate the curvature tensor in a coordinate basis:

Let $U = \partial_{x^i}$ and $V = \partial_{x^j}$. Then we need to evaluate

$$\begin{aligned}
&(\nabla_j\nabla_i - \nabla_i\nabla_j)A^k \\
&= (\nabla_j(\partial_i A^k + \Gamma^k_{il}A^l) - (i \leftrightarrow j) \\
&= (\partial_j(\partial_i A^k + \Gamma^k_{il}A^l) - \Gamma^l_{ji}\nabla_l A^k + \Gamma^k_{jl}(\partial_i A^l + \Gamma^l_{im}A^m) - (i \leftrightarrow j) \\
&= (\partial_j\partial_i A^k + (\partial_j\Gamma^k_{il})A^l) + (\Gamma^k_{il})\partial_i A^l - \Gamma^l_{ji}\nabla_l A^k + \Gamma^k_{jl}(\partial_i A^l + \Gamma^l_{im}A^m) - (i \leftrightarrow j)
\end{aligned}$$

The sum of the third, fourth and fifth terms in the expression above, is symmetric in i, j. Thus when we antisymmetrize in i, j these three terms drop out and we get

$$\begin{aligned}
&(\partial_j\Gamma^k_{il} - \partial_i\Gamma^k_{jl})A^l + (\Gamma^k_{jl}\Gamma^l_{im} - \Gamma^k_{il}\Gamma^l_{jm})A^m \\
&= (\partial_j\Gamma^k_{il} - \partial_i\Gamma^k_{jl} + \Gamma^k_{jm}\Gamma^m_{il} - \Gamma^k_{im}\Gamma^m_{jl})A^l \\
&\equiv R^k{}_{lji}A^l
\end{aligned} \tag{123}$$

This is the expression for the **Riemann** curvature tensor in a coordinate basis. Note the similarity with Yang-Mills field strength: Let us use greek indices for the derivatives, i.e $[\nabla_\nu, \nabla_\mu]$ and write $\Gamma^a_{\mu b}$. Then we have

$$R^a{}_{b\mu\nu} = \partial_\mu(\Gamma_\nu)^a_b - \partial_\nu(\Gamma_\mu)^a_b + [\Gamma_\mu, \Gamma_\nu]^a_b$$

This is like a gauge field $(A_\mu)^a_b$ where a, b are now internal indices. The difference is that for Γ all the indices are space time indices.

9. If $\Gamma^\alpha_{\beta\gamma} = 0 \implies$ geodesics are straight lines i.e. flat space time. Also $R^\alpha{}_{\beta\gamma\delta} = 0$. So flat space $\implies R^\alpha{}_{\beta\gamma\delta} = 0$.The reverse is also true:

....

10. A compact derivation using covariant exterior derivative: Let $\vec{V}$ be a vector. We evaluate $\mathbf{dd}\vec{V}$. Note that it ($\mathbf{dd} \neq 0$) is not zero as it would be for ordinary exterior derivatives ($d^2 = 0$), because this is a covariant version.

$$\begin{aligned}
\mathbf{d}\vec{V} &= \mathbf{d}(V^a\vec{e}_a) = \underline{dV}^a\vec{e}_a + V^a\underline{\mathbf{d}\vec{e}_a} = \underline{dV}^a\vec{e}_a + V^a\underline{\omega}^c{}_a\vec{e}_c \\
&= \underline{dx}^\mu\frac{\partial V^c}{\partial x^\mu}\vec{e}_c + V^a\underline{dx}^\mu\Gamma_\mu{}^c{}_a\vec{e}_c = \underline{dx}^\mu\nabla_\mu V^c\vec{e}_c
\end{aligned}$$

$$\begin{aligned}
\mathbf{dd}\vec{V} = \mathbf{d}(\underline{dV}^a\vec{e}_a + V^a\underline{\mathbf{d}\vec{e}_a}) &= \mathbf{d}(\underline{dV}^a\vec{e}_a + V^a\underline{\omega}^b{}_a\vec{e}_b) = -dV^a\wedge\mathbf{d}\vec{e}_a + d(\underline{\omega}^a{}_bV^b)\vec{e}_a - (\underline{\omega}^a{}_bV^b)\wedge\mathbf{d}\vec{e}_a \\
&= -dV^a\wedge\underline{\omega}^b{}_a\vec{e}_b + d\underline{\omega}^a{}_bV^b\vec{e}_a - \underline{\omega}^a{}_b\wedge dV^b\vec{e}_a - \underline{\omega}^a{}_bV^b\wedge(\underline{\omega}^c{}_a\vec{e}_c) \\
&= d\underline{\omega}^a{}_bV^b\vec{e}_a - \underline{\omega}^c{}_bV^b\wedge(\underline{\omega}^a{}_c\vec{e}_a) = d\underline{\omega}^a{}_bV^b\vec{e}_a + \underline{\omega}^a{}_c\wedge\underline{\omega}^c{}_bV^b(\vec{e}_a)
\end{aligned}$$

