General Relativity PHYS 4473.
Chapter 1: Review of Special Relativity

Definition of an observer's frame of reference:
(1) A clock
(2) A measurement system for distance (e.g. rod)
(3) A set of coordinate axes (preferably cartesian)

The Newtonian principle of relativity can be stated as follows:

"All inertial observers are equivalent as far as dynamical experiments are concerned"

"Dynamical" because that was the dominant mode of investigation during Newton's time.

"Problems" with this statement:
In the 1880's Hertz showed that light was an EM wave travelling at speed c. The waves were assumed to travel through a fixed "ether" that provided a distinct frame of reference, namely, at rest relative to the ether. Michelson-Morley experiment showed c was invariant regardless of motion relative to "ether", and the preferred reference frame was discarded.

Einstein realised that the Newtonian principle must be extended to include all of physics

Postulate I: Principle of Special Relativity

All inertial observers are equivalent

This is the logical completion of Newtonian thought.

Given the constancy of c, the second postulate is

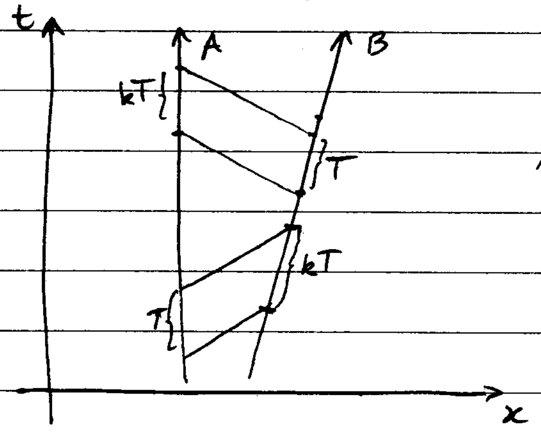
Postulate 2: Constancy of the velocity of light

The velocity of light is the same in all inertial frames.

Now consider arrival of pulses of light between observers in different inertial frames. This will follow the rules of the "k-calculus."

The k-calculus

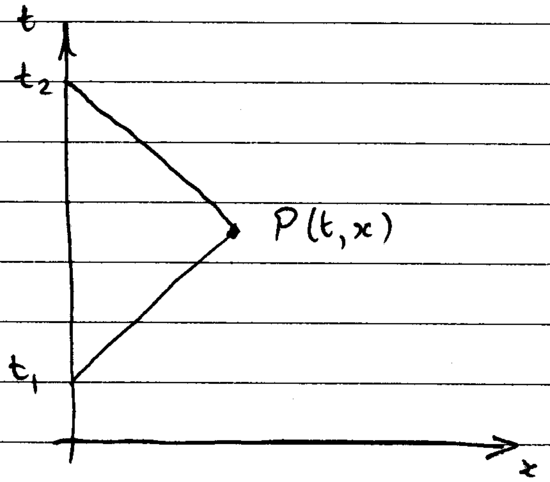
Suppose we have two observers A & B in inertial frames, S_A & S_B, in 2-dimensions for simplicity.
A sends out pulses separated by T
B recieves pulses separated by kT in its frame.
k is the "k-factor" & is a characteristic of the motion of B relative to A.



The same rule applies for pulse separated by T in B's frame that are sent to A (follows from postulate I)

We can assign coordinates to events by an idea similar to radar ranging.

Suppose we bounce light off an event $P(t,x)$, what are its coordinates?

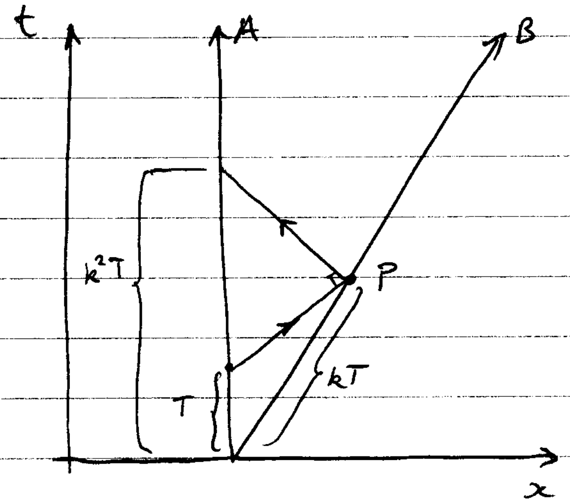


Clearly: $t = t_1 + \frac{1}{2}(t_2 - t_1)$
$= \frac{1}{2}(t_2 + t_1)$
$x = \frac{1}{2}(t_2 - t_1)\tan\frac{\pi}{4}$
$= \frac{1}{2}(t_2 - t_1)$

Thus $P(t,x) = (\frac{1}{2}(t_2+t_1), \frac{1}{2}(t_2-t_1))$

Measuring Relative Speeds:

Consider two inertial observers A & B



After time T A sends pulse to P, & B back to A. Coordinates of P are thus

$$P(t,x) = \left(\tfrac{1}{2}(k^2+1)T, \tfrac{1}{2}(k^2-1)T\right)$$

For variable T this describes B's worldline.

Thus B's velocity is

$$v = \frac{x}{t} = \frac{k^2-1}{k^2+1}$$

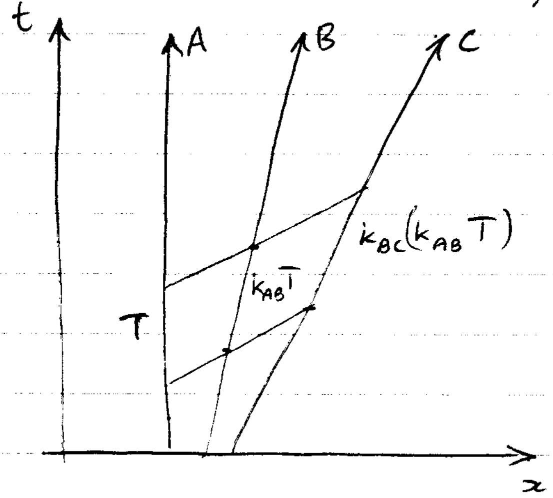
Rearranging for k as a function of v

$$\boxed{k = \sqrt{\frac{1+v}{1-v}}}$$

Notes: (1) relativistic formula for Doppler shift.
(2) Under $v \to -v$ then

$$k \to \frac{1}{k}$$

To look at relative velocities, extend to include a third reference frame/observer (C).



Thus we see that

$$k_{AC} = k_{BC}\, k_{AB}$$

Using our previous ranging argument to describe C's worldline

$$V_{AC} = \frac{x_c}{t_c} = \frac{k_{AC}^2 - 1}{k_{AC}^2 + 1} = \frac{k_{BC}^2 k_{AB}^2 - 1}{k_{BC}^2 k_{AB}^2 + 1}$$

Substitute velocities into k-factors to get

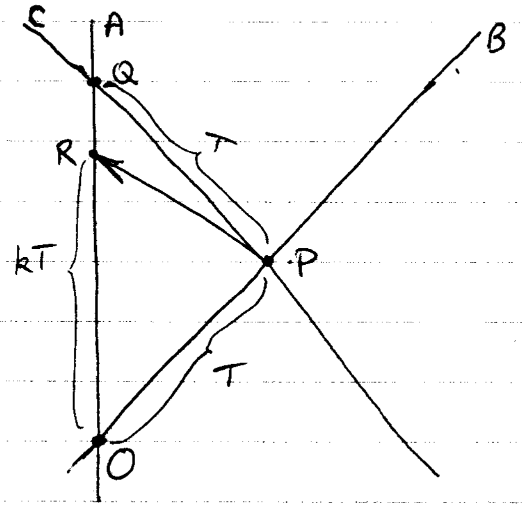
$$\boxed{V_{AC} = \frac{V_{BC} + V_{AB}}{1 + V_{AB} V_{BC}}}$$

While the composition of k-factors is very easy, the composition of velocities is more complex.

Question: What happens if one of V_{BC} or V_{AB} is "c"?

The Clock Paradox.

Consider three observers A, B, C with $v_{AC} = -v_{AB}$



A & B synchronize at O
A & C cross at Q.
B & C meet at point P, which B sees as occuring at time T after synching with A. Send pulse to A at this time

A sees pulse arrive at kT

Since $v_{AC} = -v_{AB}$, the k-factor associated with the distance QR must be k^{-1}

Total time observed by A between O & Q is

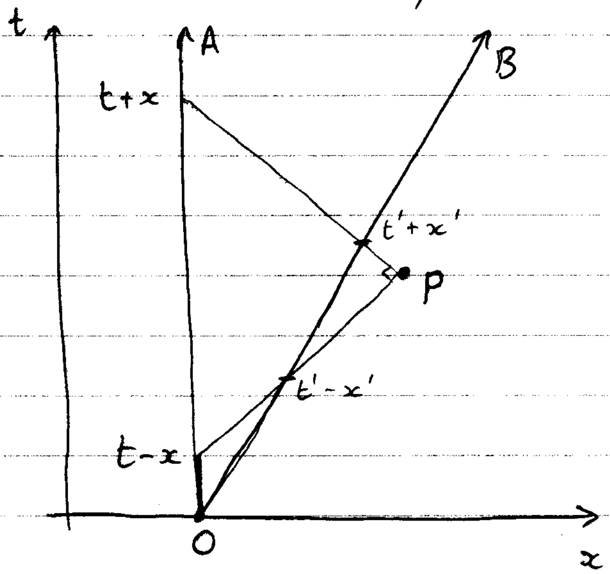
$$t = \left(k + \frac{1}{k}\right) T$$

This is greater than 2T for any $k = 1$.

Resolution: Time is route dependent, and further the 2T is measured by 2 observers, not. Passing the clock from B to C requires an acceleration & the clock is no longer in an inertial frame.

This is essentially the twin paradox.

Lorentz Transformations via k-calculus



Coordinates of an event P
$P(t,x)$ & $P(t',x')$

To illuminate P, A must send signal at $t-x$, bounced signal received at $t+x$.

