Contents (preliminary)

Notations

Natural units: $c = \hbar = 1$ $\quad 1eV = 5.07\cdot 10^{6} m^{-1} = 1.52\cdot 10^{15} s^{-1}$

Sum convention: $\sum_{\nu=0}^{3} A^{\mu}{}_{\nu} u^{\nu} \equiv A^{\mu}{}_{\nu} u^{\nu}$ $\quad$ Greek indices $\alpha, \beta, \gamma \ldots = 0,1,2,3$

$\sum_{i=1}^{3} A^{\mu}{}_{i} u^{i} \equiv A^{\mu}{}_{i} u^{i}$ $\quad$ Latin indices $a, b, c \ldots = 1,2,3$

Signature: $ds^2 = -dt^2 + dx^2 + dy^2 + dz^2$

1. Special Relativity and flat spacetime

General Relativity (GR) is a classical theory of gravity. It identifies gravity as the curvature of spacetime. The spacetime means the space (set of points) spanned by temporal and spatial coordinates (t, x, y, z).

In the absence of gravity the spacetime is called flat and GR reduces to Special Relativity (SR). Let us briefly review SR and introduce basic concepts of differential geometry which we need later in discussing GR.

Even in the absence of gravity, the flat spacetime has a specific non-trivial structure. In Newtonian physics there is a unique concept of time which is the same for all inertial observers (= observers with no acceleration $\bar{F} = \frac{d^2\bar{x}}{dt^2} = 0$). This is not true in the nature: time and space transform into each other in the way described by the Special Relativity.

Newtonian laws of physics are invariant under Galilean transformations:

- Boosts: $(t', \bar{x}') = (t, \bar{x} - \bar{v}t)$

relates two frames K and K' moving with constant velocity $\bar{v}$ w.r.t. each other

$t = t'$ time and not affected!

- Shifts: $(t', \bar{x}') = (t + \underset{\uparrow}{\alpha}, \bar{x} + \underset{\nearrow}{\bar{\beta}})$

constants

- Rotations $(t', \bar{x}') = (t, R\bar{x})$

$R^T R = 1$, R is 3×3 rotation matrix

This invariance is realised in nature only in the limit of small velocities $v \ll c = 3{,}00 \cdot 10^8 \, m/s$.

In particular, Maxwell equations are not invariant under Galilean boosts:

$$\begin{cases} t' = t \\ x' = x - vt \end{cases} \qquad \text{(take a boost along } x\text{-axis for simplicity}, y'=y, z'=z)$$

but instead under the Lorentz transformation (boost along x-axis):

$$(1.1) \quad \begin{cases} t' = \gamma(t - vx) \\ x' = \gamma(x - vt) \end{cases} , \; \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \qquad t' = t'(t,x) \neq t \quad \text{time \& space mix}.$$

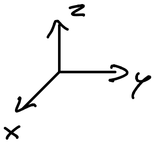
This lead Einstein to postulate Special Relativity which states that:

- Inertial frames related by Lorentz transformations, laws of physics look the same for all inertial observers ⇒ laws of physics must be Lorentz invariant
- Speed of light same in all inertial frames ⇒ must be a property of spacetime

1.1 Spacetime of Special Relativity

Note that for the transformation (1.1) we have $-t'^2 + x'^2 = -t^2 + x^2$. This can be taken as the defining property of SR.

Consider a collection of inertial frames $\begin{cases} K, & (t,x,y,z) \\ K', & (t',x',y',z') \end{cases}$ which may move with a constant velocity $\bar{v}$ wrt each other but for which $\frac{d^2x}{dt^2} = \frac{d^2y}{dt^2} = \frac{dz^2}{dt^2} = 0$.

Use for definiteness cartesian coordinates 

and define t by imagining a stationary clock at each point (x,y,z).

Define the spacetime interval Δs between any two events* A and B as:

$$(1.2) \quad \Delta s^2 \equiv -(\underset{\substack{\uparrow \\ \text{constant}}}{c}\Delta t)^2 + \Delta x^2 + \Delta y^2 + \Delta z^2 \qquad \begin{aligned} &\Delta t = t_A - t_B \\ &\Delta x = x_A - x_B \text{ etc.} \end{aligned}$$

* event = spacetime point

Special Relativity states that Δs is invariant under $K \to K'$

$$(1.3) \quad \Delta s^2 = \Delta s'^2 \iff -(c\Delta t)^2 + \Delta x^2 + \Delta y^2 + \Delta z^2 = -(c\Delta t')^2 + \Delta x'^2 + \Delta y'^2 + \Delta z'^2$$

This gives rise to Lorentz transformations between inertial frames as we will see below. The constant c is the same c that appears in Lorentz transformations, i.e. the speed of light.

The spacetime interval (1.2) can be written more compactly as:

$$\Delta s^2 = -(c\Delta t^2) + \Delta x^2 + \Delta y^2 + \Delta z^2 \equiv \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu \quad ,$$

where we introduced a 4x4 matrix called metric,

$$(1.4) \quad \eta_{\mu\nu} = \begin{pmatrix} -1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{pmatrix}$$

Minkowski metric

(Could also choose another convention $\eta_{\mu\nu} = \begin{pmatrix} +1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix}$, this is common in particle physics.)

and introduced the notation: $x^0 = ct,\ x^1 = x,\ x^2 = y,\ x^3 = z$

We further use the Einstein sum convention which just means that repeated upper and lower indices are summed over

$$(1.5) \quad \Delta s^2 = \sum_{\mu,\nu} \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu \equiv \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu$$

The spacetime interval Δs^2 actually defines the concept of distance in the 4d spacetime, i.e. set of points (t, x, y, z). By stating that Δs^2 is given by (1.5), or equivalently the metric by (1.4), we specify a certain geometry for the spacetime. The metric (1.4) is called Minkowski metric and it defines a Minkowski spacetime. The geometry of the Minkowski space is called flat and it is the spacetime in SR.

Structure of the Minkowski space

The invariance of $\Delta s^2 = \eta_{\mu\nu}\Delta x^\mu \Delta x^\nu$ between any spacetime points uniquely classifies all possible curves connecting different points into three categories:

(1.6)
$\Delta s^2 < 0$ timelike curves, paths travelled by massive particles
$\Delta s^2 = 0$ lightlike curves, paths travelled by light and massless particles
$\Delta s^2 > 0$ spacelike curves, paths along which no information can travel

Consider motion of light along the x-axis $(\Delta y = \Delta z = 0)$

$$\Delta s^2 = -(c\Delta t)^2 + \Delta x^2 = 0 \quad \Rightarrow \quad \frac{\Delta x}{\Delta t} = c$$

and since Δs^2 invariant under transformation to another inertial frame $K \to K'$:

$$\Delta s^2 = -(c\Delta t')^2 + \Delta x'^2 = 0 \quad \Rightarrow \quad \frac{\Delta x'}{\Delta t'} = c$$

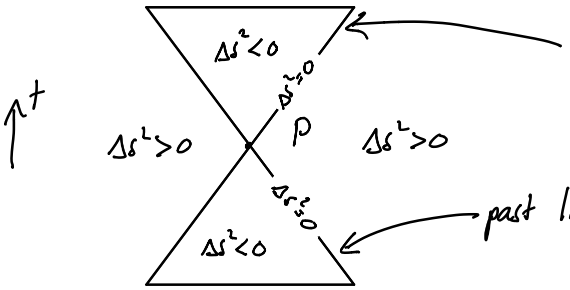
Hence, stating that light travels along the lightlike, or null, curves $\Delta s^2 = 0$ directly gives us the desired property that the speed of light is the same constant in any inertial frame.