$$= (d\underline{\omega}^a{}_b + \underline{\omega}^a{}_c \wedge \underline{\omega}^c{}_b)V^b\vec{e}_a = \underbrace{\underline{\mathbb{R}}^a{}_b}_{two-form} V^b\vec{e}_a \tag{124}$$

$$\underline{\mathbb{R}}^a{}_b = R^a{}_{b\mu\nu}dx^\mu \wedge dx^\nu$$

The object we obtain is a vector valued two form - which is at it should be since we acted on a vector with **dd**. The two-form $\mathbb{R}^a{}_b$ is the Riemann curvature two form. This is very similar to the YangMills field strength two form $(F_{\mu\nu})^a{}_b dx^\mu \wedge dx^\nu$. $\underline{\omega}^a{}_b$ plays the role of the gauge field $(A_\mu)^a{}_b dx^\mu$ written as a form.

11. **Example of Curvature Computation**

12.4 Useful Properties of Curvature and Connection:

1.

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2}g^{\sigma\rho}[\partial_\mu g_{\rho\nu} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}]$$

$$\Gamma^\mu_{\mu\nu} = \frac{1}{2}g^{\mu\rho}[\partial_\mu g_{\rho\nu} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}] = \frac{1}{2}g^{\mu\rho}\partial_\nu g_{\rho\mu}$$

$$= \frac{1}{2}Tr[g^{-1}\partial_\nu g] = \partial_\nu ln\ \sqrt{g} = \frac{1}{\sqrt{g}}\partial_\nu\sqrt{g}$$

2.

$$\nabla_\mu V^\mu = \partial_\mu V^\mu + \Gamma^\mu_{\nu\mu}V^\nu = \partial_\mu V^\mu + \frac{1}{\sqrt{g}}(\partial_\nu\sqrt{g})V^\nu$$

$$= \frac{1}{\sqrt{g}}\partial_\mu(\sqrt{g}V^\mu)$$

$$\int d^4x\sqrt{g}\nabla_\mu V^\mu = \int d^4x\partial_\mu(\sqrt{g}V^\mu)$$

This is a surface term - just as in flat space.

3.

$$R^a{}_{bcd} = \frac{\partial\Gamma^a{}_{bd}}{\partial x^c} - \frac{\partial\Gamma^a{}_{bc}}{\partial x^d} + \Gamma^a{}_{ec}\Gamma^e{}_{bd} - \Gamma^a{}_{ed}\Gamma^e{}_{bc}$$

In an inertial coordinate system we can choose $\Gamma = 0$ at a point. So

$$R^a{}_{bcd} = \frac{\partial\Gamma^a{}_{bd}}{\partial x^c} - \frac{\partial\Gamma^a{}_{bc}}{\partial x^d}$$

Suppose in this coordinate system, $\Gamma \to \Gamma + \delta\Gamma$. Then

$$\delta R^a{}_{bcd} = \frac{\partial\delta\Gamma^a{}_{bd}}{\partial x^c} - \frac{\partial\delta\Gamma^a{}_{bc}}{\partial x^d}$$

But we know that $\delta\Gamma$ is a tensor. (Why? Because the non tensorial term in the transformation law of Γ does not depend on Γ. So in $\Gamma_1 - \Gamma_2$ they cancel out.) So we should be able to write the above equation as a tensor equation:

$$\delta R^a{}_{bcd} = \nabla_c(\delta\Gamma^a{}_{bd}) - \nabla_d(\delta\Gamma^a{}_{bc})$$

Contracting a, c we get Palatini's equation:

$$\delta R^a{}_{bad} = \delta R_{bd} = \nabla_a(\delta\Gamma^a{}_{bd}) - \nabla_d(\delta\Gamma^a{}_{ba}) \tag{125}$$

4. Consider the following

$$\int d^4x\sqrt{g}g^{ab}\delta R_{ab} = \int d^4x\sqrt{g}g^{ab}[\nabla_e(\delta\Gamma^e{}_{ab}) - \nabla_b(\delta\Gamma^e{}_{ae})]$$

$$= \int d^4x\sqrt{g}[\nabla_e(g^{ab}\delta\Gamma^e{}_{ab}) - \nabla_b(g^{ab}\delta\Gamma^e{}_{ae})]$$

These are of the form

$$\sim \int d^4x[\partial_e V_1^e + \partial_b V_2^b]$$

and is thus a total divergence.