Synchronize at O, then

$$k(t-x) = t'-x'$$

Similarly

$$k(t'+x') = t+x$$

Using $k = \left(\frac{1+v}{1-v}\right)^{1/2}$ can solve for t' & x' in terms of t, v, x

Get:

$$t' = \frac{t - vx}{(1-v^2)^{1/2}}$$

$$x' = \frac{x - vt}{(1-v^2)^{1/2}}$$

Straightforward algebra shows that

$$t'^2 - x'^2 = t^2 - x^2$$

ie. $t^2 - x^2$ is invariant under the Lorentz boost.

The argument can be extended to 4-dimensions to give,

$$t' = \frac{t - vx}{(1-v^2)^{1/2}}, \quad x' = \frac{x - vt}{(1-v^2)^{1/2}}, \quad y' = y, \quad z' = z$$

Where the invariant is defined as

$$s^2 = (t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2$$

for two events (t_2, x_2, y_2, z_2) and (t_1, x_1, y_1, z_1).

If the points are infinitesimally separated ie (t, x, y, z) and $(t + dt, x + dx, y + dy, z + dz)$ Then

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2$$

Is the invariant.

A four dimensional spacetime in which the above form is invariant is called a <u>Minkowski spacetime</u>.

Key Features of Special Relativity

- Review Section 3.1 in your own time, especially derivation of LT using complex time.

Lorentz transformations map one frame to another in a linear fashion. This necessarily implies that the translation between frames S & S' (coords (t, x, y, z), (t', x', y', z')) can be written in matrix form

$$\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = L \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix}$$ where L is a 4×4 matrix

It can also be shown that the matrix L is a product of

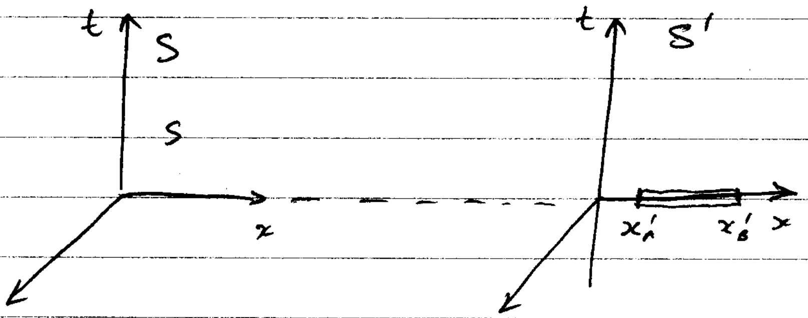
(1) a 3D rotation
(2) Lorentz boost
(3) a second 3D rotation

The special Lorentz transformations (boosts) form a group, recall a group is defined by
(1) A set G and multiplication $g_1 g_2 \in G$ $\forall\, g_1, g_2 \in G$
(2) Associative multiplication: $g_1, g_2 \in G$
then $g_1(g_2 g_3) = (g_1 g_2) g_3$
(3) Identity "e" $e \in G \Rightarrow eg = ge = g$ $\forall\, g \in G$
(4) Inverse $g^{-1} \in G$ $g^{-1} g = g g^{-1} = e$

Two boosts combine to give a third
What is the identity element? Given by $v = 0$
" " " inverse " ? " " $-v$
Associativity from addition of imaginary angles.

Length Contraction

Consider a fixed rod in S' with end points x_A' & x_B' at rest in S'



Use Lorentz boost formula to give

$$x_A' = \gamma(x_A - vt_A)$$
$$x_B' = \gamma(x_B - vt_B)$$

Define the rest length in S': $l_0' = x_B' - x_A'$

The length in S at an instant $t = t_A = t_B$ is then $l = x_B - x_A$

$$\therefore \quad l = \frac{x_B'}{\gamma} - \frac{x_A'}{\gamma} = \frac{1}{\gamma} l_0' \qquad (l < l_0)$$

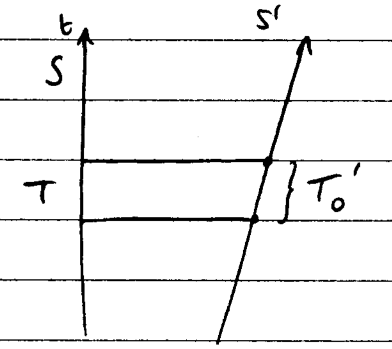
(1) Length in moving frame is reduced by $\sqrt{1-\frac{v^2}{c^2}}$

(2) Greatest length in rest frame

(3) Rest frame length is called the "proper length"

(4) As $v \to c$ $l \to 0$

Time dilation



Clock is fixed in S' at $x' = x_A'$
Records two events in S'
(x_A', t_1') & $(x_A', t_1' + T_0')$

Under the LT from S' to S
(change to primes and $v \to -v$)

$$t = \gamma\left(t' + vx'/c^2\right)$$

Define

$$t_1 = \gamma\left(t_1' + \frac{x_A' v}{c^2}\right), \quad t_2 = \gamma\left(t_1' + T_0' + \frac{x_A' v}{c^2}\right)$$

Interval in S is thus

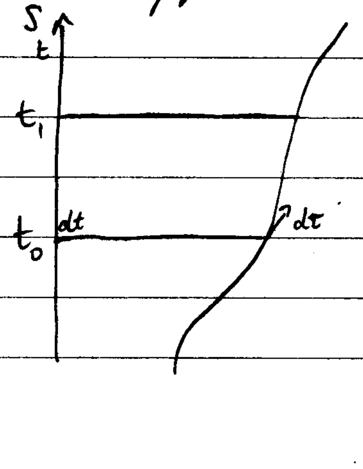
$$T = t_2 - t_1 = \gamma T_0' \qquad (T > T_0')$$

So moving clocks go slower by a factor $\left(1 - \frac{v^2}{c^2}\right)^{-1/2}$
& the effect is reciprocal.

Accelerated Clock

Ideal Clock: Its rate depends only upon its instantaneous speed v (rate is unaffected by acceleration)

Suppose time coordinate on path is τ



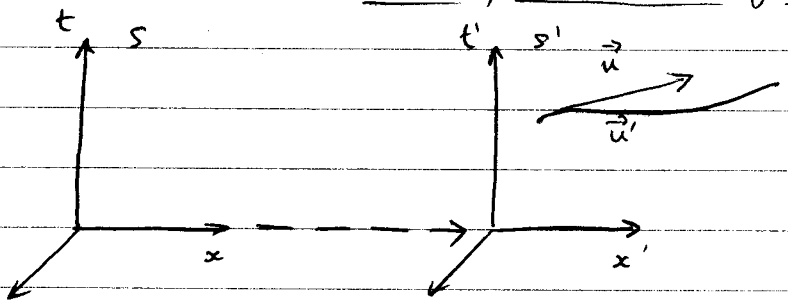
then $\gamma\, d\tau = dt$

$$\therefore \quad \tau = \int_{t_0}^{t_1} \frac{dt}{\gamma} = \int_{t_0}^{t_1} \left(1 - \frac{v^2}{c^2}\right)^{1/2} dt$$

τ is the "proper time"

Note defining an ideal clock is easy, but actually finding one is not. Closest is probably an atomic clock.

Transformations of velocities



Consider a particle in motion with Cartesian three-velocity components u_i:

$$(u_1, u_2, u_3) = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right) \quad \text{in } S$$

$$(u_1', u_2', u_3') = \left(\frac{dx'}{dt'}, \frac{dy'}{dt'}, \frac{dz'}{dt'} \right) \quad \text{in } S'$$

then under the LT

$$t' = \gamma\left(t - \frac{vx}{c^2}\right), \quad x' = \gamma(x - vt), \quad y' = y, \quad z' = z$$

taking differentials gives ~~two~~

$$u_1' = \frac{dx'}{dt'} = \frac{\gamma(dx - v\,dt)}{\gamma\left(dt - v\frac{dx}{c^2}\right)} = \frac{\frac{dx}{dt} - v}{1 - \frac{1}{c^2}\left(v\frac{dx}{dt}\right)} = \frac{u_1 - v}{1 - \frac{u_1 v}{c^2}}$$

$$u_2' = \frac{dy'}{dt'} = \frac{dy}{\cancel{\gamma}\left(dt - \frac{v\,dx}{c^2}\right)} = \frac{dy/dt}{\gamma\left\{1 - \frac{1}{c^2}\left(v\frac{dx}{dt}\right)\right\}} = \frac{u_2}{\cancel{\gamma}\left(1 - \frac{u_1 v}{c^2}\right)}$$

Similarly $u_3' = \dfrac{u_3}{\gamma\left(1 - \frac{u_1 v}{c^2}\right)}$

Key point: Components transverse to motion are also changed!

Acceleration in SR

We can also derive relations between the coordinate acceleration in different frames.

Begin from $$u_1 = \frac{u_1' + v}{1 + u_1'v/c^2}$$ (from previous section)

and again take differentials, e.g.

$$du_1 = \frac{1}{\gamma^2} \frac{du_1'}{(1 + u_1'v/c^2)^2}$$

Since $$dt = \gamma\left(1 + \frac{u_1'v}{c^2}\right) dt'$$ (from LT)

$$\frac{du_1}{dt} = \frac{1}{\gamma^3\left(1 + \frac{u_1'v}{c^2}\right)^3} \frac{du_1'}{dt'}$$

Exercise: Derive the relations for $\frac{du_2}{dt}$ & $\frac{du_3}{dt}$

Important point: If $\frac{du_1'}{dt} \neq 0$ in one frame then there is no other inertial frame for which $\frac{du_1}{dt} = 0$. <u>Acceleration is absolute in SR</u>

Uniform Acceleration

$\frac{du}{dt}$ = constant doesn't work in SR – limited by c.

