Invariants and four-vectors in Special Relativity

We have seen some basic consequences of the two postulates of Special Relativity: length contraction, time dilation, etc. We were able to derive them from mathematical transformation relating two inertial frames — the Lorentz transformation. The mathematics gives us an insight how space and time are inextricably mixed. From the mathematical (and perhaps even from the physical one, too) viewpoint the formalism starts making more sense and is more convenient if we introduce a space-time continuum — a four dimensional space consisting of one temporal and three spatial dimensions. Moreover, one can merge in 4D not only time and space, but also other quantities such as the energy and three components of the momentum. The common thing about merging is that the resulting 4-vectors transform in a specific way under Lorentz transformation.

Let us introduce a few definitions

A scalar — a quantity (number) that is the same for all observers in inertial frames.

An event — a "point" in space-time. An event is defined by four coordinates:

$$(ct, x, y, z) \equiv (x^0, x^1, x^2, x^3)$$

We use ct rather than just t to make all four components having the same units (e.g. meters), while c is a natural constant that let us connects time and spatial distance.

The superscripts 0, 1, 2, 3 above indicate the order of the four components. Note that we use upper indicies. Lowes indicies will have another mission (more on that later). In some literature the arrangement of the time and space components is different, e.g. (x, y, z, ct) and the numbering may go from 1 to 4 (not from 0 to 3). This is all matter of convention. We will stick to the modern form as indicated in the definition above.

The set of 4-coordinates can be written simply as

x^μ $\mu = 0, 1, 2, 3$ (Greek letters are normally indicate an index running from 0 to 3)

The 4-dimensional space-time is called the Minkowski space-time.

If we consider two events

$$(ct_1, x_1, y_1, z_1) \qquad (ct_2, x_2, y_2, z_2)$$

the Galilean transformation leaves time intervals

$$\Delta t = t_2 - t_1$$

and space intervals (3D distances)

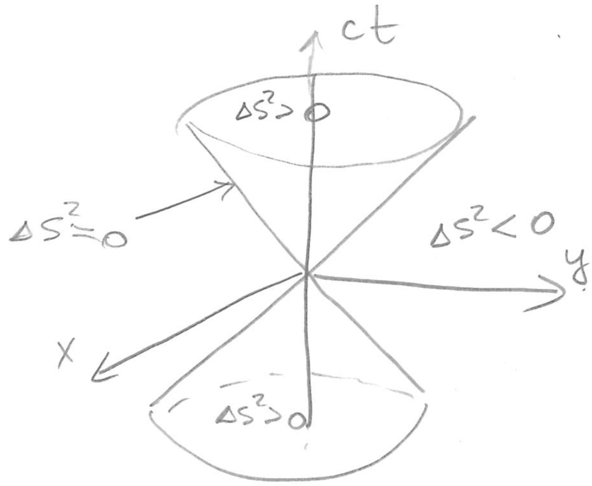
$$\Delta l^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$

invariant. Contrary to that, the Lorentz transformation leaves none of these invariant. The quantity that is invariant under the Lorentz transformation is the 4-dimensional space-time interval:

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

Note that Δs^2 is not positive definite. If $\Delta s^2 > 0$ we say the interval is time-like, while if $\Delta s^2 < 0$ we say the interval is space-like. Because one event cannot cause another event when the spatial distance between them exceeds $c\Delta t$ we can state that there could not be causal connection between events separated by a space-like interval.

Any set of physical quantities that transforms under the Lorentz transformations exactly like space and time is called a four-vector.



The light cone

$\left(\frac{E}{c}, p_x, p_y, p_z\right)$ form another important four-vector. The invariant relation

$$c^2 m^2 = \frac{E^2}{c^2} - \vec{p}^{\,2}$$

indicates that the energy could be considered as the 0-th component and the momentum as the other three components of a four-vector. Other examples of 4-vectors include

4-current $\quad j^\mu = (c\rho, \vec{j})$ $\quad$ 4-potential (electromagnetism) $\quad A^\mu = \left(\frac{\phi}{c}, \vec{A}\right)$

4-wavevector $\quad k^\mu = \left(\frac{\omega}{c}, \vec{k}\right)$ $\quad$ 4-velocity $\quad U^\mu = (\gamma c, \gamma\vec{u})$

4-gradient $\quad \partial^\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\vec{\nabla}\right)$ $\quad$ 4-force $\quad F^\mu = \left(\frac{1}{c}\frac{dE}{dt}, \frac{d\vec{p}}{dt}\right)$

The dot product of two four-vectors a^μ and b^μ is defined as

$$(a)\cdot(b) \equiv a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3$$

This dot product is invariant (scalar) as long as a^μ and b^μ are indeed 4-vectors.