5. Using

$$\delta\sqrt{g} = \sqrt{g}\delta(ln\ \sqrt{g}) = \sqrt{g}\frac{1}{2}\delta ln\ g = \frac{\sqrt{g}}{2}g^{ab}\delta g_{ab} = -\frac{\sqrt{g}}{2}(\delta g^{ab})g_{ab}$$

we have

$$\delta(\sqrt{g}g^{ab}) = \sqrt{g}\delta g^{ab} - \frac{\sqrt{g}}{2}g_{cd}\delta g^{cd}g^{ab}$$

6. Consider the following action

$$\int d^4x\ \sqrt{g}g^{ab}R_{ab} \tag{126}$$

Its variation is thus

$$\int d^4x\ (\sqrt{g}\delta g^{ab} - \frac{\sqrt{g}}{2}g_{cd}\delta g^{cd}g^{ab})R_{ab}$$

$$= \int d^4x\ [\sqrt{g}(R_{ab} - \frac{1}{2}g_{ab}R)\delta g^{ab}] \tag{127}$$

(since δR_{ab} gives a total divergence.) The coefficient of δg^{ab}, tensor inside the curved brackets is the Einstein tensor.

7. The Riemann tensor in an inertial frame where $\Gamma = 0$ is

$$R_{abcd} = \frac{1}{2}(\frac{\partial^2 g_{ac}}{\partial x^b \partial x^d} - \frac{\partial^2 g_{ad}}{\partial x^b \partial x^c} - \frac{\partial^2 g_{bc}}{\partial x^a \partial x^d} + \frac{\partial^2 g_{bd}}{\partial x^a \partial x^c})$$

Note the pattern: g_{ab} cannot occur because it has to be antisymmetric in a, b. Start with g_{ac} then anti-symmetrize in the various indices.

8. **Bianchi Identity:** Check the following by explicitly substituting the expression above or even simpler, the following expression also valid in an inertial frame:

$$R^a{}_{bcd} = \partial_{[c}\Gamma^a_{d]b}$$

Then since $\partial_e\partial_c = \partial_c\partial_e$,

$$\partial_{[e}R^a{}_{|b|cd]} = \partial_{[e}\partial_c\Gamma^a_{d]b} = 0$$

The covariant version of the above is then valid in all frames:

$$\nabla_{[e}R^a{}_{|b|cd]} = \nabla_{[e}g^{af}R_{|fb|cd]} = g^{af}\nabla_{[e}R_{|fb|cd]} = 0$$

9. Symmetries:

   (a) R_{abcd} is anti symmetric in a, b and in c, d.

   (b) $R_{abcd} = R_{cdab}$ and $R_{[abcd]} = 0$

(c) $R^a{}_{[bcd]} = 0$

(d) Bianchi: $R^a{}_{b[cd;e]} \equiv 0$

10. Contract indices of Bianchi identity:

$$\delta^e_a [\nabla_e R^a{}_{bcd} + \nabla_d R^a{}_{bec} + \nabla_c R^a{}_{bde}] = 0$$

$R^a{}_{bac} = R_{bc}$ is called the **Ricci tensor**. $R^a{}_a \equiv R$ is called the **Ricci scalar**. Contract one more:

$$g^{bc} [\nabla_a R^a{}_{bcd} + \nabla_d R_{bc} - \nabla_c R_{bd}] = 0$$

$$[\nabla_a R^{ab}{}_{bd} + \nabla_d R - \nabla_b R^b_d] = [-\nabla_a R^{ba}{}_{bd} + \nabla_d R - \nabla_b R^b{}_d] = [-2\nabla_b R^b{}_d + \nabla_d R] = 0$$

$$\implies \nabla_b [R^b{}_d - \frac{1}{2} \delta^b_a R] = 0 \qquad (128)$$

$R_{bd} - \frac{1}{2} g_{bd} R \equiv G_{bd}$ is the **Einstein tensor** defined earlier and (128) expresses the fact that it is *covariantly conserved.*

11. **Einstein's equation:**

The simplest way to get Einstein's equation, (whose LHS is (127)) is to start with this "Einstein-Hilbert" action (126). LHS is the generalization of (121) which is the Newtonian equations for gravity. The normalization of the E-H action can be obtained by requiring that it should reduce to (121) in the Newtonian limit.