Uniform acc^n in SR is <u>defined</u> as being constant in a frame that moves with the accelerating system (co-moving). Particle is described by $u(t)$.
Velocity of accelerating system in the co-moving frame, u', must be zero, acceleration is a'.

Velocity of the comoving frame is $v = u(t)$

Then $\frac{du_1}{dt} = \frac{1}{\gamma^3}\frac{du_1'}{dt'} = \left(1 - \frac{u^2}{c^2}\right)^{3/2} a'$

Separate the variables

$$\frac{du_1}{\left(1 - \frac{u^2}{c^2}\right)^{3/2}} = a'\,dt$$

Assuming the particle starts from rest at $t = t_0$

$$\frac{u}{\left(1 - \frac{u^2}{c^2}\right)^{1/2}} = a'(t - t_0)$$

Solve for u:

$$u = \frac{a'(t - t_0)}{\left[1 + a'^2 (t - t_0)^2 / c^2\right]^{1/2}}$$

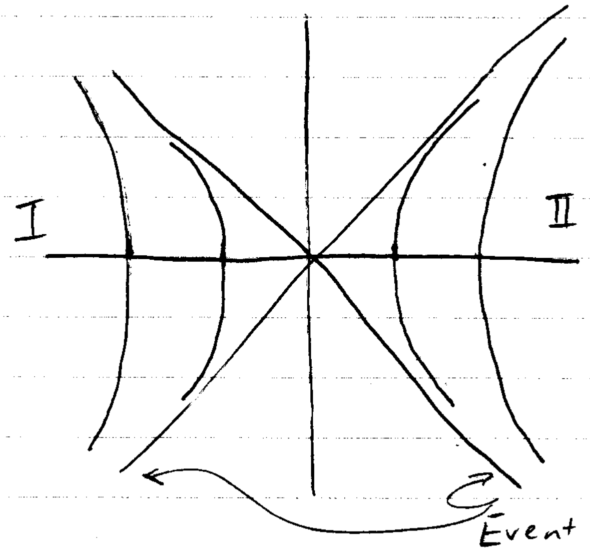
Let $x = x_0$ at $t = t_0$ and integrate $u = \frac{dx}{dt}$

$$(x - x_0) = \frac{c}{a'}\left[c^2 + a'^2(t - t_0)^2\right]^{1/2} - \frac{c^2}{a'}$$

Which can be re-written as

$$\frac{(x - x_0 + c^2/a')^2}{(c^2/a')^2} - \frac{(ct - ct_0)^2}{(c^2/a')^2} = 1$$

Draws out hyperbolae in (x, ct) space.



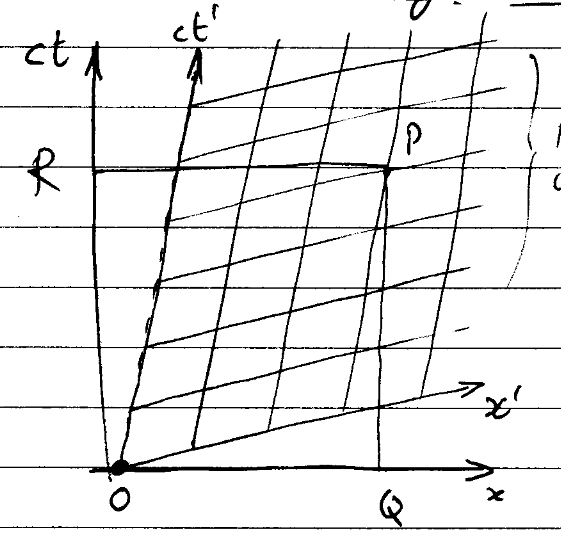
Asymptotically approach c

Notice that light signals from region I cannot reach region II and vice versa.

So the light worldlines are called "event horizons".

(For the interested: Look-up Event horizons in de Sitter space)

Relationship between space-time diagrams of inertial observers



Two frames, S & S', then from LT $t' = \gamma(t - \frac{vx}{c^2})$, $x' = \gamma(x - vt)$, $y' = y$, $z' = z$

$ct' = 0 \Leftrightarrow ct = \frac{vx}{c} = \left(\frac{v}{c}\right)x$

Similarly,

$x' = 0 \Leftrightarrow ct = \frac{cx}{v} = \left(\frac{c}{v}\right)x$

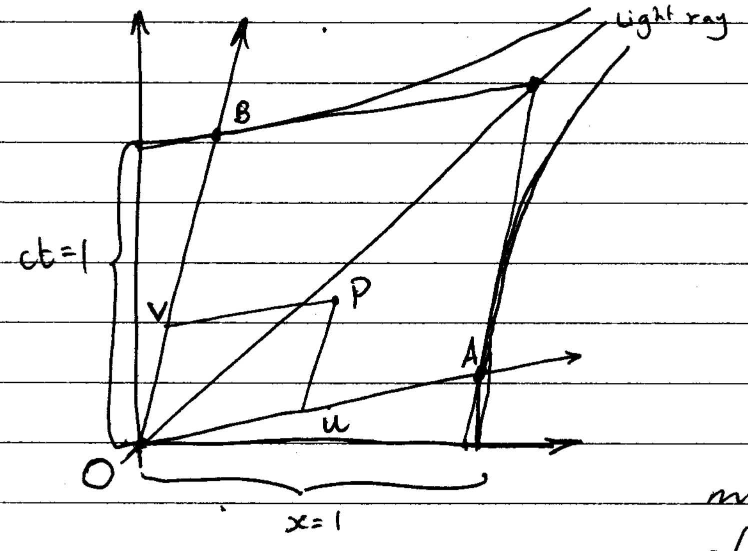
The lines parallel to Ox' (constant t') are called lines of simultaneity in S'.

How do we find the coordinate of point $P(t, x)$ in S' if the coordinate scales are different?

Use the Lorentz invariant, $c^2t^2 - x^2 = c^2t'^2 - x'^2$ although it is better to write the negative form & equal to 1

$$x^2 - c^2t^2 = x'^2 - c^2t'^2 = \pm 1$$

These equation(s) correspond to hyperbolae:



Thus

$$(ct', x') = \left(\frac{OV}{OB}, \frac{Ou}{OA}\right)$$

Note: book diagram is misleading in terms of slope of tangents to hyperbolae.

Important points:

(1) As shown in 3.1 boosts can be thought of as a rotation through an imaginary angle where $\tan\theta = \frac{iv}{c}$. This is thus equivalent to skewing the coordinate axes inward toward the light-like path, by the real angle θ.

(2) The hyperbolae are invariant across frames - they calibrate all frames.

(3) Length contraction & time dilation follow by drawing (projecting) the cross-points of the S' frame & the hyperbolae back to the t, x axes. Increasing skewness of the axes as $v \to c$ leads to stronger length contraction & time dilation.

(4) How big is the calibration factor OA?

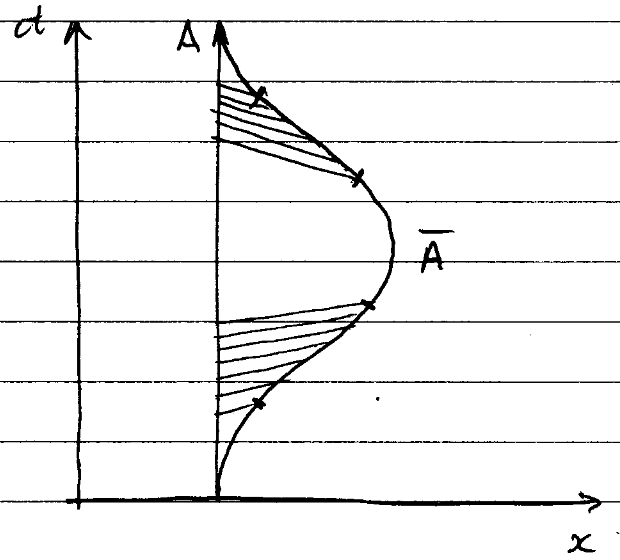
In S' A has coordinates $(0, 1)$, thus it has coords in S of $(ct, x) = (\gamma \frac{v}{c}, \gamma)$

$$\text{Thus} \quad OA = (c^2t^2 + x^2)^{1/2} = \left(\gamma^2 \frac{v^2}{c^2} + \gamma^2\right)^{1/2} = \gamma\left(1 + \frac{v^2}{c^2}\right)^{1/2}$$

$$= \left(\frac{1 + v^2/c^2}{1 - v^2/c^2}\right)^{1/2}$$

Twin Paradox via simultaneity lines

Paradox: Twins labelled A & $\bar{A}$. $\bar{A}$ leaves in spaceship under uniform acceleration. Reaches cruising velocity, decelerates & reverses direction, the follows same path home.
On arrival $\bar{A}$'s clock has elapsed less time.



Draw on lines of simultaneity, note the change around the reversal of direction.

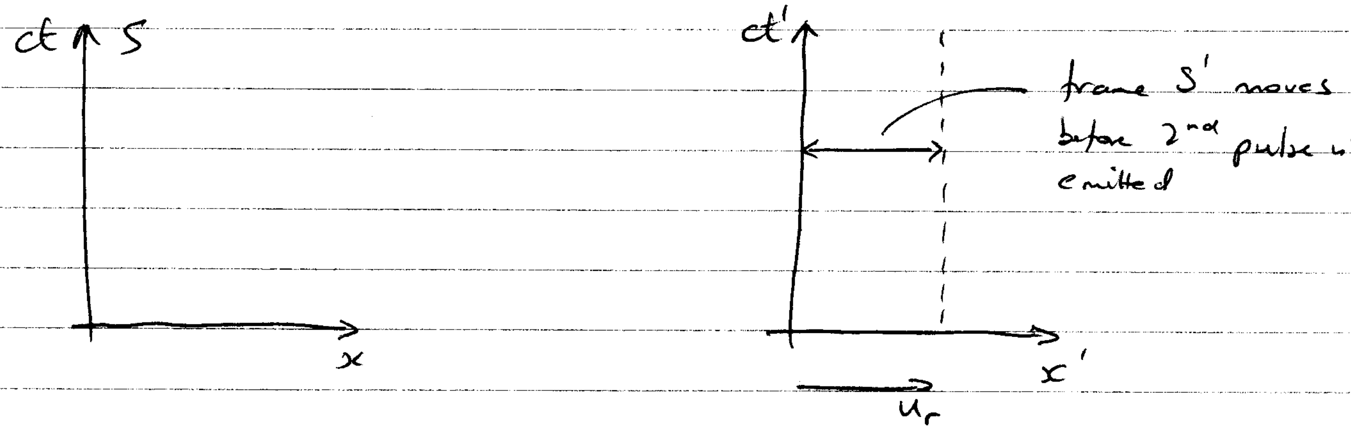
We've already discussed the resolution ($\bar{A}$ does not travel in an inertial frame while A does).