From now on we choose units such that:

$$c = 3.00 \cdot 10^8 \, m/s \equiv 1 \quad \Leftrightarrow \quad 1s = 3.00 \cdot 10^8 m$$

This is indeed a very natural choice since temporal and spatial coordinates transform into each other and are therefore on equal grounds.

The classification of curves (1.6) divides the spacetime into different causal regions with respect to any spacetime point P. This is represented by the light cone:



future light cone, set of points which can interact with P in the future

past light cone, set of points which can have exchanged information with P

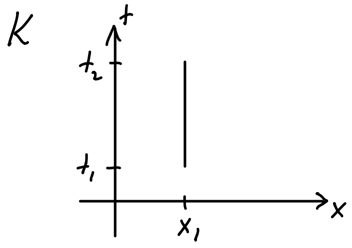
Points outside the lightcone are causaly disconnected from P.

The lightcone is invariant under $K \to K'$, transformations btw different inertial frames. This means that the causal structure is the same for all inertial observers

Proper time

The proper time τ_{AB} is the time measured by an observer moving between two spacetime events A and B.

For an inertial observer in her rest frame K this is just the coordinate time t:



$$\Delta\tau_{12} = t_2 - t_1 = \Delta t_{12} = \sqrt{-\eta_{\mu\nu}\Delta x^\mu \Delta x^\nu},$$

since $\Delta x^i = 0$, $i = 1,2,3$

The benefit of the last expression is that this is manifestly invariant under inertial transformations. In general, we define the proper time as

(1.7) $$(\Delta\tau)^2 \equiv -(\Delta s^2) = -\eta_{\mu\nu}\Delta x^\mu \Delta x^\nu$$ For timelike curves only!

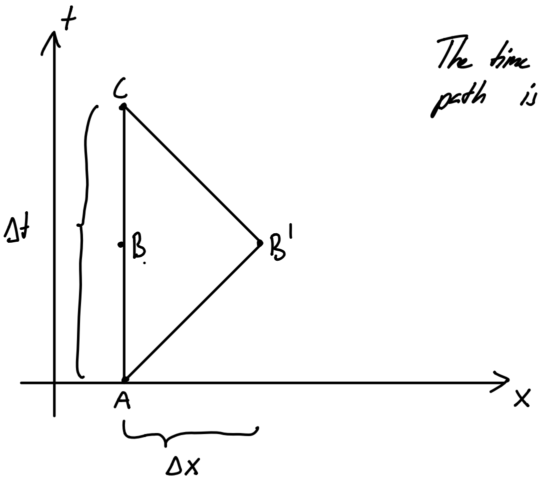
$$\Rightarrow \Delta\tau = \sqrt{-\eta_{\mu\nu}\Delta x^\mu \Delta x^\nu}$$

Using this expression we can directly express the proper time in terms of coordinates of any inertial frame:

$$\Delta\tau = \sqrt{-\eta_{\mu\nu}\Delta x^\mu \Delta x^\nu} = \underset{\substack{\uparrow \\ \text{rest frame}}}{\Delta t} = \sqrt{-\eta_{\mu'\nu'}\Delta x'^{\mu'} \Delta x'^{\nu'}} = \Delta t' \sqrt{1 - \frac{\Delta x'^2}{\Delta t'^2} - \frac{\Delta y'^2}{\Delta t'^2} - \frac{\Delta z'^2}{\Delta t'^2}}$$

coordinates of another inertial frame K' which moves wrt K.

Ex. Twin paradox, who ages more ABC or AB'C?



The time measured by an observer moving along any timelike path is given by the proper time along that path.

$$\Delta\tau_{AB'} = \sqrt{\left(\frac{1}{2}\Delta t\right)^2 - \Delta x^2} = \frac{1}{2}\Delta t\sqrt{1 - \Big(\underbrace{\frac{\Delta x}{\frac{1}{2}\Delta t}}_{=v}\Big)^2}$$

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

$$\Delta\tau_{AB'C} = \Delta\tau_{AB'} + \Delta\tau_{B'C} = \Delta t\sqrt{1-v^2}$$

$$\Delta\tau_{ABC} = \sqrt{\Delta t^2 - 0} = \Delta t$$

$\Rightarrow \Delta\tau_{ABC} > \Delta\tau_{AB'C}$, ABC ages more!

A straight path maximises proper time.

Going to infinitesimal limit, we can define the line element

(1.8) $$ds^2 \equiv \eta_{\mu\nu} dx^\mu dx^\nu$$

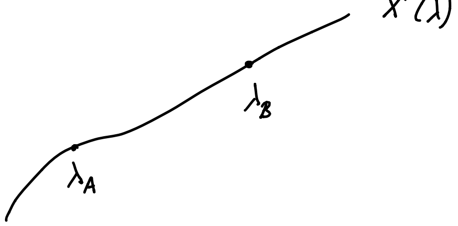
and further for $ds^2 < 0$ the proper time differential

(1.9) $$d\tau \equiv \sqrt{-\eta_{\mu\nu} dx^\mu dx^\nu}$$

Integrading this along a timelike curve $x^\mu(\lambda)$ (λ is some curve paramter)

we get:

(1.10) $$\tau_{AB} = \int_{\lambda_A}^{\lambda_B} d\lambda \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}}$$



Time measured by an observer moving along $x^\mu(\lambda)$ from A to B.

For spacelike curves $ds^2 > 0$ we can define the proper length (physical distance)

$$s_{AB} = \int_{\lambda_A}^{\lambda_B} d\lambda \sqrt{+\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}}$$

Lorentz transformations

The defining property of Minkowski spacetime is that the line element

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$$

is invariant under transformations $K \to K'$ from one inertial frame to another.