13 Mathematical Digression IV

13.1 Conformal Transformations

Reference: Wald's book

1.
$$\tilde{g}_{ab} = \Omega^2 g_{ab} \tag{129}$$
with Ω a smooth strictly positive function is a conformal transformation.

2. If $\exists$ a difeomorphism
$$\psi : M \to M : (\psi^* g)_{ab} = \Omega^2 g_{ab}$$
this diffeo is called a "conformal symmetry". But *not all* Ω's are associated with diffeo's. [eg in 2 dim. $g_{ab} \to e^{2\sigma} g_{ab}$ is a conformal transf on a metric - "Weyl Transf". Only if $e^{2\sigma} = |\frac{\partial w}{\partial z}|^2$ is it associated with a conformal transf or conf diffeo on coordinates $(w(z))$.

3. $\tilde{\nabla}, \tilde{R}_{abcd}$ associated with $\tilde{g}_{ab}$ are related to ∇, R_{abcd} associated with g_{ab}.

4. **(a)** Because $\Omega^2 > 0$, a vector v^a is timelike, spacelike or null w.r.t g_{ab}, $\implies$ it has the same prop. w.r.t $\tilde{g}_{ab}$. $\implies$ (M, g_{ab}) and $(M, \tilde{g}_{ab}$ have identical causal structure.

   (b) Converse is also true. If causal structure is the same at a point, the metrics are conformal to each other at that point. Proof:

   Basic idea of proof is that null vectors remain null vectors if causal structure is same .

   So let $t^a, x_1^a, x_2^a, ...x_n^a$ be a set of orthonormal vectors at a point P w.r.t g_{ab}. Then $t^a \pm x_i^a$ is a null vector w.r.t g_{ab} and also therefore w.r.t $\tilde{g}_{ab}$ because causal structure is same. Now $t^a + \frac{x_i^a + x_j^a}{\sqrt{2}}$ is also null in g_{ab} *because* x_i and x_j are orthogonal. Conversely, if $t^a + \frac{x_i^a + x_j^a}{\sqrt{2}}$ is null w.r.t g_{ab}, and $t^a \pm x_i^a$ is null, for all i, then it must be that x_i^a and x_j^a are orthogonal.

   Similarly $t^a + \frac{x_i^a + x_j^a}{\sqrt{2}}$ is also null in $\tilde{g}_{ab}$, because they are null w.r.t g_{ab} - and we have said the causal structures are the same. Also $t^a \pm x_i^a$ is null, for all i. But that implies x_i^a and x_j^a must also be orthogonal w.r.t $\tilde{g}_{ab}$. Thus $t^a, x_1^a, x_2^a, ...x_n^a$ form an orthogonal set w.r.t $\tilde{g}_{ab}$ and their norms are all equal. That means $\tilde{g}_{ab}$ and g_{ab} are related by an overall constant.

   Notation: $\tilde{g}_{ab} = \Omega^2 g_{ab}$. $\tilde{g}^{ab}$ is inverse of $\tilde{g}_{ab}$ and g^{ab} is inverse of g_{ab}. So $\tilde{g}^{ab} = \Omega^{-2} g^{ab}$.

13.2 Covariant Derivative,Geodesics

5. Derivative: Let ∇_a be covariant derivative w.r.t g_{ab}. $\tilde{\nabla}_a$ be derivative w.r.t $\tilde{g}_{ab}$. Any two covariant derivatives are related by
$$\tilde{\nabla}_a \omega_b = \nabla_a \omega_b - \underbrace{C^c{}_{ab}}_{tensor} \omega_c$$
$\tilde{\nabla}_a \tilde{g}_{bc} = 0$ gives a constraint on $C^c{}_{ab}$:
$$C^c{}_{ab} = \frac{1}{2} \tilde{g}^{cd} (\nabla_a \tilde{g}_{bd} + \nabla_b \tilde{g}_{ad} - \nabla_d \tilde{g}_{ab}) \tag{130}$$
(Think of ∇ as ∂, then this is the usual expression for Γ- Christoffel.)