The implication that a longer path in spacetime takes less time to travel is because of the negative sign on the spatial coordinates in the Minkowski line element

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$$

Relativistic Doppler Effect

A source of light is at rest in frame S' & has wavelength λ_0. Frame S' travels at speed u_r relative to frame S. Pulses of light in S' are emitted, what does frame S observe?



Pulse in S' are separated by dt', what is this time period in S?

Suppose $dt' = t_2' - t_1'$ so $t_1' = t'$ & $t_2' = t' + dt'$

Then Inverse LT gives

$$t_2 = \gamma\left(t_2' + \frac{vx'}{c^2}\right) \qquad t_1 = \gamma\left(t_1' + \frac{vx'}{c^2}\right)$$

Thus

$$dt = t_2 - t_1 = \gamma\left(t_2' + \frac{vx'}{c^2}\right) - \gamma\left(t_1' + \frac{vx'}{c^2}\right)$$

$$= \gamma\, dt'$$

Also need additional delay for second pulse associated with movement of frame S' in time dt'

So need distance Δx in frame S

Inverse LT $\quad x = \gamma(x' + vt')$

Thus $\quad \Delta x = \gamma(x' + u_r t_2') - \gamma(x' + u_r t_1') = \gamma u_r dt'$

Time for light to travel $= \frac{\Delta x}{c} = \frac{\gamma u_r dt'}{c}$

Thus difference in pulse arrival for frame S:

$$\Delta t = dt + \frac{\Delta x}{c} = \gamma\, dt' \left(1 + \frac{u_r}{c}\right)$$

Using $\quad \lambda = c \Delta t \quad$ and $\quad \lambda_0 = c\, dt'$

$$\frac{\lambda}{\lambda_0} = \gamma\left(1 + \frac{u_r}{c}\right) = \left(\frac{1 + u_r/c}{1 - u_r/c}\right)^{1/2}$$

ie the k-factor formula ~~from~~ lecture 1.

Note: This is the <u>radial</u> Dopple shift.

Can also have transverse Dopple shift in a rotating frame.

Summary of relative nature of phenomena.

Theory	Position	Velocity	Time	Acceleration
Newtonian	Relative	Relative	Absolute	Absolute
SR	Relative	Relative	Relative	Absolute
GR	Relative	Relative	Relative	Relative

4-vector approach to Special Relativity

The 4-vectors are described using a basis of unit vectors $\underset{\sim}{e}_0, \underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3$. $\underset{\sim}{e}_0$ being the unit vector along the ct axis.

General form is the

$$\underset{\sim}{a} = a^0 \underset{\sim}{e}_0 + a^1 \underset{\sim}{e}_1 + a^2 \underset{\sim}{e}_2 + a^3 \underset{\sim}{e}_3$$

$$\underset{\sim}{a} = a^t \underset{\sim}{e}_t + a^x \underset{\sim}{e}_x + a^y \underset{\sim}{e}_y + a^z \underset{\sim}{e}_z$$

By definition components are written with a upper index, basis vectors with a lower index.

$$\therefore \quad \underset{\sim}{a} = \sum_{\alpha=0}^{3} a^\alpha \underset{\sim}{e}_\alpha$$

Labelling conventions:

Greek indices: run from 0 to 3
Roman indices: run from 1 to 3

Note: be careful, not all texts follow this rule & occassionally it may be broken.

Summation convention:

In an equation where ~~greek~~ letters are repeated this is assumed to imply summation over the repeated index

Thus, by definition

$$\sum_{\alpha=0}^{3} a^\alpha \underset{\sim}{e}_\alpha = a^\alpha \underset{\sim}{e}_\alpha$$

For 3-vectors

$$\sum_{i=1}^{3} a^i \underset{\sim}{e}_i = a^i \underset{\sim}{e}_i$$

Note: where repeated index occur we can replace this with another "dummy" index

$$a^\alpha \underset{\sim}{e}_\alpha = a^\beta \underset{\sim}{e}_\beta = a^\mu \underset{\sim}{e}_\mu$$

This is very important & used frequently in many derivations.

Simple 4-vector example:
Given 2 points in a 4D spacetime
$\underset{\sim}{x}_A = (ct_A, x_A, y_A, z_A)$ and $\underset{\sim}{x}_B = (ct_B, x_B, y_B, z_B)$
then

$$\Delta x^\alpha = x_B^\alpha - x_A^\alpha$$

describes the components of the displacement 4-vector.

Transformations of vectors

We've already mentioned that an LT can be written as

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = L \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}$$

where L is 4×4 matrix

In component form, we can write the elements of a Lorentz boost as ~~A~~ $L^{\mu}{}_{\nu}$ where μ is the y^{th} component, ν is the x^{th} component

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\frac{v\gamma}{c} & 0 & 0 \\ -\frac{v\gamma}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & \\ 0 & 0 & & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}$$

Then

$$x'^{\beta} = L^{\beta}{}_{\alpha} x^{\alpha}$$

In general Λ is used rather than L, to represent the components of the LT. We follow that from now on.

Mathematical notes: Scaler product of 4-vectors

$$\underset{\sim}{a} \cdot \underset{\sim}{b} = \underset{\sim}{b} \cdot \underset{\sim}{a}$$
$$\underset{\sim}{a} \cdot (\underset{\sim}{b} + \underset{\sim}{c}) = \underset{\sim}{a} \cdot \underset{\sim}{b} + \underset{\sim}{a} \cdot \underset{\sim}{c}$$
$$(\alpha \underset{\sim}{a}) \cdot \underset{\sim}{b} = \alpha (\underset{\sim}{a} \cdot \underset{\sim}{b})$$

~~$\forall \alpha$~~ where α is number

Scalar product of basis vectors:

Since $\underset{\sim}{a} \cdot \underset{\sim}{b} = a^{\alpha} \underset{\sim}{e}_{\alpha} \cdot b^{\beta} \underset{\sim}{e}_{\beta}$ (two implied summations here)

$$= a^{\alpha} b^{\beta} (\underset{\sim}{e}_{\alpha} \cdot \underset{\sim}{e}_{\beta})$$

Define $\eta_{\alpha\beta} = \underset{\sim}{e}_{\alpha} \cdot \underset{\sim}{e}_{\beta}$

Since $\underset{\sim}{a}\cdot\underset{\sim}{b} = \underset{\sim}{b}\cdot\underset{\sim}{a}$ we must have

$$a^\alpha b^\beta (\underset{\sim}{e}_\alpha\cdot\underset{\sim}{e}_\beta) = a^\alpha b^\beta (\underset{\sim}{e}_\beta\cdot\underset{\sim}{e}_\alpha)$$

$$\therefore \quad \eta_{\alpha\beta} = \eta_{\beta\alpha} \quad \Rightarrow \eta_{\alpha\beta} \text{ is symmetric}$$

We can constrain the values of $\eta_{\alpha\beta}$ using the Minkowski line element. By definition, the square of the displacement vector describes the square of the distance in 4D spacetime.

$$(\Delta s)^2 = \Delta\underset{\sim}{x}\cdot\Delta\underset{\sim}{x} = \Delta x^\alpha \Delta x^\beta\, \underset{\sim}{e}_\alpha\cdot\underset{\sim}{e}_\beta$$

$$\Rightarrow \quad (\Delta s)^2 = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$$

Since there are no terms that are products of $\Delta x^\alpha \Delta x^\beta$ where $\alpha \neq \beta$ then the only non-zero elements are the diagonal elements $\eta_{\alpha\alpha}$ (no summation implied).

We can then read off the diagonal elements as $(\cancel{0}1, -1, -1, -1)$

$$\therefore \quad \eta_{\alpha\beta} \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Clearly $\underset{\sim}{a}\cdot\underset{\sim}{b} = a^\alpha b^\beta \eta_{\alpha\beta} = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3$

Since $\eta_{\alpha\beta}$ effectively describes the measurement of distance in 4D spacetime it is the called the metric. As we have already see

$$(ds)^2 = \eta_{\alpha\beta}\, dx^\alpha\, dx^\beta$$

In SR. $\eta_{\alpha\beta}$ is very simple, however in GR it can be highly complex.