Let us now find how $K \to K'$ must act on the coordinates x^μ in order to meet this condition. An infinitesimal coordinate transformation can be represented by a linear matrix multiplication:

$$x^{\mu'} = \Lambda^{\mu'}{}_\nu x^\nu \quad , \quad \Lambda^{\mu'}{}_\nu \text{ is a } 4\times4 \text{ matrix.}$$

Under this transformation the line element changes as:

$$ds'^2 = \underset{\substack{\uparrow \\ \text{still the same Minkowski metric } \eta_{\mu'\nu'} = \begin{pmatrix} -1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{pmatrix}}}{\eta_{\mu'\nu'}} dx^{\mu'} dx^{\nu'} = \eta_{\mu'\nu'} \Lambda^{\mu'}{}_{\alpha} dx^{\alpha} \Lambda^{\nu'}{}_{\beta} dx^{\beta} = \Lambda^{\mu'}{}_{\alpha} \eta_{\mu'\nu'} \Lambda^{\nu'}{}_{\beta} dx^{\alpha} dx^{\beta}$$

Requiring that $ds^2 = ds'^2$ we get:

$$\Lambda^{\mu}{}_{\alpha} \eta_{\mu\nu} \Lambda^{\nu}{}_{\beta} dx^{\alpha} dx^{\beta} = \eta_{\alpha\beta} dx^{\alpha} dx^{\beta}$$

$$\Rightarrow \quad \Lambda^{\mu}{}_{\alpha} \eta_{\mu\nu} \Lambda^{\nu}{}_{\beta} = \eta_{\alpha\beta}$$

[note that we are free to name the dummy indices summed over as we wish:
$\eta_{\mu\nu} dx^{\mu} dx^{\nu} = \eta_{\alpha\beta} dx^{\alpha} dx^{\beta}$
$\Lambda^{\mu'}{}_{\alpha} \eta_{\mu'\nu'} \Lambda^{\nu'}{}_{\beta} = \Lambda^{\mu}{}_{\alpha} \eta_{\mu\nu} \Lambda^{\nu}{}_{\beta}$]

This is a condition for the form of Λ which generate the transformations btw. inertial frames.

Written in matrix form this reads

$$(1.11) \qquad \Lambda^T \eta \Lambda = \eta \qquad \Leftrightarrow \qquad \Lambda^{\mu}{}_{\alpha} \eta_{\mu\nu} \Lambda^{\nu}{}_{\beta} = \eta_{\alpha\beta}$$

The matrices Λ which satisfy (1.11) are generators of Lorentz transformations. They form a group called Lorentz group.

It is illustrative to compare (1.11) to rotations in 3d: $\bar{x}' = R\bar{x}$, where the rotation matrices R satisfy $R^T R = \mathbb{1} \Rightarrow \mathbb{1} = R^T \mathbb{1} R$. These matrices form a group $O(3)$, and imposing an extra condition $\det R = 1$ to exclude parity transformation $(\bar{x} \to -\bar{x})$, the group becomes $SO(3)$.

In the Lorentz group condition $\Lambda^T \eta \Lambda = \eta$, the unit matrix of $SO(3)$ $\mathbb{1} = R^T \mathbb{1} R$ is replaced by the Minkowski metric $\eta = \mathrm{diag}(-1, 1, 1, 1)$. The Lorentz group is denoted by $O(3,1)$ and it includes:

- 3d rotations (fixed t)
- 4d rotations called boosts (transformations btw inertial frames with different velocity)
- time reversals $t \to -t$
- parity transformations $\bar{x} \to -\bar{x}$

We are not interested in the latter two transformations and exclude them by imposing an additional condition: $\det \Lambda = 1$, $\Lambda^0{}_0 \geq 1$

The conditions:

(1.12) $$\Lambda^T \eta \Lambda = \eta \quad , \det \Lambda = 1 \quad , \Lambda^0{}_0 \geq 1$$

specify Lorentz transformations which form the restricted Lorentz group $SO(3,1)$. In the following we refer to this group when talking about Lorentz transformations.

In addition to the Lorentz transformations: $x^{\mu'} = \Lambda^{\mu'}{}_\nu x^\nu$, also constant shifts

$$x^\mu \to x^\mu + \overset{\swarrow \text{ constant, 4 degrees of freedom.}}{a^\mu}$$

leave the line element $ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$ invariant. Lorentz transformations + shifts together form the Poincare group which generates Poincare transformations

(1.13) $$x^{\mu'} = \underset{\substack{\uparrow \\ \text{3 rotations} \\ \text{3 boosts}}}{\Lambda^{\mu'}{}_\nu} x^\nu + \underset{\substack{\uparrow \\ \text{4 translations}}}{a^{\mu'}} \qquad 6+4=10 \text{ parameters}$$

This is the most general transformation $K \to K'$ which leaves ds^2 invariant.

The Jacobian of the Poincare transformation is the same as for Lorentz transformation since they only differ by constants:

$$\frac{\partial x^{\mu'}}{\partial x^\nu} = \Lambda^{\mu'}{}_\nu$$

Finally, we denote components of the inverse transformation Λ^{-1} by $\Lambda^\mu{}_{\nu'}$:

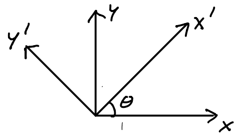
$K \to K'$: $x^{\mu'} = \Lambda^{\mu'}{}_\nu x^\nu$

$K' \to K$: $x^\mu = \Lambda^\mu{}_{\nu'} x^{\nu'} \Rightarrow x^\mu = \Lambda^\mu{}_{\nu'} \Lambda^{\nu'}{}_\sigma x^\sigma \Leftrightarrow \underbrace{\Lambda^\mu{}_{\nu'} \Lambda^{\nu'}{}_\sigma}_{= \frac{\partial x^\mu}{\partial x^{\nu'}} \frac{\partial x^{\nu'}}{\partial x^\sigma} = \frac{\partial x^\mu}{\partial x^\sigma} = \delta^\mu{}_\sigma} = \delta^\mu{}_\sigma = \begin{cases} 1, & \mu = \sigma \\ 0, & \mu \neq \sigma \end{cases}$

Examples:

Rotation by an angle θ in xy-plane:

$$\Lambda^{\mu'}{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

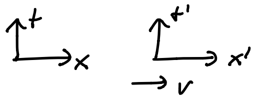


Boost in x-direction = rotation in tx-plane:

$$\Lambda^{\mu'}{}_{\nu} = \begin{pmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad x^{\mu'} = \Lambda^{\mu'}{}_{\nu} x^{\nu} \Rightarrow \begin{aligned} t' &= t\cosh\phi - x\sinh\phi \\ x' &= -t\sinh\phi + x\cosh\phi \\ y' &= y \\ z' &= z \end{aligned}$$

To see what this means, consider e.g. the point $x'=0$ in K-frame:

$x'=0 \Rightarrow x = t\frac{\sinh\phi}{\cosh\phi} = t\tanh\phi$, $x'=0$ moves with constant velocity $v = \tanh\phi$ wrt. K frame:



Using that $\tanh\phi = v$ we get:

$$\cosh^2\phi - \sinh^2\phi = 1$$
$$1 - \tanh^2\phi = \frac{1}{\cosh^2\phi} \Rightarrow \cosh\phi = \frac{1}{\sqrt{1-v^2}} \equiv \gamma$$

$$\sinh\phi = \sqrt{\cosh^2\phi - 1}$$
$$= \sqrt{\frac{1}{1-v^2} - \frac{1-v^2}{1-v^2}} = v\gamma$$

$$\Rightarrow \quad \begin{aligned} t' &= \gamma(t - vx) \\ x' &= \gamma(x - vt) \end{aligned}$$

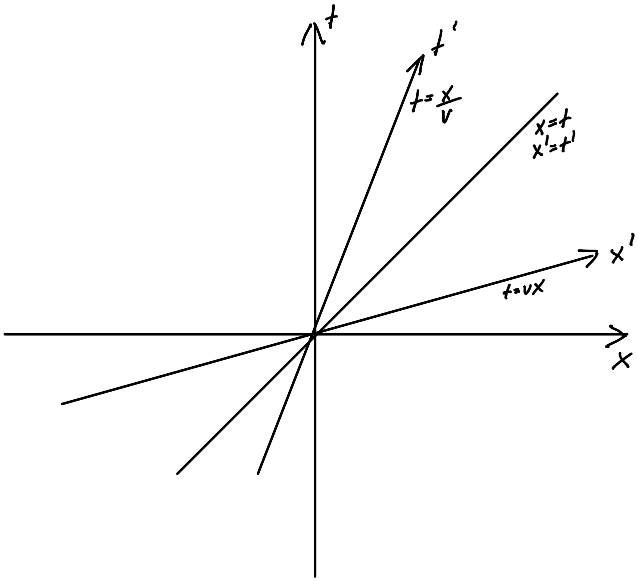
This yields the familiar time dilatation and lenght contraction results.

