

In special relativity, many quantities may appear similar, e.g. with upper and lower indices, but they can play very different roles.

We can break them down into 3 categories:

i) Tensors: The kinematic and dynamical quantities we work with will (almost always) be represented by some (p, q) -tensor where p (q) is the number of upper (lower) indices. The distinction between upper and lower indices is exactly the distinction between a representation ρ of $SO(1,3)$ (upper) and a dual representation (lower).

Tensors can be combined with or without index contraction:

$$V_\mu T^{\mu\nu} = G^\nu$$

$$V_\mu T^{\lambda\nu} = H^\lambda{}^\nu$$

Some tensors can be represented by matrices, e.g. $V^\mu = (\cdot)$, $V_\mu = (\dots)$, $T^{\mu\nu} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$
but many cannot, e.g. $H_{\mu\nu}$. $T_{\mu\nu} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$

The indices on a tensor essentially tell us how it transforms: $T_{\mu\nu} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$

Each upper index transforms like a vector, i.e. $V^\mu \rightarrow V'^\mu = \Lambda^{\mu'}{}_{\mu} V^\mu$

Each lower index transforms like a dual vector, i.e. $V_\mu \rightarrow V_{\mu'} = \Lambda^{\mu}{}_{\mu'} V_\mu$

ii) The metric tensor: The metric tensor $g_{\mu\nu}$ takes an element of V (a vector index) to an element of V^* (a dual vector index). This is often called "lowering" the index, i.e. $g_{\mu\nu} V^\nu = V_\mu$.
The inverse metric $g^{\mu\nu}$ does the opposite, $g^{\mu\nu} V_\nu = V^\mu$.

Aside: Secretly $g_{\mu\nu}$ is also a dynamical field, but it takes general relativity to see that!

The metric is a true tensor and so transforms accordingly, i.e. $g_{\mu\nu} \rightarrow g'_{\mu'\nu'} = \Lambda^{\mu}{}_{\mu'} \Lambda^{\nu}{}_{\nu'} g_{\mu\nu}$
However we know: $\Lambda^T g \Lambda = g \Rightarrow g = \Lambda^{-T} g \Lambda^{-1} \Rightarrow g = g'$
Hence, the Λ 's are called "isometries" of the metric. They do not change its form.

In special relativity we often use $\eta_{\mu\nu}$ for $g_{\mu\nu}$.

The components of $\eta_{\mu\nu} = \begin{pmatrix} -1 & & \\ & 1 & \\ & & 1 \end{pmatrix}$ and happen also to be the components of $\eta^{\mu\nu} = \begin{pmatrix} -1 & & \\ & 1 & \\ & & 1 \end{pmatrix}$

Then: $\eta_{\mu\nu} \eta^{\nu\lambda} = \delta_{\mu}{}^{\lambda} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}$, $\eta_{\mu\nu} \eta^{\nu\mu} = 4$

Note: An arbitrary $(0,1)$ -tensor may seem like it can lower indices, e.g. $M_{\mu\nu} V^\nu$ but the result is not V_μ , it is a new tensor, i.e. $M_{\mu\nu} V^\nu = G_\mu$. Only the metric provides the unique map from a vector to its corresponding dual, $g_{\mu\nu} V^\nu = V_\mu$.

iii) Transformations: $\Lambda^{\mu'}{}_{\mu}$, $\Lambda^{\mu}{}_{\mu'} = (\Lambda^{\mu'}{}_{\mu})^{-1T}$ these always operate on tensors, can always be represented by a matrix, always carry one index from the old coordinates and one from the new

We never transform transformations! That would be silly!

Index notation and matrices

Okay, for real let's look at an example:

We can rotate a 2D vector $dx^{\mu} = \begin{pmatrix} dx \\ dy \end{pmatrix}$ with a matrix: $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} \cos\theta dx + \sin\theta dy \\ -\sin\theta dx + \cos\theta dy \end{pmatrix} = \begin{pmatrix} dx' \\ dy' \end{pmatrix}$
the order here matters! ← same thing!