Note: Definitions of $\eta_{\alpha\beta}$ vary

D'Inverno: $$\eta_{\alpha\beta} = \begin{bmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{bmatrix}$$

Hartle $$\eta_{\alpha\beta} = \begin{bmatrix} -1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{bmatrix}$$

The number of +ve & -ve signs is known as the signature of the metric, ie (1, 3) or (3, 1). This freedom of choice arises due to a factor of -1 that can multiply both sides of distance definitions:

ie. if $(\Delta s)^2 = 1$
then $-(\Delta s)^2 = -1$

Classification of 4-vectors

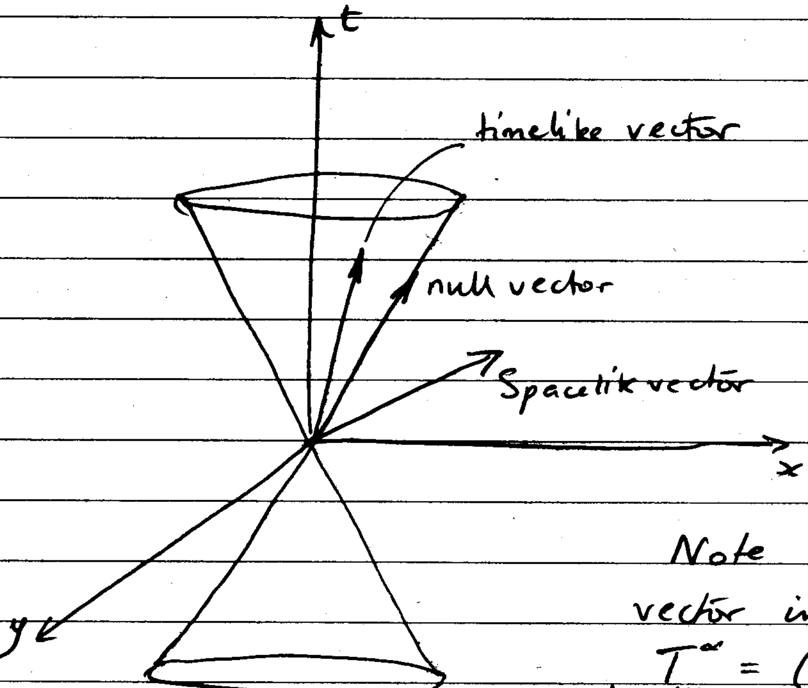
The square of the length, or norm of a 4-vector is given by

$$x^2 = \eta_{\alpha\beta} x^\alpha x^\beta = \underset{\sim}{x} \cdot \underset{\sim}{x}$$

The vector is is then classified according to:

timelike	if	$x^2 > 0$
spacelike	if	$x^2 < 0$
null or lightlike	if	$x^2 = 0$

Orthogonal vectors have 0 scalar product. Thus all null vectors are orthogonal to themselves.



Note if we define a unit vector in the time direction $T^\alpha = (1,0,0,0)$ then timelike vectors are classified by

future pointing	if	$\eta_{\alpha\beta} x^\alpha T^\beta > 0$
past pointing	if	$\eta_{\alpha\beta} x^\alpha T^\beta < 0$

4-vector description of mechanics

We can describe 4D worldlines by $x^\alpha = x^\alpha(\sigma)$ where σ is a parameter describing the evolution.

The most natural choice for σ is the proper time τ

$$\therefore \quad x^\alpha = x^\alpha(\tau)$$

The derivative wrt to τ then defines the four-velocity $u^\alpha(\tau)$

$$u^\alpha(\tau) \equiv \frac{dx^\alpha(\tau)}{d\tau}$$

$u^\alpha(\tau_A)$ is clearly tangent to the point $x^\alpha(\tau_A)$.

The components of $u^\alpha(\tau)$ are given by:

$$u^t(\tau) = c\frac{dt}{d\tau} = c\gamma$$ from the definition of proper time in lecture 2.

$$u^i(\tau) = \frac{dx^i}{d\tau} = \frac{dt}{d\tau}\frac{dx^i}{dt} = \gamma v_3^i$$

Thus $\quad u^\alpha \equiv (c\gamma, \gamma \vec{v_3})$ where $\vec{v_3}$ is the three-velocity.

Important point:

$$\underset{\sim}{u} \cdot \underset{\sim}{u} = c^2\gamma^2 - \gamma^2 v^2 = c^2\left\{1-\frac{v^2}{c^2}\right\}\left\{1-\frac{v^2}{c^2}\right\}^{-1}$$

$$= c^2$$

$\therefore$ $\underset{\sim}{u} \cdot \underset{\sim}{u}$ is a Lorentz invariant.

Note that if $\underset{\sim}{u}.\underset{\sim}{u} = c^2$

$$\underset{\sim}{u}.\underset{\sim}{u} = \eta_{\alpha\beta} \frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = \left(\frac{ds}{d\tau}\right)^2$$

implies

$$ds^2 = c^2 d\tau^2$$

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

Which is equivalent to our earlier definition of proper time.

Note since $\underset{\sim}{u}.\underset{\sim}{u} = c^2$ it is a timelike vector.

<u>Newton's First Law in SR</u>

$$\boxed{\frac{d\underset{\sim}{u}}{d\tau} = 0}$$

Implies that $\vec{v}_3$ must be constant in an inertial frame.

<u>Newton's Second Law in SR</u>

$$\boxed{m\frac{d\underset{\sim}{u}}{d\tau} = \underset{\sim}{f}}$$

What m do we use? Most straightforward (& correct) idea is to use the rest mass. Note that the rest mass is an <u>invariant across inertial frames</u>.

If we define the 4-acceleration in the logical way,

$$\underset{\sim}{a} \equiv \frac{d\underset{\sim}{u}}{d\tau}$$

Then

$$\underset{\sim}{f} = m\underset{\sim}{a}$$

Note since $\underset{\sim}{u}\cdot\underset{\sim}{u} = c^2$

$$m\frac{d}{d\tau}(\underset{\sim}{u}\cdot\underset{\sim}{u}) = 0 \Rightarrow m\underset{\sim}{u}\cdot\underset{\sim}{a} = 0 \Rightarrow \underset{\sim}{u}\cdot\underset{\sim}{f} = 0$$

This implies that there are not 4 independent components in $\underset{\sim}{f}$ & infact $\underset{\sim}{f} = m\underset{\sim}{a}$ describes 3 independent equations.

<u>Four momentum</u>

Defined by $\underset{\sim}{p} = m\underset{\sim}{u}$ then

$$\frac{d\underset{\sim}{p}}{d\tau} = \underset{\sim}{f}$$

Since $\underset{\sim}{u}\cdot\underset{\sim}{u} = c^2$, we have $\underset{\sim}{p}\cdot\underset{\sim}{p} = m^2c^2$

Given $\underset{\sim}{u} = (c\gamma, \gamma\vec{v}_3)$ from earlier, $\underset{\sim}{p} = (mc\gamma, m\gamma\vec{v}_3)$

For the t-component

$$p^t = mc\gamma = mc\left(1 - \frac{v^2}{c^2}\right)^{-1/2}$$

Taylor expand to get

$$p^t = \frac{1}{c}\left\{ mc^2 + \frac{1}{2}mv^2 + \text{h.o.t.} \right\} = \frac{1}{c}E$$

by defn.

$$p^i = \gamma m v_3^i = m v_3^i + \text{h.o.t.}$$

Writing out components of $\underline{p}\cdot\underline{p}$:

$$(p^0)^2 - (p^1)^2 - (p^2)^2 - (p^3)^2 = m^2c^2$$

$$\Rightarrow \quad \frac{1}{c^2}E^2 = m^2c^2 + m^2\gamma^2\vec{v}_3^{\,2}$$

$$\therefore \quad E^2 = (m^2c^4 + c^2\vec{p}^{\,2})$$

where we have <u>defined</u> $\vec{p} = m\gamma\vec{v}_3$

Note the equivalence $\vec{p} = m\gamma\vec{v}_3$ leads to the idea of "relativistic mass":

Write $m_r = \gamma m$ and then

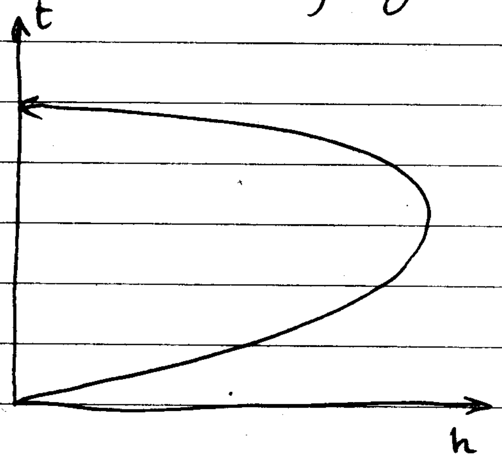
$$\vec{p} = m_r\vec{v}_3$$

This is a potentially confusing terminology which we will avoid. It is better to think of p^0 as defining the total energy, which the relativistic mass is just a proxy for.

The Key idea of General Relativity

The absence of any other force than gravity particles follow a unique worldline in spacetime

Equivalently, knowing paths of freely-falling objects ≡ knowing gravity



Parabola of a ball thrown into the air. In GR the particle is following a straight path in the curved spacetime produced by Earth.

"Gravity is Geometry"

Traditional approach to geometry is Euclidean. Begin with Euclids postulates (ie there exists a unique line between two points) and then derive properties of a 2D plane.

Alternatively, we can think of geometry as a experimental science. What would we measure?

(1) Circumference of a circle compared to radius
(2) What do the angles in a triangle add up to?
(3) Do initially parallel lines ever cross?
(4) Is the geometry the same in all directions? isotropy
(5) " " " " " " " places? homogeneity

For a 2D plane:

$C = 2\pi R$
$\sum$ angles $= \pi$
parallel lines never meet
space is homogeneous & isotropic

You can actually encapsulate all of these measurements into Euclids 5 postulates.

The best way to encapsulate the geometry of a surface is look at how the distance between nearby points changes.

This is "differential geometry"

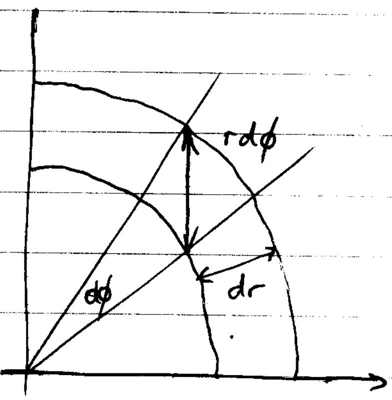
For the Euclidean geometry the line element is ds, or equivalently

ds_{12} = distance between nearby points

$= (dx^2 + dy)^{1/2}$

Or in polar coordinates $x = r\cos\phi$, $y = r\sin\phi$

$= (dr^2 + r^2 d\phi^2)^{1/2}$



We can test homogeneity very quickly

$y \quad x' = x + a \qquad y' = y + b$

$\Rightarrow \quad dx' = dx \quad , \quad dy' = dy$

$\therefore \quad ds'_{12} = ds_{12}$

For isotropy, we rotate the coordinates

$$x' = x\cos\theta + y\sin\theta$$
$$y' = -x\sin\theta + y\cos\theta$$

It's straightforward to show that

$$ds'_{12} = ds_{12}$$

What about circumference of circle?