Boosts and spacetime diagrams

Draw the boosted frame K' in crd's of K:

t'-axis $\quad x'=0: \quad x = t\tanh\phi = vt$

x'-axis $\quad t'=0: \quad t = x\tanh\phi = xv$



Note that the light cone is invariant:

$$x = \pm t \qquad \begin{aligned} x' &= t(-\sinh\phi \pm \cosh\phi) \\ t' &= t(\cosh\phi \mp \sinh\phi) \\ &= \pm t(-\sinh\phi \pm \cosh\phi) = \pm x' \end{aligned}$$

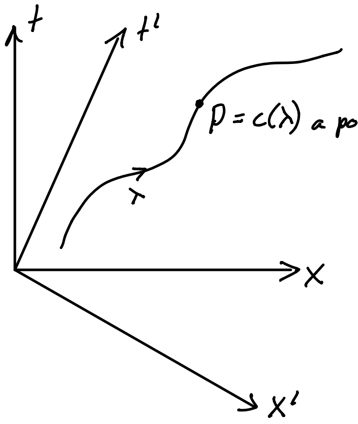
so that $x = \pm t$ maps to $x' = \pm t'$

1.2 Vectors in Minkowski spacetime

In the Euclidean space we are used to thinking vectors as arrows pointing to some direction. This concept needs to be reformulated in a more precise, geometric way to define the concept of vectors in more general spacetimes, i.e. manifolds. We start by discussing vectors in Minkowski space M, the definitions directly generalize to other manifolds (manifold is a spacetime which can be divided in patches which can be mapped onto $\mathbb{R}^n$ and connected in a smooth differentiable way, a more precise definition will follow later).

In Minkowski space M we define vectors as tangents of smooth curves $c : \mathbb{R} \to M$.

$$\underset{\substack{\uparrow \\ \text{curve parameter } \lambda \in \mathbb{R}}}{c(\lambda)} \in M$$



$P = c(\lambda)$ a point of M which lies on the curve c.

$X^\mu(P)$ crd's in frame K

$X^{\mu'}(P)$ crd's in frame K'

P is a physical point, $X^\mu(P) \neq X^{\mu'}(P)$ are its different crd representations.

Want to define vectors in a crd invariant manner!

Define the tangent vector of $c(\lambda)$ as the directional derivative along the curve

$$(1.13) \quad v = \frac{d}{d\lambda} = \frac{d X^\mu(c(\lambda))}{d\lambda} \frac{\partial}{\partial X^\mu} \qquad \text{The vector } v \text{ is an operator!}$$

$$\equiv \underset{\substack{\uparrow \\ \text{components}}}{v^\mu} \underset{\substack{\nwarrow \\ \text{basis vectors}}}{\frac{\partial}{\partial X^\mu}}$$

The vector v is an operator which acts on functions $f : M \to \mathbb{R}$

$$v[f] : M \to \mathbb{R}$$

(1.14) $V[f] \equiv \frac{dx^\mu}{d\lambda}\frac{\partial f}{\partial x^\mu} = \frac{df}{d\lambda}$ so $V[f]$ just gives the derivative of f along the curve $c(\lambda)$.

The vector v is a geometric object which is invariant under coordinate transformations.

$$V = \underbrace{\frac{dx^\mu}{d\lambda}}_{\equiv V^\mu}\underbrace{\frac{\partial}{\partial x^\mu}}_{\equiv e_\mu} = \underbrace{\frac{dx^{\mu'}}{d\lambda}}_{\equiv V^{\mu'}}\underbrace{\frac{\partial}{\partial x^{\mu'}}}_{\equiv e_{\mu'}} = \frac{d}{d\lambda}$$

components basis vectors

Notation: $\partial_\mu = \frac{\partial}{\partial x^\mu}$

Its components however change, just like in the Euclidean space.

Consider a Lorentz transformation: $x^{\mu'} = \Lambda^{\mu'}{}_\nu x^\nu$

$$V = V^\mu \frac{\partial}{\partial x^\mu} = V^\mu \frac{\partial x^{\nu'}}{\partial x^\mu}\frac{\partial}{\partial x^{\nu'}}$$

recall $\Lambda^{\nu'}{}_\mu = \frac{\partial x^{\nu'}}{\partial x^\mu}$

$$= V^\mu \Lambda^{\nu'}{}_\mu \frac{\partial}{\partial x^{\nu'}}$$

$$= V^{\nu'} \frac{\partial}{\partial x^{\nu'}} \quad \Rightarrow \quad V^{\mu'} = \Lambda^{\mu'}{}_\nu V^\nu$$

We thus find that the components V^μ transform as:

(1.15) $V^{\mu'} = \Lambda^{\mu'}{}_\nu V^\nu = \frac{\partial x^{\mu'}}{\partial x^\nu} V^\nu$ under $x^{\mu'} = \Lambda^{\mu'}{}_\nu x^\nu$

What about the basis vectors $e_\mu \equiv \frac{\partial}{\partial x^\mu}$?

(1.16) $e_{\mu'} = \frac{\partial}{\partial x^{\mu'}} = \frac{\partial x^\nu}{\partial x^{\mu'}}\frac{\partial}{\partial x^\nu} = \Lambda^\nu{}_{\mu'} e_\nu$

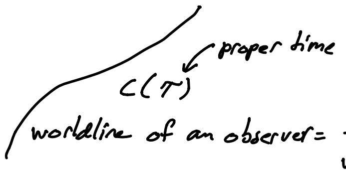
We can now check that v is indeed invariant:

$$V = V^{\mu'} e_{\mu'} = \underbrace{\Lambda^{\mu'}{}_\nu \Lambda^\sigma{}_{\mu'}}_{= \delta^\sigma{}_\nu} V^\nu e_\sigma = V^\mu e_\mu \quad \text{OK.}$$

Example:

4-velocity

$$u \equiv \frac{d}{d\tau} = \frac{dx^\mu}{d\tau}\frac{\partial}{\partial x^\mu}$$



worldline of an observer = the path along which she travels through the spacetime.

Apply u to the coordinate function $x^\mu(c(\tau))$ ← a point on the worldline

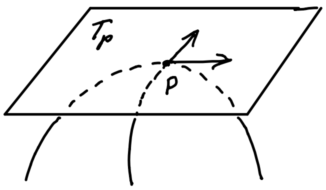
$$u[x^\mu] = \frac{dx^\nu}{d\tau}\frac{\partial x^\mu}{\partial x^\nu} = \frac{dx^\mu}{d\tau} = u^\mu$$

↑ μ - component in this crd system evaluated at x^μ.