   Now
$$\nabla_a \tilde{g}_{bc} = \nabla_a (\Omega^2 g_{bc}) = 2\Omega (\nabla_a \Omega) g_{bc} \quad [\because \nabla_a g_{bc} = 0]$$
(Note that $\nabla_a ln\ \Omega = \partial_a ln\ \Omega$.)
$$\therefore C^c{}_{ab} = \Omega^{-1} g^{cd} [(\nabla_a \Omega) g_{bd} + (\nabla_b \Omega) g_{ad} - (\nabla_d \Omega) g_{ab}]$$

Finally

$$C^c{}_{ab} = \nabla_{(a} ln\ \Omega \delta^c_{b)} - (\nabla_d ln\ \Omega) g^{cd} g_{ab} \tag{131}$$

where

$$\boxed{\tilde{\nabla}_a \omega_b = \nabla_a \omega_b - C^c{}_{ab} \omega_c}$$

6. Geodesics: Compare geodesics of the two metrics:

$$v^a \nabla_a v^b = 0$$

where v^a is the tangent vector field of along a geodesic γ of g_{ab}.

$$\begin{aligned} v^a \tilde{\nabla}_a v^b &= \underbrace{v^a \nabla_a v^b}_{=0} + v^a C^b{}_{ac} v^c \\ &= v^a [\nabla_{(a} ln\ \Omega \delta^b_{c)} v^c - (\nabla_d ln\ \Omega) g^{bd} g_{ac} v^c] \\ &= v^a (\nabla_a ln\ \Omega) v^b + (\nabla_c ln\ \Omega) v^c v^b - (\nabla_d ln\ \Omega) g^{bd} v.v \\ &= 2v^b (v.\nabla ln\ \Omega) - (\nabla_d ln\ \Omega) g^{bd} v.v \end{aligned} \tag{132}$$

So if geodesic is null,i.e. $v.v = 0$, then the equation is of the form $v^a \nabla_a v^b = \alpha v^b$ which is a geodesic with a non-affine parametrization. In more detail:

7. Non affine parameter for geodesic:
Let $v^a = \frac{dx^a}{ds}$. In flat space $x^a = v^a s$ and $v^a \frac{\partial}{\partial x^a} = \frac{d}{ds}$. Let us choose $g_{ab} = \eta_{ab}$. Then

$$\begin{aligned} v^a \tilde{\nabla}_a v^b &= 2v^b v.\nabla (ln\ \Omega) \\ &= 2v^b \frac{d}{ds} (ln\ \Omega) \end{aligned} \tag{133}$$

(In curved space we define

$$D_s A^a \equiv \frac{d}{ds} A^a + \Gamma^a_{bc} \frac{dx^b}{ds} A^c$$

Then

$$D_s A^a(x(s)) = v^b \nabla_b A^a$$

Geodesic eqn is $D_s \frac{dx^a}{ds} = 0$. Note that for a scalar $D_s = \frac{d}{ds}$ so (133) would continue to hold in curved space also.)

Now instead of changing matric let us change parametrization: Let $x^a = u^a t(s)$ where $t(s)$ is a non linear fn of s. Then

$$v^a = \frac{dx^a}{ds} = u^a \frac{dt}{ds}$$

$$\begin{aligned} v^a \nabla_a v^b &= \frac{d}{ds} v^b \\ &= \frac{d}{ds} (u^b \frac{dt}{ds}) \\ &= u^b \frac{d^2 t}{ds^2} = v^b \frac{\frac{d^2 t}{ds^2}}{\frac{dt}{ds}} = v^b \frac{d}{ds} (ln\ \frac{dt}{ds}) \end{aligned} \tag{134}$$

Comparing (133) with (134) we see that $\Omega^2 = \frac{dt}{ds}$ makes them identical. Thus the effect of a conformal rescaling of a flat metric is the same as an affine reparametrization *for a null geodesic*. The choice of flat metric for g_{ab} was not required - it was only made for simplicity.

Conclusion: Null geodesics continue to be null geodesics under conformal rescaling.

13.3 Riemann Tensor, Weyl Tensor

8. How do $C^c{}_{ab}, R_{abc}{}^d$ change under $g \to \tilde{g}$?

Recapitulation:

(a) If (∇_a, g_{ab}) with $\nabla_a g_{bc} = 0$ and $\tilde{\nabla}$ is another arbitrary covariant derivative (such as ∂), then

$$\nabla_a \omega_b = \tilde{\nabla}_a \omega_b - C^c{}_{ab}\omega_c$$

for some $C^c{}_{ab}$. Also

$$C^c{}_{ab} = \frac{1}{2} g^{cd}(\tilde{\nabla}_a g_{bd} + \tilde{\nabla}_b g_{ad} - \tilde{\nabla}_d g_{ab})$$

(This is just (130) with tilde and non tilde interchanged.)