Consider $x^2 + y^2 = R^2$

$$C = \oint ds = \oint (dx^2 + dy^2)^{1/2}$$

$$= 2\int_{-R}^{+R} dx \left(1 + \left(\frac{dy}{dx}\right)^2\right)^{1/2}_{x^2+y^2=R^2}$$

$$= 2\int_{-R}^{R} dx \sqrt{\frac{R^2}{R^2 - x^2}}$$

which follows

$x^2 + y^2 = R^2$

$\Rightarrow x\,dx + y\,dy = 0$

$\Rightarrow \dfrac{dy}{dx} = -\dfrac{x}{y} = \dfrac{-x}{(R^2 - x^2)^{1/2}}$

Change variables $x = R\xi$

$$\Rightarrow \quad C = 2R\int_{-1}^{1} \frac{d\xi}{\sqrt{1-\xi^2}} = 2\pi R.$$

The result is even easier in polar coordinates

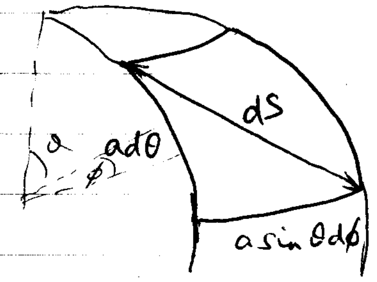
$$C = \oint dS = \int_0^{2\pi} R\, d\phi = 2\pi R$$

Note: This trivial example beautifully demonstrates how some coordinate systems are better than others for a given problem.

What about non-Euclidean geometries?

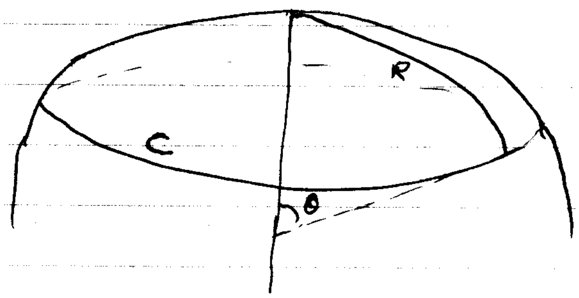
Easiest example is the surface of a sphere of radius a. (θ, ϕ) polar coordinates can label points.

What is the line element?



So $\quad ds^2 = a^2(d\theta^2 + \sin^2\theta\, d\phi^2)$

Defining circles as loci of points a fixed distance from another point, we then orient our coordinate system so that the polar axis is centered on the circle.



For a circle $\theta = \Theta \quad d\theta = 0$

$$\therefore \quad C = \oint dS = \int_0^{2\pi} a \sin\Theta \, d\phi$$

$$= 2\pi a \sin\Theta$$

Radius is distance from point to circle

$$r = \int_{\text{center}}^{\text{circle}} dS = \int_0^{\Theta} a \, d\theta = a\Theta$$

Eliminating Θ we find

$$C = 2\pi a \sin\left(\frac{r}{a}\right)$$

$$\simeq 2\pi r \left(1 - \frac{1}{3}\left(\frac{R^2}{a^2}\right) + \dots\right)$$

So for large $\frac{a}{r}$, ie small r/a this reduces to Euclidean geometry.

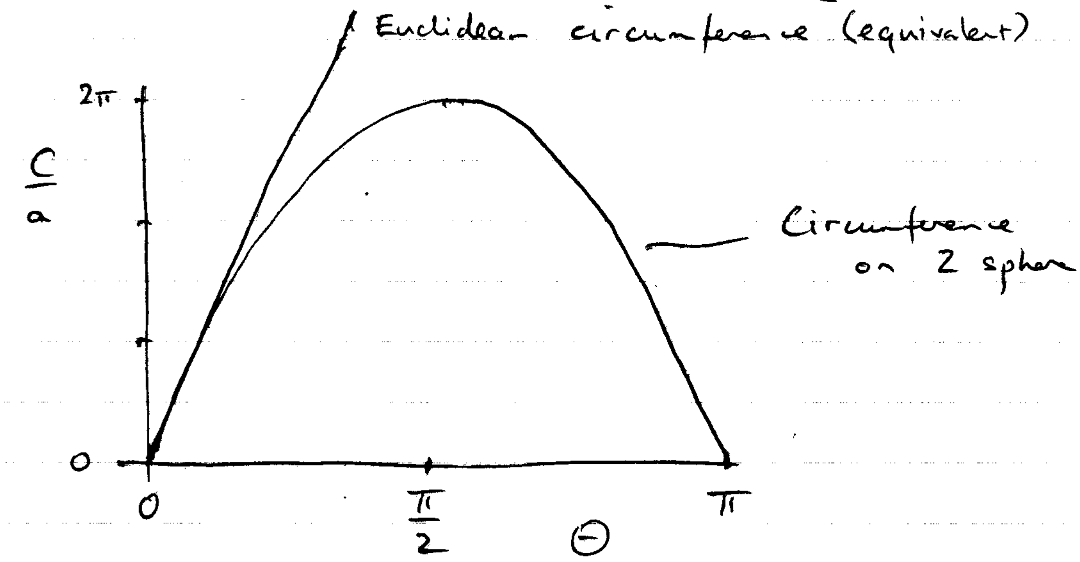
Also find that

Sum of angles $= \pi + \frac{\text{area}}{a^2}$

Parallel lines meet in distance $\frac{\pi a}{2}$

Line element is homogeneous & isotropic.

Interesting point: The circumference of circles reaches a maximum at $\Theta = \frac{\pi}{2}$



Under general coordinate transformation:

$$\theta = f(\theta', \phi') \qquad \phi = g(\theta', \phi')$$

then

$$d\theta = \frac{\partial f}{\partial \theta'} d\theta' + \frac{\partial f}{\partial \phi'} d\phi'$$

Substitution gives

$$ds^2 = a^2 \left[\left(\left(\frac{\partial f}{\partial \theta'} \right)^2 + \sin^2(f(\theta', \phi')) \left(\frac{\partial g}{\partial \phi'} \right)^2 \right) d\theta'^2 + \ldots \right]$$

which clearly demonstrates this can be non-trivial!

We can still attempt to project the sphere's geometry into a plane, but there is of course no transformation that will produce $ds^2_{sphere} \to ds^2_{plane}$.

This is analogous to map making, & there are occasions in GR where we want to map the infinite structure of spacetime (Penrose Diagrams).

As an example, let us map the sphere into a plane.

For familiarity use lattitude & longitude

(lattitude) $\lambda = \frac{\pi}{2} - \theta \Rightarrow \sin\theta \equiv \cos\lambda$

Then $ds^2 = a^2(d\theta^2 + \sin^2\theta\, d\phi^2)$

$$= a^2(d\lambda^2 + \cos^2\lambda\, d\phi^2)$$

Introduce new coordinates $x = x(\lambda, \phi)$, $y = y(\lambda, \phi)$

Different projections $\equiv$ different functions for x & y

Simplest example:

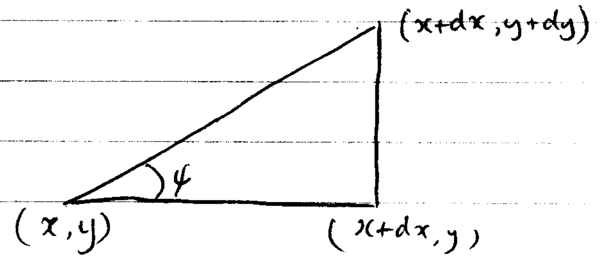
$x = \frac{L\phi}{\pi}$ $\quad y = \frac{L\lambda}{\pi}$ where L is "size" of map

x extent will be $2L$, y extent L.

Substitute for $d\lambda$ & $d\phi$ to get

$$ds^2 = \left(\frac{\pi a}{L}\right)^2 \left(dy^2 + \cos^2\left(\frac{\pi y}{L}\right) dx^2\right)$$

Does this transformation preserve the angles on a sphere when mapped into the plane?



In y-direction

$$dS_y = \frac{\pi a}{L}\, dy$$

In x-direction

$$dS_x = \frac{\pi a}{L} \cos\left(\frac{\pi y}{L}\right) dx$$

So $\tan \psi = \dfrac{dS_x}{dS_y} = \dfrac{dy}{\cos\left(\frac{\pi y}{L}\right) dx}$

$y = \frac{L}{2} \Rightarrow$ equator $\Rightarrow \cos\left(\frac{\pi y}{L}\right) = 1$

So angles are preserved on equator but elsewhere they are modified.

What about other projections?

Consider the following

$$x = \frac{L\phi}{2\pi} \qquad y = y(\lambda) \qquad [\lambda = \lambda(y)]$$

$\Rightarrow \quad d\phi = \frac{2\pi}{L}\, dx$ but leave $d\lambda = \frac{d\lambda}{dy} dy$

On substituting

$$ds^2 = a^2\left[\left(\frac{2\pi}{L}\cos\lambda(y)\right)^2 dx^2 + \left(\frac{d\lambda}{dy}\right)^2 dy^2\right] \quad \text{—— (1)}$$

Kramer (1569) – Is there a way to map the functions so as to preserve angles? Yes, the Mercator projection is one example (which Kramer found).

To preserve angles we need $\tan\psi = \frac{dy}{dx}$

A line element of the form

$$ds^2 = \Omega^2(x,y)\,[dx^2 + dy^2]$$

will do this for us!