So this rather abstract approach indeed reproduces the concept of vectors as we are used to thinking of them.

Tangent space:

The set of all vectors at some point $P \in M$ (i.e. tangents of all possible curves passing through P) defines the tangent space T_p



$u, v \in T_p$; $a, b \in \mathbb{R}$

$(a+b)(u+v) = au + bu + av + bv \in T_p$

The vector space $T_p(M)$ is the set of T_p's over the entire manifold M.

1.3 Dual vectors in Minkowski spacetime

In addition to vectors, we can define other geometrical quantities on a spacetime. Dual vectors, or one forms, are defined through their action on vectors. For every vector space $T_p(M)$ there exists a dual vector space $T_p^*(M)$ s.t.

$\omega \in T_p^*(M)$, $V \in T_p(M)$

$\omega : T_p(M) \to \mathbb{R}$ a map from the vector space to real numbers, ω is linear in its arguments

$$\omega = \underset{\substack{\uparrow \\ \text{components}}}{\omega_\mu} \underset{\substack{\nwarrow \\ \text{dual basis vectors}}}{dx^\mu} \quad , \quad V = \underset{\substack{\uparrow \\ \text{components}}}{V^\mu} \underset{\substack{\nwarrow \\ \text{basis vectors}}}{\partial_\mu} \quad , \quad \partial_\mu \equiv \frac{\partial}{\partial x^\mu}$$

$$\omega[V] \equiv \omega_\mu V^\nu \underbrace{dx^\mu[\partial_\nu]}_{= \delta^\mu{}_\nu}$$

this defines the dual basis vectors

$$(1.17) \quad \omega[V] = \omega_\mu V^\mu \equiv V[\omega]$$

Here we have chosen the so called coordinate basis where $\omega = \omega_\mu \theta^\mu$, $V = V^\mu e_\mu$ s.t. $e_\mu = \frac{\partial}{\partial x^\mu}$, $\theta^\mu = dx^\mu$. In general we are free to choose the basis vectors in a different way but here we mostly stick to this choice.

An important one-form is the gradient of a function $f: M \to \mathbb{R}$

$$df \equiv \underset{\substack{\nearrow \\ \text{components of} \\ \text{a gradient}}}{\frac{\partial f}{\partial x^\mu}} dx^\mu \qquad df[V] = V^\mu \frac{\partial f}{\partial x^\mu} = \frac{dx^\mu}{d\lambda} \frac{\partial f}{\partial x^\mu} = \frac{df}{d\lambda} = V[f]$$

Like vectors, the dual vectors are coordinate independent objects but their components change under crd transformations:

$x^{\mu'} = \Lambda^{\mu'}{}_{\nu}\, x^{\nu}$

$$\omega = \omega_{\mu'} dx^{\mu'} = \omega_\mu dx^\mu$$

$$\omega_{\mu'} dx^{\mu'} = \omega_\mu \underbrace{\frac{\partial x^\mu}{\partial x^{\nu'}}}_{= \Lambda^{\mu}{}_{\nu'}} dx^{\nu'}$$

(just write the
differential dx^μ
in terms of $dx^{\mu'}$)

$$\Rightarrow \quad \underline{\omega_{\mu'} = \Lambda^{\nu}{}_{\mu'}\, \omega_\nu} \qquad (1.18)$$

1.4 Tensors in Minkowski spacetime

So far we have defined : vectors $V \in T_p(M)$ $\quad V : T_p^*(M) \longrightarrow \mathbb{R}$ linear

$V = V^\mu \partial_\mu \qquad V[\omega] = V^\mu \omega_\mu$

$V^{\mu'} = \Lambda^{\mu'}{}_\nu V^\nu$

duals $\omega \in T_p^*(M)$ $\quad \omega : T_p(M) \to \mathbb{R}$ linear

$\omega = \omega_\mu dx^\mu \qquad \omega[V] = \omega_\mu V^\mu$

$\omega_{\mu'} = \Lambda^\nu{}_{\mu'} \omega_\nu$

functions, i.e. scalars $\phi : M \to \mathbb{R}$, $\phi' = \phi$ under $x^\mu \to x^{\mu'}$

These are all examples of tensors: scalar = tensor of type $(0,0)$, 0 indices

vector = tensor of type $(1,0)$, 1 upper index

dual = tensor of type $(0,1)$, 1 lower index

A general tensor of (m,n) type is defined as a multilinear map from m dual vectors and n vectors to real numbers

$$T : \underbrace{T_p^* \times \dots \times T_p^*}_{m \text{ times}} \times \underbrace{T_p \times \dots \times T_p}_{n \text{ times}} \longrightarrow \mathbb{R}$$

$T(\omega^{(1)}, \dots, \omega^{(m)}, v^{(1)}, \dots, v^{(n)}) \in \mathbb{R}$; (m,n) tensor is an operator which eats m duals and n vectors, and results a number

To construct a basis for general tensors, we define the tensor product $\otimes$

$$\underset{\substack{\uparrow \\ (m,n)}}{T} \otimes \underset{\substack{\uparrow \\ (p,q)}}{S}(\omega^{(1)}, \dots, \omega^{(m)}, \dots \omega^{(m+p)}, v^{(1)}, \dots, v^{(n)}, \dots v^{(n+q)}) \equiv T(\omega^{(1)}, \dots, \omega^{(m)}, v^{(1)}, \dots v^{(n)}) \underset{\substack{\uparrow \\ \text{usual product}}}{\times} S(\omega^{(m+1)}, \dots, \omega^{(n+p)}, v^{(n+1)}, \dots, v^{(n+q)})$$

In general the tensor product does not commute $T \otimes S \neq S \otimes T$

Using the coordinate basis for vectors and duals, the basis of (m,n) tensors is given by:

$$\partial_{\mu_1} \otimes \dots \otimes \partial_{\mu_m} \otimes dx^{\nu_1} \otimes \dots dx^{\nu_n}$$

and a (m,n) tensor can be written as:

$$(1.19) \qquad T = \underbrace{T^{\mu_1 \dots \mu_m}{}_{\nu_1 \dots \nu_n}}_{\text{Components}} \underbrace{\partial_{\mu_1} \otimes \dots \otimes \partial_{\mu_m} \otimes dx^{\nu_1} \otimes \dots dx^{\nu_n}}_{\text{basis}}$$

Using that $dx^\mu[\partial_\nu] = \partial_\nu[dx^\mu] = \delta^\mu{}_\nu$ we get:

$$(1.20) \qquad T(\omega^{(1)}, \dots, \omega^{(m)}, V^{(1)}, \dots, V^{(n)}) = T^{\mu_1 \dots \mu_m}{}_{\nu_1 \dots \nu_n} \omega^{(1)}_{\mu_1} \dots \omega^{(m)}_{\mu_m} V^{\nu_1}_{(1)} \dots V^{\nu_n}_{(n)}$$

$\therefore$ The action of a tensor on duals & vectors amounts to just multiplying the components and summing over indices.