(b) Suppose we start with (∇_a, g_{ab}) and change to $g_{ab}(\lambda)$. Need to find C. Write

$$\nabla^\lambda_a \omega_b = \tilde{\nabla}^0_a \omega_b - C^{c(\lambda)}_{ab}\omega_c$$

with

$$C^{c(\lambda)}_{ab} = \frac{1}{2} g^{cd}(\lambda)(\nabla^0_a g_{bd}(\lambda) + \nabla^0_b g_{ad}(\lambda) - \nabla^0_d g_{ab}(\lambda))$$

Thus we have ∇^λ in terms of ∇^0 and $C^{c(\lambda)}_{ab}$. So use

$$\nabla_a \nabla_b \omega_c - \nabla_b \nabla_a \omega_c = R_{abc}{}^d \omega_d = -R_{ab}{}^d{}_c \omega_d$$

to get $R^\lambda_{abc}{}^d$ in terms of $R^0_{abc}{}^d$ and $C^{c(\lambda)}_{ab}$. This gives

$$R^\lambda_{abc}{}^d = R^0_{abc}{}^d - \nabla^0_{[a} C^{d(\lambda)}_{b]c} + C^{e(\lambda)}_{c[a} C^{d(\lambda)}_{b]e}$$

The first term would be zero if we start with (∂, η_{ab}) and then the second and third terms give the usual expression for R in terms of Γ.

--- Plug

in for C in terms of Ω from (131).

$$\begin{aligned}\tilde{R}_{abc}{}^d &= R_{abc}{}^d + \delta^d_{[a}\nabla_{b]}\nabla_c ln\ \Omega - g^{de} g_{c[a}\nabla_{b]}\nabla_e ln\ \Omega + (\nabla_{[a} ln\ \Omega)\delta^d_{b]}\nabla_c ln\ \Omega \\ &\quad -(\nabla_{[a} ln\ \Omega) g_{b]c} g^{df}\nabla_f ln\ \Omega - g_{c[a}\delta^d_{b]} g^{ef}(\nabla_e ln\ \Omega)(\nabla_f ln\ \Omega)\end{aligned} \tag{135}$$

$$\begin{aligned}\tilde{R}_{ac} &= R_{ac} - (n-2)\nabla_a\nabla_c ln\ \Omega - g_{ac} g^{de}\nabla_d\nabla_e ln\ \Omega \\ &\quad +(n-2)(\nabla_a ln\ \Omega)(\nabla_c ln\ \Omega) - (n-2) g_{ac} g^{de}(\nabla_d ln\ \Omega)(\nabla_e ln\ \Omega)\end{aligned} \tag{136}$$

Contracting with $\tilde{g}^{ac} = \Omega^{-2} g^{ac}$

$$\tilde{R} = \Omega^{-2}[R - 2(n-1) g^{ac}\nabla_a\nabla_c ln\ \Omega - (n-2)(n-1) g^{ac}(\nabla_a ln\ \Omega)(\nabla_c ln\ \Omega)] \tag{137}$$

Weyl Tensor C_{abcd} is the trace free part of Riemann:

$$R\underbrace{ab}_{AS}\underbrace{cd}_{AS} = C\underbrace{ab}_{AS}\underbrace{cd}_{AS} + \frac{2}{n-2}(g_{a[c}R_{d]b} - g_{b[c}R_{d]a}) - \frac{2}{(n-1)(n-2)} R g_{a[c} g_{d]b}$$

$$\tilde{C}_{abc}{}^d = C_{abc}{}^d \quad ; \quad \tilde{C}_{abcd} = \Omega^2 C_{abcd}$$

9. Conformal Invariance of EOM etc:

   Conformal weight is determined by no. of upper indices and lower indices. eg $\int d^n x \underbrace{\sqrt{-g}}_{\Omega^n}\underbrace{g_{ab}}_{\Omega^2} T^{ab}$.

   Conformal inv requires $T^{ab} \to \Omega^{-n-2} T^{ab}$. Thus T^{ab} has conf wt $-n-2$ and T_{ab} then has conf wt $-n+2$. But T^{ab} has definite conf wt only if $T^a_a = 0$ - condn for conf invariance of theory.

   One can show that (see Wald) $\nabla_a T^{ab} = 0$ is conformally invariant with the above conformal dimension provided $T^{ab} = T^{ba}$ and $T^a_a = 0$.

   Finally the scalar wave eqn

$$g^{ab}\nabla_a\nabla_b\phi - \frac{n-2}{4(n-1)}R\phi = 0$$

   is conf invariant with conf dim of ϕ being $1 - \frac{n}{2}$.