Therefore, for (1) above we need

$$\frac{d\lambda}{dy} = \frac{2\pi}{L}\cos\lambda(y)$$

$$\therefore \quad y(\lambda) = \int^{\lambda} \frac{d\lambda'}{\cos\lambda'} = \frac{L}{2\pi}\ln\left[\tan\left(\frac{\pi}{4} + \frac{\lambda}{2}\right)\right]$$

together with $x = \frac{L\phi}{2\pi}$, this defines the <u>Mercator projection</u>.

The line element is

$$ds^2 = \left[\frac{2\pi a}{L} \cos \lambda(y)\right]^2 (dx^2 + dy^2)$$

Notice what happens to areas:

$$dS_x dS_y = (\Omega(y)\Delta x)(\Omega(y)\Delta y) = \Omega^2(y)\,\Delta x \Delta y$$

This explains why Greenland looks so large when compared with South America, as one travels towards the poles the $\Omega(y)$ factor goes to zero. ⇒ Thus the y positions become more & more extended.

More general surfaces (Example)

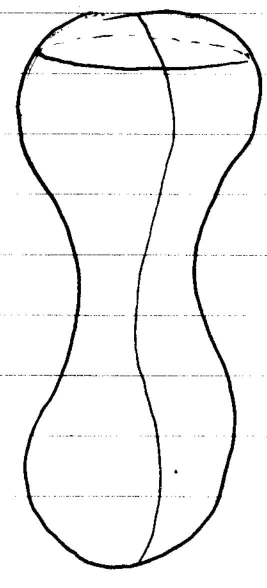
Hartle gives an interesting example of a "peanut" geometry

$$dS^2 = a^2(d\theta^2 + f^2(\theta)\,d\phi^2)$$

let $$f(\theta) = \sin\theta\left(1 - \frac{3}{4}\sin^2\theta\right)$$

$$\text{Circumference} = \int_0^{2\pi} a f(\theta)\,d\phi$$
$$= 2\pi a f(\theta)$$
$$= 2\pi a \sin\theta\left(1 - \frac{3}{4}\sin^2\theta\right)$$

Pole to Pole distance $= a\int_0^{\pi} d\theta = \pi a$



Tensor Algebra

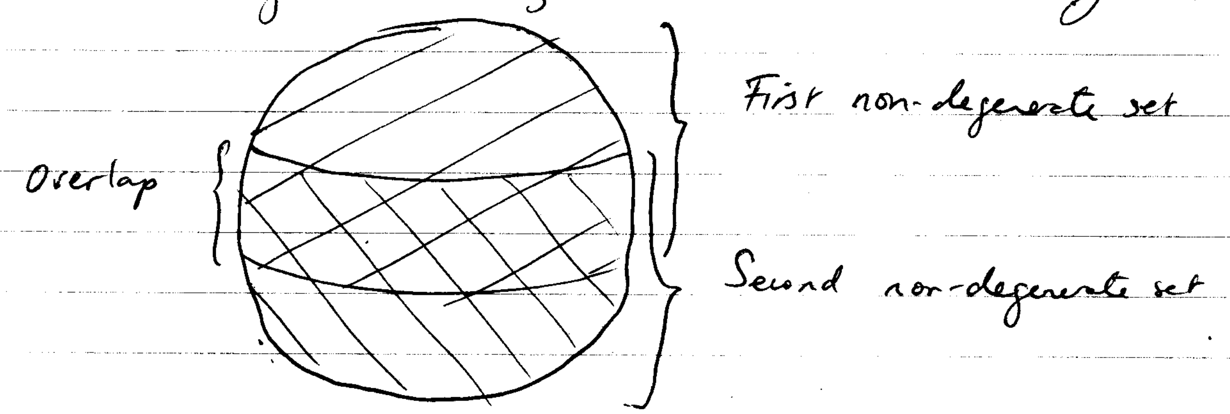
While Newtonian mechanics can be described effectively using vectors, General Relativity requires us to use tensors. The discussion of tensor algebra can be undertake via an abstract (index free) approach, or alternatively a more conventional approach based on indices. We'll use the latter approach.

Manifolds

Tensors are defined on differential manifolds. Manifolds appear locally like Euclidean space ($\mathbb{R}^n$) but are globally different (note that the global properties define the topology of the manifold).

For the purpose of the course, an n-dimensional manifold is a set of points, which each point labelled $(x^1, x^2, \dots, x^n)$. It may not be possible to cover the manifold with a single non-degenerate coordinate system (NOTE: polar coordinates are degenerate at $(r=0, \theta)$).

Example The sphere cannot be covered by a single non-degenerate coordinate system.



Coordinate systems the cover only a portion of the manifold are called patches. A set of patches that covers the manifold is an Atlas.

In the overlaps we can convert between different coordinate sets. Describing how geometric quantities behave under the coordinate transformation is the underlying goal of tensor calculus.

Defining curves & surfaces on a manifold

Parametric approach:
A curve has a single parameter u, and defines coordinates in an n-dimensional manifold via

$$x^a = x^a(u) \qquad (a = 1, \dots, n)$$

A surface (with two degrees of freedom) is

$$x^a = x^a(u, v)$$

in general for a surface with $m < n$ degrees of freedom it can be defined by

$$x^a = x^a(u^1, \dots, u^m)$$

Such a surface may be called a subspace, & surfaces with $m = n-1$ degrees of freedom are called hypersurfaces e.g. $x^a(u^1, \dots, u^{n-1})$

Constraint approach:
Beginning from $x^a = x^a(u^1, \dots, u^{m-1})$ we could eliminate all u^1 to u^{m-1}; (where $m = n-1$) to get

$$f(x^1, x^2, \dots, x^n) = 0$$

This is equivalent to simply specifying the constraint. The constraint must obviously describe how the n^{th} degrees of freedom are reduced to $n-1$ (simple example: $x^2+y^2=R^2$ in $\mathbb{R}^2$).

More constraints are required to further reduce the dimension of the subspace. An m-dimensional subspace must have $n-m$ constraints:

$$f^1(x^1,\dots,x^n) = 0$$
$$f^2(x^1,\dots,x^n) = 0$$
$$\vdots$$
$$f^{n-m}(x^1,\dots,x^n) = 0$$

This is equivalent to the parametric approach.

<u>Coordinate Transforms</u>

(NOTE: Notation is VERY important in tensor algebra. It is the key to most derivations.)

Tensor equations are coordinate covariant (as we'll see later). This necessarily implies we must consider coordinate transformations. Consider a change of coordinates described by

$$x^a \rightarrow x'^a = f^a(x^1,\dots,x^n) \qquad (a=1,\dots,n)$$

Note: It is a requirement that f be at least singly differentiable for all its parameters. Note for ease of notation we will write

$$x^a \rightarrow x'^a = f^a(x^1,\dots,x^n) = f^a(x) = x'^a(x)$$

Since we can differentiate all the f^a wrt x^b we can define the matrix

$$\left[\frac{\partial x'^a}{\partial x^b}\right] = \begin{bmatrix} \frac{\partial x'^1}{\partial x^1} & \frac{\partial x'^1}{\partial x^2} & \cdots & \frac{\partial x'^1}{\partial x^n} \\ \frac{\partial x'^2}{\partial x^1} & \frac{\partial x'^2}{\partial x^2} & \cdots & \frac{\partial x'^2}{\partial x^n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x'^n}{\partial x^1} & \frac{\partial x'^n}{\partial x^2} & \cdots & \frac{\partial x'^n}{\partial x^n} \end{bmatrix}$$

You should recognize that the determinant of this matrix is the Jacobian of the coordinate transform

$$J' = \left|\frac{\partial x'^a}{\partial x^b}\right|$$

Provided J' is non-singular we can invert the coordinate transform to get

$$x^a = x^a(x')$$

and for this transform $\quad J = \left|\frac{\partial x^a}{\partial x'^b}\right| = \frac{1}{J'}$

Under $x^a \longrightarrow x'^a(x)$ the Chain Rule gives

$$dx^a = \sum_{b=1}^{n} \frac{\partial x^a}{\partial x'^b}\, dx'^b$$

$$= \frac{\partial x^a}{\partial x'^b}\, dx'^b \qquad \text{(Summation convention)}$$

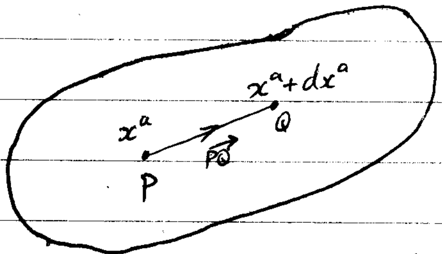
Helpful definition: The Kronecker delta δ^a_b is given by

$$\delta^a_b = \begin{cases} 1 & \text{if } a=b \\ 0 & \text{if } a\neq b \end{cases}$$

then

$$\frac{\partial x'^a}{\partial x'^b} = \frac{\partial x^a}{\partial x^b} = \delta^a_b$$

Definition of Contravariant Tensors



Two points in a coordinate patch are separated by an infinitesimal displacement vector $\vec{PQ}$.

Note: The vector $\vec{PQ}$ is associated with the point P, it is not "freely floating."

For two different coordinate systems

$$\vec{PQ} \equiv dx^a \equiv dx'^a$$

by the Chain Rule

$$dx'^a = \left[\frac{\partial x'^a}{\partial x^b}\right]_P dx^b$$

Subscript P emphasizes the transform is associated with the point P. From now on we drop the P.

This is the prototype definition for a contravariant tensor of rank 1, written X^a:

$$X'^a = \frac{\partial x'^a}{\partial x^b} X^b$$

(also called a contravariant vector)

Since X'^a is associated with the point P it is distinct from the tangent vector $\frac{dx^a}{du}$ on a curve $x^a(u)$. (One is a coordinate dependent object at P while the other is a tangent vector along a curve).