The transformation of tensor components under $x^{\mu'} = \Lambda^{\mu'}{}_\nu x^\nu$ follows from the invariance of T just like for vectors and duals:

$$\begin{aligned}
&T^{\mu_1' \dots \mu_m'}{}_{\nu_1' \dots \nu_n'} \partial_{\mu_1'} \otimes \dots \otimes \partial_{\mu_m'} \otimes dx^{\nu_1'} \otimes \dots dx^{\nu_n'} \\
&= T^{\mu_1 \dots \mu_m}{}_{\nu_1 \dots \nu_n} \partial_{\mu_1} \otimes \dots \otimes \partial_{\mu_m} \otimes dx^{\nu_1} \otimes \dots dx^{\nu_n} \\
&= T^{\mu_1 \dots \mu_m}{}_{\nu_1 \dots \nu_n} \frac{\partial x^{\alpha_1'}}{\partial x^{\mu_1}} \dots \frac{\partial x^{\alpha_m'}}{\partial x^{\mu_m}} \frac{\partial x^{\nu_1}}{\partial x^{\beta_1'}} \dots \frac{\partial x^{\nu_n}}{\partial x^{\beta_n'}} \partial_{\alpha_1'} \otimes \dots \otimes \partial_{\alpha_m'} \otimes dx^{\beta_1'} \otimes \dots dx^{\beta_n'}
\end{aligned}$$

$$\Rightarrow \qquad T^{\mu_1' \dots \mu_m'}{}_{\nu_1' \dots \nu_n'} = \Lambda^{\mu_1'}{}_{\alpha_1} \dots \Lambda^{\mu_m'}{}_{\alpha_m} \Lambda^{\beta_1}{}_{\nu_1'} \dots \Lambda^{\beta_n}{}_{\nu_n'} T^{\alpha_1 \dots \alpha_m}{}_{\beta_1 \dots \beta_n} \qquad (1.21)$$

We can construct new tensors of T by acting with it on vectors & duals e.g. $T^\mu{}_\nu : V^\mu \to T^\mu{}_\nu V^\nu$ maps a vector V^μ to another vector $T^\mu{}_\nu V^\nu$

(1,1) tensor $\to T(\ ,V) = T^\mu{}_\nu\, \partial_\mu \otimes dx^\nu [V^\sigma \partial_\sigma] = T^\mu{}_\nu V^\sigma \underbrace{dx^\nu[\partial_\sigma]}_{=\delta^\nu{}_\sigma} \partial_\mu$

feed nothing here, i.e. act on T only on part of its arguments

$= \underbrace{T^\mu{}_\nu V^\nu}_{\text{components of the vector}} \partial_\mu$ (1,0) tensor = vector

Inner product and metric

The Minkowski metric $\eta_{\mu\nu}$ is a (0,2) tensor. It specifies the geometry of the Minkowski space (more on this later) and defines the inner product of vectors $V, U \in T_p(M)$

(1.22) $$V\cdot U \equiv \eta(V,U) = \eta_{\mu\nu} V^\mu U^\nu$$ crd invariant quantity

Using the inner product we can define the norm, or length, of a vector

(1.23) $$V\cdot V = \eta(V,V) = \eta_{\mu\nu} V^\mu V^\nu = -(V^0)^2 + (V^1)^2 + (V^2)^2 + (V^3)^2$$

Since (1.23) is a tensor product, the norm is crd invariant as it of course should be. The vectors are classified according to the sign of the norm:

$V\cdot V < 0$	timelike vector	$\Leftrightarrow$	tangent of a timelike curve
$V\cdot V = 0$	lightlike vector	$\Leftrightarrow$	tangent of a lightlike curve
$V\cdot V > 0$	spacelike vector	$\Leftrightarrow$	tangent of a spacelike curve

Lorentz transformations and the Minkowski metric

Recall that in Minkowski spacetime $\eta = \Lambda^T \eta \Lambda$ where $\eta = \begin{pmatrix} -1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{pmatrix}$

$\Leftrightarrow \eta_{\alpha\beta} = \Lambda^{\mu'}{}_\alpha \Lambda^{\nu'}{}_\beta\, \eta_{\mu'\nu'}$

This is just the transformation rule of (0,2) tensor

Since $\eta_{\alpha\beta} = \eta_{\alpha'\beta'} = \begin{pmatrix} -1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{pmatrix}$ we see that the components of the Minkowski metric are unchanged under Lorentz transformations.* This is related to the fact that global Lorentz transformations are a symmetry of the Minkowski spacetime.

Inverse metric

We define the inverse metric $\eta^{\mu\nu}$ by $\eta^{\mu\nu}\eta_{\nu\alpha} = \delta^{\mu}{}_{\alpha}$, i.e. the components of $\eta^{\mu\nu}$ are just the inverse of the matrix $\eta_{\mu\nu}$.

Since the components $\eta_{\mu\nu}$ do not change under Lorentz transformations, also the components of $\eta^{\mu\nu}$ remain unchanged.

Manipulating tensors

Raising and lowering indices

From a tensor T we can construct new tensors by multiplying with $\eta^{\mu\nu}$ or $\eta_{\mu\nu}$. This appears so frequently that it is convenient to introduce a specific notation:

$$\eta^{\mu\gamma} \underset{(2,2)\text{ tensor}}{T^{\alpha\beta}{}_{\gamma\delta}} \equiv \underset{(3,1)\text{ tensor}}{T^{\alpha\beta\mu}{}_{\delta}} \quad \text{raising of the index}$$

$$\eta_{\mu\alpha} T^{\alpha\beta}{}_{\gamma\delta} \equiv T_{\mu}{}^{\beta}{}_{\gamma\delta} \quad \text{lowering of the index}$$

So raising/lowering of indices $\Longleftrightarrow$ multiply with $\eta^{\mu\nu}/\eta_{\mu\nu}$

---

* in Cartesian coordinates.

Consider the inner product of vectors: $v, u \in T_p(M)$

$$\eta(v,u) = \eta_{\mu\nu} v^\mu u^\nu = v_\mu u^\mu$$

Here: $v_\mu = \eta_{\mu\nu} V^\nu$ is a dual vector constructed from v

Similarly: $\omega^\mu = \eta^{\mu\nu}\omega_\nu$ is a vector constructed from a dual $\omega \in T_p^*(M)$

Also note that with this notation:

$$\eta^{\mu\nu}\eta_{\nu\alpha} = \eta^\mu{}_\alpha = \delta^\mu{}_\alpha$$

because $\eta^{\mu\nu}$ is the inverse of $\eta_{\mu\nu}$ $\quad \eta^{\mu\nu}\eta_{\nu\alpha} \equiv \delta^\mu{}_\alpha$

Contraction

A contraction means summing over a pair of indices. It turns a (m,n) tensor into $(m-1, l-1)$ tensor:

$$\underset{(1,1)\text{ tensor}}{T^{\alpha\beta}{}_{\beta\delta}} = \underbrace{T^{\alpha\beta}{}_{\gamma\delta}}_{(2,2)\text{ tensor}}\,\eta^\gamma{}_\beta$$

sum over a pair of indices to get a lower rank tensor

Division into symmetric and antisymmetric parts

A tensor is called symmetric if $A_{\mu\nu} = A_{\nu\mu}$ and antisymmetric if $A_{\mu\nu} = -A_{\nu\mu}$. A generic tensor is neither symmetric nor antisymmetric but it can be unique divided into symmetric and antisymmetric parts.