Or we could say: $\Lambda^{\mu'}_{\mu} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \Rightarrow \Lambda^{1'}_1 = \cos\theta, \Lambda^{1'}_2 = \sin\theta, \Lambda^{2'}_1 = -\sin\theta, \Lambda^{2'}_2 = \cos\theta$
 then $dx^{\mu} \rightarrow dx^{\mu'} = \Lambda^{\mu'}_{\mu} dx^{\mu} = \Lambda^{\mu'}_1 dx + \Lambda^{\mu'}_2 dy$

or

$$\left. \begin{aligned} dx' &= \cos\theta dx + \sin\theta dy \\ dy' &= -\sin\theta dx + \cos\theta dy \end{aligned} \right\}$$

But we could also say: $dx^{\mu'} = dx^{\mu} \Lambda^{\mu'}_{\mu} = dx \Lambda^{\mu'}_1 + dy \Lambda^{\mu'}_2$

We end up with the same thing even though we switched order!

So one huge advantage to index notation is that we don't have to worry about order, But if we do want to use matrices (sometimes they are useful) we have to get the order right!

$$dx^{\mu'} = \Lambda^{\mu'}_{\mu} dx^{\mu} = dx^{\mu} \Lambda^{\mu'}_{\mu} = \Lambda dx \quad \text{but not } dx \Lambda!$$

The trick is to always get repeated indices directly next to each other.

So: $M_{\mu\nu} T^{\lambda\mu} = T^{\lambda\mu} M_{\mu\nu} = TM$ but not MT !

Sometimes we need to reorder indices to get this to work. Some important things to note:

$$\begin{aligned} \Lambda_{\mu\nu} = \Lambda_{\nu\mu} \quad , \quad \text{starting w/ } \Lambda \sim \Lambda^{\mu'}_{\mu} \Rightarrow & \Lambda_{\mu}^{\mu'} = (\Lambda^{\mu'}_{\mu})^T \\ & \Lambda^{\mu}_{\mu'} = (\Lambda^{\mu'}_{\mu})^{-1} \\ & \Lambda_{\mu'}^{\mu} = (\Lambda^{\mu'}_{\mu})^{-T} \end{aligned}$$

In nonrelativistic physics we often use the spatial derivative $\vec{\nabla}$ which operates on fields, e.g. $\vec{\nabla}\phi$, $\vec{\nabla}\cdot\vec{E}$, and time derivative $\frac{\partial}{\partial t}$ which operates on both fields, e.g. $\frac{\partial\vec{v}}{\partial t}$, and particle degrees of freedom, e.g. $\frac{d\vec{r}}{dt}$.

When we move to special relativity, one might naively think that we just combine these to make a 4-derivative. We do, but this is only part of the story.

First: $\partial_\mu = \frac{\partial}{\partial x^\mu} = (c\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$ is the 4-derivative that acts on fields.

Just as the index placement suggests, this is in fact a dual vector, i.e. $\partial_\mu \rightarrow \partial_{\mu'} = \Lambda^{\mu'}_{\mu} \partial_\mu$.
Hence the 4-gradient $\partial_\mu \phi$ is a dual vector } similar to 3D
and the 4-divergence $\partial_\mu A^\mu$ is a scalar

We will encounter this type of derivative a lot when we work in terms of fields, particularly when constructing Lagrangians and equations of motion.

But that is not the end of the story...

When we actually calculate probabilities for comparison with experiment we work in a limit where things look more like particles. Recall that for particles in 3D the main degree of freedom is position $\vec{r}(t)$, and interesting kinematic quantities are obtained by taking derivatives with respect to time, e.g. $\vec{v} = \frac{d\vec{r}}{dt}$, $\vec{p} = m\vec{v}$, $E = \frac{1}{2} m v^2$, $\vec{a} = \frac{d\vec{v}}{dt}$.

So a naive way to generalize this to 4D is: $d\vec{r} \rightarrow dx^\mu = \begin{pmatrix} cdt \\ dx \\ dy \\ dz \end{pmatrix}$, $\vec{v} \rightarrow \frac{dx^\mu}{dt}$, etc.

But we immediately encounter a problem.

An important question to ask is "Does our new object transform like a tensor?"

The answer is no!

$$\frac{dx^\mu}{dt} \rightarrow \frac{dx'^\mu}{dt'} = \Lambda^{\mu'}_{\mu} \frac{dx^\mu}{dt}$$

this guy dt' does not transform like any tensor (scalar, vector, dual, etc.).