A contravariant tensor of rank 2 transforms according to

$$X'^{ab} = \frac{\partial x'^a}{\partial x^c} \frac{\partial x'^b}{\partial x^d} X^{cd}$$

[Note: positions of indices matter]

Example: The product of two rank 1 contravariant tensors, Y^a, Z^a is $Y^a Z^b$ and it defines a rank 2 contravariant tensor.

Note: a rank zero tensor is a scalar, and

$$\phi' = \phi$$

Covariant tensors

For a scalar $\phi = \phi(x^a(x'))$ by function of a function we can write

$$\frac{\partial \phi}{\partial x'^b} = \frac{\partial \phi}{\partial x^a} \frac{\partial x^a}{\partial x'^b}$$

which can be re-labelled

$$\frac{\partial \phi}{\partial x'^a} = \frac{\partial \phi}{\partial x^b} \frac{\partial x^b}{\partial x'^a} = \frac{\partial x^b}{\partial x'^a} \frac{\partial \phi}{\partial x^b}$$

We see that the transform involves $\frac{\partial x^b}{\partial x'^a}$ the inverse coordinate transform.

This is the prototype definition of how a ~~contra~~covariant tensor of rank 1 transforms

$$X'_a = \frac{\partial x^b}{\partial x'^a} X_b$$

The rank n transformation follows in the natural manner.
To remember the position of the index
" co goes below"

The distinction between how dx^a and $\frac{\partial}{\partial x^a}$ transform is at the heart of the co-contravariant distinction, and explains why we always write coordinate components with an upper index (although it is the differentials that are tensorial in nature).

Mixed tensors

Objects containing both contravariant & covariant parts can be defined, e.g.

$$X'^{a}{}_{bc} = \frac{\partial x'^a}{\partial x^d}\frac{\partial x^e}{\partial x'^b}\frac{\partial x^f}{\partial x'^c} X^d{}_{ef}$$

this would be a type $(1,2)$ tensor.

We can now demonstrate the coordinate invariance of tensor equations:

Consider $X_{ab} = Y_{ab}$
transform both sides by multiplying by $\frac{\partial x^a}{\partial x'^c}\frac{\partial x^b}{\partial x'^d}$

$$\therefore \quad \frac{\partial x^a}{\partial x'^c}\frac{\partial x^b}{\partial x'^d} X_{ab} = \frac{\partial x^a}{\partial x'^c}\frac{\partial x^b}{\partial x'^d} Y_{ab}.$$

$$\Rightarrow \quad X'_{ab} = Y'_{ab} \qquad \text{(after relabelling)}$$

Tensor fields

To date, we have strictly been speaking about tensors defined at a single point.
Tensor fields are defined over regions of a manifold, i.e.

$$T^{a\ldots}_{b\ldots}(P)$$

would define the tensor at point P.

In practice we deal mostly with tensor fields, which transform at a given point in the manners already discussed.
Note: A useful property of tensor fields is that they be smooth: all components are differentiable to all orders.

Operations on tensors

It should be clear that

$$X^a{}_{bc} = Y^a{}_{bc} + Z^a{}_{bc}$$

is an allowed equation under the transformation rules. Subtraction & scalar multiplication follow as well.

Symmetry properties of type 2 tensors

A type 2 tensor (either contra or covariant) is symmetric if

$$X_{ab} = X_{ba}$$

This necessarily implies that there are $\frac{1}{2}n(n+1)$ independent components.

A tensor is anti-symmetric (or skew-symmetric) if

$$X_{ab} = -X_{ba}$$

which necessarily implies there are $\frac{1}{2}n(n-1)$ independent components.

Given a general type-2 tensor we can define the symmetric part of the tensor $X_{(ab)}$ where

$$X_{(ab)} = \frac{1}{2}(X_{ba} + X_{ab})$$

The anti-symmetric part is denoted $X_{[ab]}$ and is given by

$$X_{[ab]} = \frac{1}{2}(X_{ab} - X_{ba})$$

Clearly $X_{[ab]} + X_{(ab)} = X_{ab}$.

The generalization for higher rank tensors is

$$X_{(a_1 a_2 \dots a_r)} = \frac{1}{r!}\left\{ \text{sum over all permutations of indicies } a_1 \text{ to } a_r \right\}$$

$$X_{[a_1 a_2 \dots a_r]} = \frac{1}{r!}\left\{ \text{alternating sum over all permutations of the indicies } a_1 \text{ to } a_r \right\}$$

<u>Example:</u> For a rank 3 covariant tensor X_{abc}

$$X_{[abc]} = \frac{1}{6}\left\{ X_{abc} - X_{acb} + X_{cab} - X_{cba} + X_{bca} - X_{bac} \right.$$

<u>Contracted Tensors & Compound Tensors</u>

We've already seen an example of a rank 2 tensor constructed from two rank 1 tensors. In general a type (p_1, q_1) tensor multiplied by a type (p_2, q_2) tensor gives a type $(p_1 + p_2, q_1 + q_2)$ tensor:

<u>Example</u> $\quad X^a{}_{bcd} = Y^a{}_b Z_{cd}$

We can reduce the rank of tensors by the contraction operation:

$$X^a{}_{bcd}\, \delta^b_a = X^a{}_{acd} \equiv Y_{cd}$$

Where δ^b_a is the Kronecker delta.

Relation to Index-free Notation

We have emphasized the difference between a geometrical object (tensor) & its components in a coordinate system. We can relate the approaches by considering a contravariant vector field as an operator on real valued functions (which maps to real valued functions).

For the tangent vector $\frac{dx^\alpha(u)}{du}$

$$\frac{dx^\alpha(u)}{du} \equiv \frac{d}{du} x^\alpha(u) \equiv \frac{dx^\beta}{du}\frac{\partial x^\alpha(u)}{\partial x^\beta}$$

we can view the tangent vector V as being defined by $\frac{d}{du} \equiv \frac{dx^\beta}{du}\frac{\partial}{\partial x^\beta}$

We then associate/relate the $\frac{\partial}{\partial x^\beta}$ as the basis of the tangent vector.

The result extends to any function & the basis of contravariant vectors is $\{\frac{\partial}{\partial x^\beta}\}$.

Thus in component form the ~~tangent~~ contravariant vector V is

$$V = X^a \frac{\partial}{\partial x^a} \equiv X^a \partial_a \qquad \left(\partial_a = \frac{\partial}{\partial x^a} \right)$$

We now examine how this object transforms under coordinate transformations.

$$X'^a \partial'_a = X'^a \frac{\partial}{\partial x'^a}$$

$$= \frac{\partial x'^a}{\partial x^b} X^b \frac{\partial x^c}{\partial x'^a} \frac{\partial}{\partial x^c}$$

$$= \frac{\partial x'^a}{\partial x^b} \frac{\partial x^c}{\partial x'^a} X^b \frac{\partial}{\partial x^c}$$

We can use the following identity

$$\delta^a_b = \frac{\partial x^a}{\partial x^b} = \frac{\partial}{\partial x^b} x^a(x'^c(x^d)) = \frac{\partial x'^c}{\partial x^b} \frac{\partial x^a}{\partial x'^c}$$

$$\therefore \quad X'^a \partial'_a = \frac{\partial x'^a}{\partial x^b} \frac{\partial x^c}{\partial x'^a} X^b \frac{\partial}{\partial x^c} = \delta^c_b X^b \partial_c$$

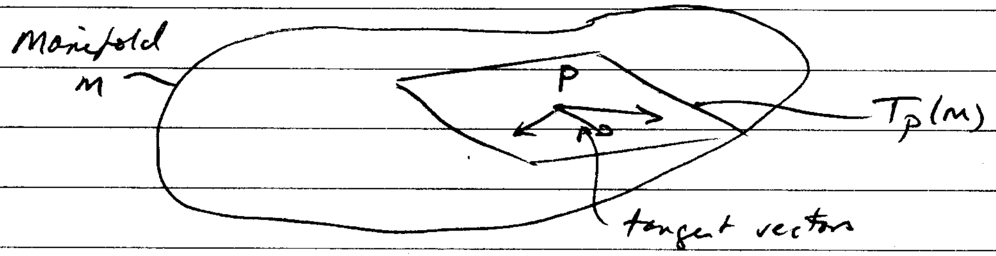
$$= X^c \partial_c = X^a \partial_a$$ after relabelling

So regardless of the coordinate system X defines the same operation.

NOTE: The basis elements transform in the opposite nature to the component!

The vector space of all the contravariant vectors a P is known as the <u>tangent space</u> at P & is written $T_P(M)$.
The tangent space is usually different from the underlying manifold, although Euclidean space & Minkowski spacetime have tangent spaces that "lie" in the manifold.

Diagrammatically:



Lie bracket

Given a vector fields defined in terms of functions, we can define a new vector field:

$$[X, Y] = XY - YX$$

which is the Lie bracket or commutator.

Straightforward algebra (exercise) shows that

$$[X, Y]f = (X^b \partial_b Y^a - Y^b \partial_b X^a)\partial_a f$$
$$- X^a Y^b (\partial_a \partial_b f - \partial_b \partial_a f)$$

Assuming commutativity of partial derivatives on f

$$[X, Y]^a = Z^a = X^b \partial_b Y^a - Y^b \partial_b X^a$$

It is straightforward to demonstrate the following:

$$[X, X] \equiv 0$$
$$[X, Y] \equiv -[Y, X]$$
$$[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$$

– the Jacobi identity

Quick note on co vs contravariant terminology.

Covariant - a characteristic being measured increases/decreases with an increase/decrease of the underlying parameter.
e.g. Suppose you measure a sharp spatial gradient in temperature, say $5\,K\,m^{-1}$
As distance increases so the temperature increases also. The characteristic is covariant.

Contravariant: If we're shown that temperature derived through the $5\,K\,m^{-1}$ gradient is covariant, then $1/\text{temp}$ has the opposite variation (gets smaller as distance increases). Therefore $1/\text{temp}$ is said to vary in a contravariant fashion.