   ($\int d^n x\ \sqrt{-g} g^{ab}\partial_a\phi\partial_b\phi$ is conf inv provided ϕ has dim $\frac{-n+2}{2} = 1 - \frac{n}{2}$. Note that we are not scaling x.)

14 Mathematical Digression V

14.1 Extrinsic Curvature

Reference: MTW

MTW notation

1. A hypersurface Σ is embedded in a higher dimensional space. $\hat{\mathbf{n}}$ is a unit normal vector field on Σ. We consider the change in $\hat{\mathbf{n}}$ as we move from a point P to P+δP, where δP is an infinitesimal vector in the hypersurface (i.e tangent to it). The change $\delta\hat{\mathbf{n}}$ is clearly proportional to δP So

$$\delta_{(\delta\vec{P})}\hat{\mathbf{n}} = -\mathbf{K}(\delta\vec{P}) \tag{138}$$

   (By $K(\vec{a})$ it is meant that K is a function of $\vec{a}$, in this case a linear function.)

2. Introduce $d\hat{\mathbf{n}}$, a vector valued one form. So (Note: $\mathbf{e_i}$ are basis vectors in Σ and $\hat{\mathbf{n}}.\mathbf{e_i} = 0$.)

$$< d\hat{\mathbf{n}}, \mathbf{e_i} >= \nabla_i\hat{\mathbf{n}}$$

   (just like $< df, \mathbf{e_i} >= \nabla_i f$).

   $\nabla_j \equiv \mathbf{e_j}.\vec{\nabla}$ is the covariant derivative along the direction specified by $\mathbf{e_j}$.

   i.e. the change in a vector can be written in terms of a vector valued form. So is $\mathbf{K}$ a vector valued one form. Introduce $\mathbf{dP} = \frac{\partial\mathbf{P}}{\partial\mathbf{x^i}}\mathbf{dx^i}$ or more formally $\frac{\partial}{\partial x^i}dx^i = \mathbf{e_i}\otimes\omega^{\mathbf{i}} = \mathbf{1}$ - a general displacement. We get an equation for for the general change in $d\hat{\mathbf{n}}$:

$$d\hat{\mathbf{n}} = -\mathbf{K}(\mathbf{dP}) = -\mathbf{K} = -\mathbf{K^i}_{\mathbf{j}}\mathbf{e_i}\otimes\omega^{\mathbf{j}}$$

   If we input the vector $dP^k\mathbf{e_k}$

$$dP^k\nabla_k\hat{\mathbf{n}} = -(K^i{}_k dP^k)\mathbf{e_i}$$

$$\nabla_j\hat{\mathbf{n}} = -K^i{}_j\mathbf{e_i}$$

   Note that the vector $\nabla_j\hat{\mathbf{n}}$ lies in the hypersurface Σ as required since $\nabla_j(\hat{\mathbf{n}}.\hat{\mathbf{n}}) = 0 = 2(\nabla_j\hat{\mathbf{n}}).\hat{\mathbf{n}}$.

3. Using $\mathbf{e_i}.\mathbf{e_j} = g_{ij}$ we get

$$\mathbf{e_m}.\nabla_j\hat{\mathbf{n}} = -K^i{}_j g_{im} = -K_{mj}$$

   $\nabla_j\hat{\mathbf{n}}$ in components is $(\nabla_j\hat{\mathbf{n}})^k\mathbf{e_k} \equiv \hat{n}^k{}_{;j}\mathbf{e_k}$. So $K_{im} = -\hat{n}_{i;m}$.

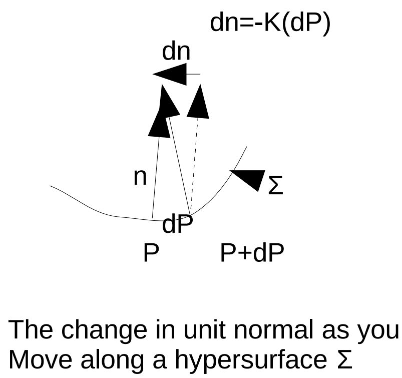


Figure 40: Extrinsic Curvature

4. **Property:** K_{im} **is symmetric:**

$$K_{mj} = -\mathbf{e_m}.\nabla_j \hat{\mathbf{n}} = +(\nabla_j \mathbf{e_m}).\hat{\mathbf{n}} = \Gamma^{\mu}_{jm}\hat{\mathbf{n}}.\mathbf{e}_{\mu} = \Gamma^{\mu}_{mj}\hat{\mathbf{n}}.\mathbf{e}_{\mu} = K_{jm}$$

Note the use of μ - which includes the normal direction to Σ.