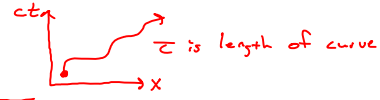
In fact we know how it transforms:

it transforms like one component of a vector, i.e. dx^μ !

Note: To transform like a tensor a quantity must only transform with "factors" of Λ .

To remedy this we need something that parameterizes the path of a particle that replaces time.

One obvious solution is the "length" of the path.



$$\text{Then for sub-luminal particles } \frac{d}{dt} \Rightarrow \frac{d}{d\tau} \text{ where } \tau = \int \sqrt{-ds^2} = \int \sqrt{c^2 dt^2 - dx^2 - dy^2 - dz^2}$$

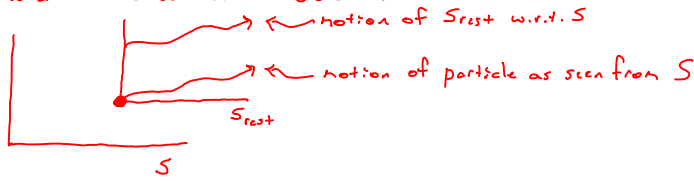
This is often called the "rest" time because in the rest frame of a particle $dx=dy=dz=0 \Rightarrow d\tau = c dt$

It solves our problem since $d\tau$ is an invariant!

So... introducing the 4-velocity $U^\mu \equiv c \frac{dx^\mu}{d\tau}$ which is a true tensor, $U^\mu \rightarrow U'^\mu = \Lambda^{\mu'}_{\mu} U^\mu$!

↑
This just gets the units right since $\frac{dx^\mu}{d\tau}$ is unitless!

Let's see what U^μ looks like in detail:



As seen in S :

$$U^0 = c \frac{dt}{dt} = c$$

$$U^1 = c \frac{dx}{dt} = c \frac{dx}{dt} \frac{dt}{dt}$$

$$U^2 = c \frac{dy}{dt} = c \frac{dy}{dt} \frac{dt}{dt}$$

$$U^3 = c \frac{dz}{dt} = c \frac{dz}{dt} \frac{dt}{dt}$$

But: $dt^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$

$$\downarrow$$

$$\frac{dt}{dt} = c \sqrt{1 - v^2/c^2}$$

$$\downarrow$$

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

$$\left. \begin{aligned} U^0 &= \gamma c \\ U^1 &= \gamma v_x \\ U^2 &= \gamma v_y \\ U^3 &= \gamma v_z \end{aligned} \right\} U^\mu = \begin{pmatrix} \gamma c \\ \gamma \vec{v} \end{pmatrix}$$

Note: The γ 's utilize all components of \vec{v} since $\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$

We can now define perhaps the most important quantity for studying collisions, 4-momentum $P^\mu = m U^\mu$

Then immediately: $P^\mu = \begin{pmatrix} m \gamma c \\ m \gamma \vec{v} \end{pmatrix} \Rightarrow P^\mu_{rest} = \begin{pmatrix} mc \\ 0 \end{pmatrix}$

includes decays
a.k.a. explosions

invariant rest mass

look like kinetic energy

To interpret this: For $v/c \ll 1$ $\gamma \approx 1 + \frac{1}{2} \frac{v^2}{c^2} \Rightarrow P^\mu = \begin{pmatrix} \frac{1}{2} (mc^2 + \frac{1}{2} m v^2 + \dots) \\ m \vec{v} + \dots \end{pmatrix} \equiv \begin{pmatrix} \frac{1}{2} E_{relativistic} \\ \vec{p}_{relativistic} \end{pmatrix}$

good old 3D momentum

henceforth we will just call these E and \vec{p}

This leads to the famous $E_{rest} = mc^2$, but more importantly ...

One interesting consequence of the negative sign in the metric of special relativity is as follows:

In 3D $ds^2 = dx^i dx^i = dx^2 + dy^2 + dz^2 \geq 0$ and only 0 when $dx^i = 0$.

In SR $ds^2 = dx^\mu dx^\mu = -c^2 dt^2 + dx^2 + dy^2 + dz^2$ can be +, -, 0 and more interestingly, can be 0 even when $dx^\mu \neq 0$!