$$A_{\mu\nu} = \underbrace{\frac{1}{2}(A_{\mu\nu} + A_{\nu\mu})}_{\equiv A_{(\mu\nu)}} + \underbrace{\frac{1}{2}(A_{\mu\nu} - A_{\nu\mu})}_{\equiv A_{[\mu\nu]}}$$

symmetric part

$A_{(\mu\nu)} = A_{(\nu\mu)}$

antisymmetric part

$A_{[\mu\nu]}$

Trace

The trace of a tensor is a scalar which is obtained by contracting over all the indices.

$$(1,1) \text{ tensor } A^\mu{}_\nu \quad , \ \overset{\text{trace}}{\swarrow} \, A \equiv A^\mu{}_\mu$$

$$(0,2) \text{ tensor } B_{\mu\nu} \quad , \ B \equiv \eta^{\mu\nu} B_{\mu\nu} = B^\nu{}_\nu$$

(note $\neq \sum_\mu B_{\mu\mu}$ so this is not the trace of the matrix $B_{\mu\nu}$)

The trace of the Minkowski metric is:

$$\eta^{\mu\nu}\eta_{\nu\mu} = \eta^\mu{}_\mu = \delta^\mu{}_\mu = 4$$

Note added, duals and tensors in Euclidean space

(Thanks to student comments!)

In the usual language of 3d Euclidean space, components of a vector would form a column matrix

$$V^i = \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} \qquad \text{vector}$$

A dual would correspond to a row matrix

$$\omega_i = (\omega_x \ \ \omega_y \ \ \omega_z) \qquad \text{dual}$$

Tensor product of a (1,1) tensor $T^i{}_j$ and v and ω would be

$$T^i{}_j v^j \omega_i = (\omega_x \ \omega_y \ \omega_z) \begin{pmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{pmatrix} \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix}$$

The metric is $g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$$ds^2 = g_{ij} dx^i dx^j$$
$$= dx^2 + dy^2 + dz^2$$

The inner product is $g_{ij} V^i V^j = V_i V^i = V^T V$ etc.

Tensor fields

A tensor space $T_p^* \times \ldots \times T_p^* \times T_g \times \ldots \times T_p$ is a product of dual and vector spaces. When we define tensor spaces over a set of spacetime points $p \in M$ (in practice over the entire spacetime), we obtain tensor fields.

E.g. $T_p(M)$ vector field

$u \in T_p(M)$, u = four velocity

$$u = \frac{dx^\mu}{d\tau}\frac{\partial}{\partial x^\mu}$$

↑ defined over the worldline of an observer.

Partial derivatives

In Minkowski space and in the Cartesian coordinates partial derivatives of tensors wrt coordinates x^μ form new tensors:

E.g. $\frac{\partial}{\partial x^\sigma} T^{\mu\nu} \equiv \partial_\sigma T^{\mu\nu}$ is a (2,1) tensor

Check this explicitly by investigating how $\partial_\sigma T^{\mu\nu}$ transforms under $x^{\mu'} = \Lambda^{\mu'}{}_\nu x^\nu$

$$\partial_{\sigma'} T^{\mu'\nu'} = \frac{\partial x^\alpha}{\partial x^{\sigma'}} \frac{\partial}{\partial x^\alpha} \left(\Lambda^{\mu'}{}_\beta \Lambda^{\nu'}{}_\gamma T^{\beta\gamma} \right)$$

↑↑ constant components for global Lorentz transformations

$$= \Lambda^\alpha{}_{\sigma'} \Lambda^{\mu'}{}_\beta \Lambda^{\nu'}{}_\gamma \, \partial_\alpha T^{\beta\gamma}$$

which is precisely the transformation rule of a (2,1) tensor

In a curved spacetime where the metric is not Minkowski this is no longer true and we need to define a covariant derivative to construct tensorial derivatives of tensors (i.e. crd independent derivative). This is so even in the Minkowski sace if we use other than Cartesian coordinates.

1.5 Relativistic mechanics

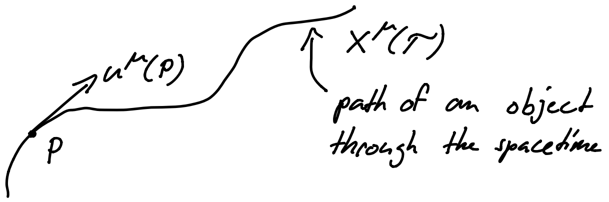
In Special Relativity (SR) the laws of physics are formulated in terms of tensorial quantities: 4-velocity, 4-acceleration, 4-force etc.

4-velocity

(1.24) $u^\mu \equiv \frac{dx^\mu}{d\tau}$ $\quad d\tau = \sqrt{-\eta_{\mu\nu}dx^\mu dx^\nu}$ proper time

↑

tangent of an objects worldline $x^\mu(\tau)$ where the curve parameter is the proper time τ.



The proper time differential $d\tau$ can be expressed in terms of inertial coordinates as:

$$d\tau = \sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}}\,dt$$

$$= \sqrt{-(-1)\left(\frac{dt}{dt}\right)^2 - \delta_{ij}v^iv^j}\,dt$$

$$= \sqrt{1-v^2}\,dt$$

↑

$v^2 \equiv v_i v^i$

(Recall $c \equiv 1$)

where we have defined:

$v^i \equiv \frac{dx^i}{dt}$ 3-velocity

and used that

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

Using this we can recast (1.24) into:

(1.25) $$u^\mu = \underbrace{\frac{1}{\sqrt{1-v^2}}}_{\equiv\gamma}\frac{dx^\mu}{dt} = \gamma(1,\vec{v})$$

Note that the γ factor is necessary for $u^\mu = \gamma \frac{dx^\mu}{dt}$ to transform as a vector. You can check by direct computation that $\frac{dx^\mu}{dt}$ alone does not transform as a vector under $x^{\mu'} = \Lambda^{\mu'}{}_{\nu} x^\nu$.

The norm of the 4-velocity is by definition -1:

$$u^\mu u_\mu = \eta_{\mu\nu} u^\mu u^\nu = \eta_{\mu\nu} \frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = -\frac{d\tau^2}{d\tau^2} = -1$$

$$\Rightarrow u^\mu u_\mu = -1 \qquad (1.26)$$

4-acceleration

(1.26) $\quad a^\mu \equiv \frac{du^\mu}{d\tau} = \frac{d^2 x^\mu}{d\tau^2}$ this is a vector since u^μ is a vector and τ is a scalar

$$a^\mu u_\mu = u_\mu \frac{du^\mu}{d\tau} = \frac{1}{2}\frac{d}{d\tau}\underbrace{(u_\mu u^\mu)}_{=-1} = 0 \Rightarrow a^\mu \text{ orthogonal to } u^\mu$$

4-momentum of massive particles

The mass m of an object is the same in all inertial frames, hence m is a scalar.