Reference: Wald

1. Let ξ^a be a unit tangent vector field of a congruence of time like **geodesics** normal to Σ. (Both **n** and ξ are unit norm vector fields orthogonal to Σ but **n** is not necessarily geodesic.)

$$K_{ab} = \nabla_a \xi_b$$
$$\xi^a K_{ab} = 0 \quad \because \xi^a \nabla_a \xi_b = 0 \;\; (geodesic\ eqn)$$
$$K_{ab}\xi^b = 0 \quad \because \nabla_a(\xi_b \xi^b) = 0 \quad (unit\ norm)$$

So K_{ab} lies entirely in the hypersurface Σ.

2. By a theorem if ξ is hypersurface orthogonal geodesic, then $\omega_{ab} = \nabla_{[a}\xi_{b]} = 0 =$ *rotation*. So K_{ab} is symmetric.

$$K_{ab} = \frac{1}{2}(\nabla_a \xi_b + \nabla_b \xi_a) = \frac{1}{2}\mathcal{L}_\xi g_{ab} = \frac{1}{2}\mathcal{L}_\xi(h_{ab} - \xi_a \xi_b) = \frac{1}{2}\mathcal{L}_\xi h_{ab}$$

$h_{ab} = g_{ab} + \xi_a \xi_b$ is the spatial metric - because $\xi^a h_{ab} = 0$. (Using $\xi^a \xi_a = -1$).

3. If we let $\xi^a = (\frac{\partial}{\partial t})^a$ then $\mathcal{L}_\xi = \frac{\partial}{\partial t}$. So

$$K_{ab} = \frac{1}{2}\frac{\partial}{\partial t}h_{ab}$$
$$TrK = h^{ab}K_{ab} = \theta - \; expansion$$

4. Along the hypersurface n and ξ match (both are orthogonal and unit length). So

$$K_{ab} = \nabla_a \xi_b = h_a{}^c \nabla_c \xi_b = h_a{}^c \nabla_c n_b = \frac{1}{2}\mathcal{L}_n h_{ab}$$

($h_a{}^b$ projects onto the hypersurface).

Some examples: A circle embedded in R^2.

$$ds^2 = dr^2 + r^2 d\theta^2$$

$(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta})$ is a basis for vectors. **n** is the unit normal to the circle.

$$\mathbf{n} = \frac{\vec{\partial}}{\partial r}$$
$$K^i_\theta \mathbf{e_i} = \nabla_\theta \mathbf{n} \qquad K^i{}_\theta = -n^i{}_{;\theta}$$
$$\frac{\vec{\partial}}{\partial r} = cos\theta \frac{\vec{\partial}}{\partial x} + sin\theta \frac{\vec{\partial}}{\partial y} \implies n^x = cos\theta, \quad n^y = sin\theta$$
$$n^x{}_{;\theta} = n^x{}_{,\theta} = -sin\theta, \quad n^y{}_{,\theta} = cos\theta$$
$$\therefore K^x{}_\theta = sin\theta, \quad K^y{}_\theta = -cos\theta$$
$$K^r{}_\theta = \hat{r}.\mathbf{K} = cos\theta K^x{}_\theta + sin\theta K^y{}_\theta = 0$$

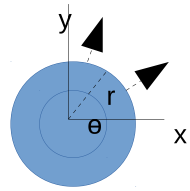


Figure 41: Extrinsic Curvature Calculation

$$\frac{\vec{\partial}}{\partial\theta} = -r\ sin\theta\frac{\vec{\partial}}{\partial x} + r\ cos\theta\frac{\vec{\partial}}{\partial y} \qquad \Longrightarrow \hat{\theta} = -sin\theta\frac{\vec{\partial}}{\partial x} + cos\theta\frac{\vec{\partial}}{\partial y}$$

$$\hat{\theta}.\mathbf{K} = K^{\hat{\theta}}{}_{\theta} = -sin\theta K^{x}{}_{\theta} + cos\theta K^{y}{}_{\theta} = -1$$

$$\Longrightarrow K^{\hat{\theta}}{}_{\hat{\theta}} = -\frac{1}{r} = K_{\hat{\theta}\hat{\theta}}$$

Another way: Use the formula (with $r = t$ and $\xi = \frac{\partial}{\partial t}$):

$$K_{\theta\theta} = \frac{1}{2}\partial_t h_{\theta\theta} = \frac{1}{2}\partial_r g_{\theta\theta} = r$$

This is consistent with $K_{\theta\theta} = r^2 K_{\hat{\theta}\hat{\theta}}$. (Wald uses the opposite sign convention for K)