Now ds^2 is an invariant so it is the same in any reference frame. That means if it is negative, it will be negative to all observers (same for positive and 0).

But notice that the sign is controlled by how the coordinate velocity compares to c .

If for example we take $dy=dz=0$ then for

$\left\{ \begin{array}{l} \frac{dx}{dt} < c \\ \frac{dx}{dt} = c \\ \frac{dx}{dt} > c \end{array} \right.$	$\underbrace{\hspace{2cm}}_{\text{consider } dx=0}$	$ds^2 < 0$	timelike
		$ds^2 = 0$	lightlike
		$ds^2 > 0$	spacelike

Since P^μ is a vector, $P_\mu P^\mu$ is an invariant.

In particular $\underbrace{P_\mu P^\mu}_{\text{Any frame}} = \underbrace{P_{\mu, \text{rest}} P^{\mu, \text{rest}}}_{\text{Rest frame}}$

$$\underbrace{-\frac{E^2}{c^2} + \mathbf{p}^2}_{\downarrow} = -m^2 c^2 \quad \text{or} \quad \boxed{E^2 - \mathbf{p}^2 c^2 = m^2 c^4}$$

This "mass-shell" condition relates relativistic energy and momentum to mass, and must be obeyed by all real particles!

Also recall: $P_\mu P^\mu \begin{cases} < 0 & \text{timelike} \Rightarrow m^2 > 0 & \text{massive} \\ = 0 & \text{lightlike} \Rightarrow m^2 = 0 & \text{massless} \\ > 0 & \text{spacelike} \Rightarrow m^2 < 0 & \text{tachyonic} \end{cases}$

To study collisions with P^μ we just define our system to include all colliding particles and then impose: $P_{\text{tot}, i}^\mu = P_{\text{tot}, f}^\mu$

However there is an incredibly useful trick at our disposal.

If we use $P_i^\mu = P_f^\mu$ then everything (both sides) must be evaluated in a single reference frame (x^μ).

However if we consider:

$$P_{i, i}^\mu = P_{i, f}^\mu = P_f^\mu \quad \text{both sides are invariants so we can evaluate them in any frame, even different ones!}$$

$$P_{i, i}^\mu = P_{i, f}^\mu = P_f^{\mu'}$$

Note: We would never consider $P_i^\mu = P_f^{\mu'}$!!

In special relativity, anything that was a vector in 3D must become a vector in 4D.

This includes coordinate related things like $d\vec{r} = \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}$, $\vec{v} = \begin{pmatrix} dx/dt \\ dy/dt \\ dz/dt \end{pmatrix}$ and also more abstract things

like $\vec{E} = \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}$, $\vec{B} = \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix}$.

To promote $\vec{E} + i\vec{B}$ to 4D we actually have two options:

1) Combine $\vec{E} + i\vec{B}$ into an antisymmetric (0,2) tensor $F_{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}$

Then $F_{\mu\nu} \rightarrow F'_{\mu\nu} = \Lambda^\alpha_\mu \Lambda^\beta_\nu F_{\alpha\beta}$

or

2) Use the scalar and vector potentials $\vec{E} = -\vec{\nabla}\phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}$ to form a vector $A^\mu = \begin{pmatrix} -\phi \\ \vec{A} \end{pmatrix}$

$\vec{B} = \vec{\nabla} \times \vec{A}$

Then $A^\mu \rightarrow A'^\mu = \Lambda^\mu_\nu A^\nu$

$A_\mu = (\phi, \vec{A}) = (\phi, A_x, A_y, A_z)$
 $= (A_0, A_1, A_2, A_3)$

These are equivalent if we say $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

Example: $F_{01} = \partial_0 A_1 - \partial_1 A_0$
 $= \frac{1}{c} \frac{\partial A_x}{\partial t} - \frac{\partial \phi}{\partial x}$
 $= -E_x$

$F_{23} = \partial_2 A_3 - \partial_3 A_2$
 $= \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}$
 $= (\vec{\nabla} \times \vec{A})_x$
 $= B_x$

$\partial_\mu = (\frac{1}{c} \frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$
 $= (\partial_0, \partial_1, \partial_2, \partial_3)$