For an object with $m \neq 0$ we define the 4-momentum:

(1.27) $\quad p^\mu \equiv m u^\mu \quad$ this is a vector

The energy and 3-momentum are not invariant but depend on the frame. We define E and $\bar{p}$ as components of the 4-momentum

(1.28) $\quad E \equiv p^0 \;, \quad \bar{p} \equiv p^i \quad$ not tensorial but frame dependent

Therefore, the 4-momentum can be written as:

$$p^\mu = m u^\mu = m\gamma(1, \bar{v}) \Rightarrow \begin{aligned} E &= m\gamma \\ \bar{p} &= m\gamma\bar{v} \end{aligned} \qquad (1.29)$$

The norm of the four-momentum is by definition invariant

$$\underbrace{p^\mu p_\mu}_{=\text{scalar}} = m^2 \underbrace{u^\mu u_\mu}_{=-1} \Rightarrow p^\mu p_\mu = \eta_{\mu\nu} p^\mu p^\nu = -E^2 + \bar{p}^2 = -m^2$$

$$E^2 = \bar{p}^2 + m^2 \qquad (1.30)$$

The relativistic dispersion relation, i.e. the relation between the energy and momentum.

In the limit of small velocities $v \ll 1$, we recover the Newtonian results:

$$\gamma = \frac{1}{\sqrt{1-v^2}} \approx 1 + \frac{1}{2}v^2 + \ldots$$

$$E = m\gamma = \underset{\substack{\uparrow \\ \text{mass contributes to energy } (E=mc^2)}}{m} + \frac{1}{2}mv^2 + \ldots \quad \text{or equivalently} \quad E = \sqrt{|\bar{p}|^2 + m^2} = m + \frac{1}{2}mv^2 + \ldots$$

Relativistic version of $\bar{F} = m\bar{a}$

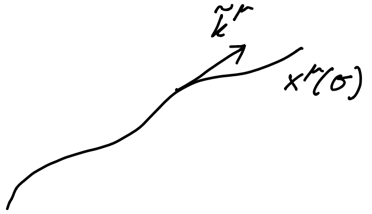
A force acting on particles is defined by the relativistic counterpart of Newton's law:

$$\underset{\substack{\uparrow \\ \text{4-force}}}{f^\mu} = \frac{dp^\mu}{d\tau} = m a^\mu = \frac{dt}{d\tau}\frac{dp^\mu}{dt} \underset{\substack{\uparrow \\ d\tau = \sqrt{1-v^2}\,dt = \frac{dt}{\gamma}}}{=} \gamma\left(\frac{dE}{dt}, \frac{d\bar{p}}{dt}\right) = \gamma(\bar{F}\cdot\underset{\substack{\uparrow \\ \text{3-velocity}}}{\bar{v}}, \underset{\substack{\uparrow \\ \text{3-force}}}{\bar{F}})$$

Massless particles $m=0$

Photons and other massless particles move along null curves $ds^2=0$. These cannot be parameterised by proper time since $d\tau = 0$ (photons experience no time!).

We can use some other parameter σ along the photon path:



and define the tangent vector $\quad ds^2 = -dt^2 + d\bar{x}^2 = 0 \Rightarrow \left|\frac{d\bar{x}}{dt}\right| = 1$

$$\tilde{k}^\mu = \frac{dx^\mu}{d\sigma} = \frac{dt}{d\sigma}\left(1, \frac{d\bar{x}}{dt}\right) = \frac{dt}{d\sigma}\left(1, \hat{x}\right)$$

unit 3-vector along the photon path

Here we are free to do reparameterisations $\sigma \rightarrow a\sigma + b$; a, b = constants and since $\tilde{k}^\mu \tilde{k}_\mu = 0$ the normalisation also does not fix σ.

Consider an inertial frame K where the source of photons is at rest. In this frame each photon has a definite energy:

$$E = \hbar\omega = \hbar\frac{2\pi}{\lambda} \quad , \; \lambda = \text{photon wavelength}$$

and momentum

$$\bar{p} = \hbar\omega\hat{k} \quad , \; \hat{k} = \text{unit 3-vector that points in the direction of photon propagation}$$

Now we can choose the curve parameter σ s.t. $\frac{dt}{d\sigma} = \omega$ and $\hbar\tilde{k}^\mu$ becomes the photon 4-momentum:

$$(1.31) \qquad p^\mu = \hbar k^\mu = \hbar\omega(1, \hat{k}) \qquad (\hat{k}\cdot\hat{k} = 1 \;, \text{unit 3-vector})$$

Here k^μ is called the wave vector. $\left(\omega = \frac{2\pi}{\lambda} = 2\pi f\right)$

The wave vector k^μ is tangent to photon path and hence a null vector:

$$k^\mu k_\mu = \omega^2(-1 + \hat{k}\cdot\hat{k}) = 0 \quad \Rightarrow p^\mu p_\mu = \hbar^2 k^\mu k_\mu = 0$$

The components of p^μ give the energy and 3-momentum .

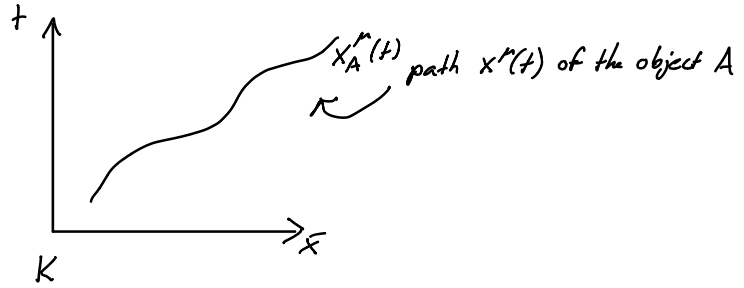
(1.32) $p^\mu = (E, \bar{p})$ (By definition of the 4-momentum)

Using that $p^\mu p_\mu = 0$ we get the dispersion relation of massless particles

(1.33) $E = |\bar{p}| = \hbar\omega$

What does an observer measure?

Consider an object A moving in the rest frame K of an observer



The energy and 3-momentum of A measured by K are given by the time and space components of the 4-momentum in the observers rest frame K.

$E_{obs} = p^0$ p^μ = four velocity of A in the observer rest frame K

$\bar{p}_{obs} = p^i$

The energy can be expressed in a manifestly coordinate invariant (=covariant) form in terms of the 4-velocity u^μ of K. In the rest frame we have just $u^\mu = \frac{dx^\mu}{d\tau} = (1,0,0,0)$, $u_\mu = (-1,0,0,0)$ so that:

(1.34) $E_{obs} = p^0 = -p^\mu u_\mu$ p^μ = 4-momentum of the object
u^μ = 4-velocity of the observer

The point is that the RHS is a scalar which has the same value in any crd system. To answer what energy an observer K measures for an object A we just write down the 4-momentum p^μ of A and the 4-velocity u^μ of K and compute E_{obs} according to (1.34).

This holds also for massless particles $m=0$:

$$(1.35) \qquad E_{obs} = \hbar\omega_{obs} = -\hbar k^\mu u_\mu \Rightarrow \underset{\substack{\nearrow \\ \text{observed wavelength}}}{\lambda_{obs}} = \frac{-2\pi}{k^\mu u_\mu}$$

k^μ = object wavevector

u^μ = observer 4-velocity