The minus sign in the formation of a Lorentz invariant scalar means that there is asymmetry in the way time and space are treated. To avoid the appearace of this asymmerty we can can introduce a metric tensor.

$$g_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad \mu, \nu = 0, 1, 2, 3$$

The invariant scalar product of two 4-vectors can be written as

$$a \cdot b = a^\mu g_{\mu\nu} b^\nu$$

where the Einstein summation rule is assumed, i.e.

$$a^\mu g_{\mu\nu} b^\nu \equiv \sum_{\mu,\nu=0}^{3} a^\mu g_{\mu\nu} b^\nu$$

The metric tensor can also be used to define covariant components:

$$a_\mu \equiv g_{\mu\nu} a^\nu$$

The components a^μ (upper indicies) are called contravariant components. For an event we have

$$x^\mu = (ct, \vec{r})$$

$$x_\mu = (ct, -\vec{r})$$

The metric tensor $g^{\mu\nu} \equiv g_{\mu\nu}^{-1} = g_{\mu\nu}$ can be used to raise indicies, e.g.

$$a^\nu = g^{\mu\nu} a_\mu$$

With that we can write a dot product of two 4-vectors a and b as

$$a \cdot b = a_\mu b^\mu = a^\mu b_\mu$$

Relativistic Lagrangian mechanics

Let us begin with an assumption that there exist a relativistic scalar quantity called the action. If we minimize the action, the invariance of this scalar guarantees that all observers in inertial frames will agree that it reaches a minimum (extremum) for a given path. Now the choice of what to minimize is forced upon us by the requirement that the result be a Lorentz invariant, i.e. all observers must get the same value when they calculate the action.

The only two Lorentz invariants that exist for a particle not acted on by forces are its rest mass and proper time. Let us assume that the action is given by

$$S = \lambda \int_A^B d\tau$$

where λ is an unknown constant, which we will choose in such a way that the action has a correct nonrelativistic limit, and $d\tau$ is a differential of proper time. Since τ is a scalar under Lorentz transformations, all observers will calculate the same action for any given path

In a particular inertial frame the connection between proper time and local time for events at the same location in the rest frame is given by

$$dt = \gamma d\tau$$

Then for an observer in this particular frame the variational principle becomes

$$\delta S = 0 = \lambda \delta \int_{t_A}^{t_B} \frac{dt}{\gamma} = \lambda \delta \int_{t_A}^{t_B} \sqrt{1-\beta^2}\, dt$$

By analogy with the definition of a Lagrangian in nonrelativistic mechanics, we define the relativistic Lagrangian L by the equation

$$S = \int_{t_A}^{t_B} L\,dt$$

L and dt depend on the observer, while S does not.

If particle's velocity is small compared to the speed of light, we can expand

$$\sqrt{1-\beta^2} \approx 1 - \frac{1}{2}\beta^2 + \dots$$

If we ignore an insignificant constant term independent of β, for small β values we have

$$\delta S \equiv \delta\left(\int L\,dt\right) = \delta \int \left(-\frac{1}{2}\lambda\beta^2\right) dt$$

For this expression to agree with the nonrelativistic limit we must choose $\lambda = -mc^2$. This choice of λ is only a convention since L can always be changed by an additive or multiplicative constant without affecting the equations of motion.

The relativistic Lagrangian for a free particle is then

$$L = -mc^2\sqrt{1-\frac{v^2}{c^2}}$$

Proceeding in the usual way we can define canonical momenta

$$\vec{p} = (p_x, p_y, p_z) \equiv \left(\frac{\partial L}{\partial v_x}, \frac{\partial L}{\partial v_y}, \frac{\partial L}{\partial v_z}\right)$$

So

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}}$$

This is identical to our previous expression for relativistic momentum.

The Lagrange equations for a free particle are

$$\frac{d\vec{p}}{dt} = 0 \longleftarrow$$ the momentum $\vec{p}$ and hence $\vec{v}$ are constants for a free particle.