Sean Weill
Kapitza Society
Lecture 2

Special Relativity

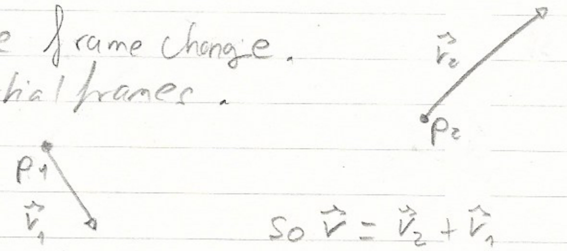
I. Lorentz Transformations:

According to the special theory of relativity, the laws of physics are equally valid in all inertial reference systems. An inertial system is one in which Newton's first law (the law of inertia) is obeyed: objects keep moving in straight lines at constant speeds unless acted upon by some force.

Thus, two inertial systems must be moving at constant velocity with respect to one another.

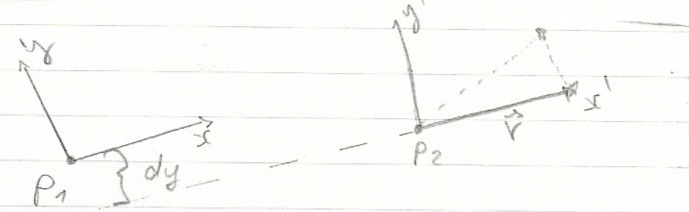
Imagine that we have two inertial frames, S and S', with S' moving at uniform velocity $\vec{v}$ with respect to S (and so S is moving at velocity $-\vec{v}$ with respect to S'). We can, of course, lay out our coordinates in such a way that the motion is along the common x/x' axis.

Before frame change. Inertial frames.



so $\vec{v} = \vec{v}_2 + \vec{v}_1$

After



Both y and y' are constant, they are related to each other by a constant dy so doesn't matter for relativity.

We can also set the master clocks at the origin in each system so that both read zero at the instant the two coincide (that is, $t = t' = 0$ when $x = x' = 0$).

Suppose, now, that some event occurs at position (x, y, z) and time t in S. What are the space-time coordinates (x', y', z') and t' of this same event in S'?

The fact that all inertial frames are physically equivalent can be expressed mathematically by requiring that the inner products be the same.

The dot product we are used to see is:

$$\vec{v}\cdot\vec{v} = \underbrace{\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}}_{\delta}\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}(v_1\ v_2\ v_3) = \begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}(v_1\ v_2\ v_3) = v_1^2+v_2^2+v_3^2$$

In special relativity, we use the Minkowski metric.

$$v\cdot v = \underbrace{\begin{pmatrix}-1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}}_{\eta}\begin{pmatrix}v_1\\v_2\\v_3\\v_4\end{pmatrix}(v_1\ v_2\ v_3\ v_4) = -v_1^2+v_2^2+v_3^2+v_4^2$$

MOLESKINE

So if we have $r = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$, $r' = \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix}$, we have $\left(\underline{\text{Note}}: \vec{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}\right)$

$$r \cdot r = r' \cdot r' \Leftrightarrow -t^2 + x^2 + y^2 + z^2 = -(t')^2 + (x')^2 + (y')^2 + (z')^2$$

Now, we are ready to answer the $S - S'$ question.

We saw that we can tweak our coordinate frames such that $y = y'$ & $z = z'$, and when $x' = 0$, $x = vt$.

So $\quad -t^2 + x^2 = -(t')^2 + (x')^2$

$\Leftrightarrow (it)^2 + x^2 = (it')^2 + (x')^2$. Letting $T = it$, and $T' = it'$, we have

$$T^2 + x^2 = (T')^2 + (x')^2$$

In a two-dimensional (x, T)-space, the quantity $x^2 + T^2$ represents the distance of a point P from the origin. This will also remain invariant under a rotation in (x, T)-space. If we denote the angle of rotation by θ, we have

$$(*) \quad \begin{pmatrix} x' \\ T' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ T \end{pmatrix}$$

Now, using the fact that when $x' = 0$, $x = vt$, we get

$$\boxed{x = \frac{vT}{i}}$$

So using $(*)$, we get $x' = x\cos\theta + T\sin\theta$

$\hookrightarrow 0 = \frac{vT}{i}\cos\theta + T\sin\theta \rightarrow \tan\theta = -\frac{v}{i} = iv$, θ is imaginary

$$\text{But } \cos\theta = \frac{1}{\sqrt{1 + \tan^2\theta}} = \frac{1}{\sqrt{1 - v^2}} = \gamma$$

$$x' = \cos\theta\,(x + T\tan\theta) = \gamma\,(x + (it)(iv)) = \boxed{\gamma(x - vt) = x'}$$

$$T' = \cos\theta\,(-x\tan\theta + T) = \gamma\,(it - xiv) = \gamma i (t - vx)$$

$\Leftrightarrow it = Bi(t - vx)$, giving $\boxed{t = \gamma(t - vx)}$

with good units $(c \neq 1)$, we have $x' = \gamma(x - vt)$, $t' = \gamma\left(t - \frac{vx}{c^2}\right)$.

1.1. Einstein's Notation:

We define the position-time four vector x^μ, $\mu = 0, 1, 2, 3$ as follows:

$x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$. Which changes the Lorentz transformations to:

$$\left.\begin{aligned}(x^0)' &= \gamma(x^0 - \beta x^1)\\ (x^1)' &= \gamma(x^1 - \beta x^0)\end{aligned}\right\} \beta \equiv v/c$$

$$(x^2)' = x^2$$

$$(x^3)' = x^3$$

We want to make this as compact as possible. We can do this by defining

$$\Lambda = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Which gives: $(x^\mu)' = \sum_{\nu=0}^{3} \Lambda^\mu_\nu x^\nu$, where $\Lambda^0_0 = \Lambda^1_1 = \gamma$

$$\Lambda^1_0 = \Lambda^0_1 = -\gamma\beta$$

To avoid writing lots of Σ's, we shall follow Einstein's "summation convention", which says that repeated Greek indices (one as subscript, one as superscript) are to be summed from 0 to 3. Giving

$$\boxed{(x^\mu)' = \Lambda^\mu_\nu x^\nu}$$

So for example, $(x^1)' = \Lambda^1_0 x^0 + \Lambda^1_1 x^1 + \Lambda^1_2 x^2 + \Lambda^1_3 x^3$

$$= -\gamma\beta x^0 + \gamma x^1 + 0 + 0$$

$$= \gamma(x^1 - \beta x^0) \checkmark$$

We can use a similar trick to express the invariant in a compact form.

$$I \equiv -(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 = -((x^0)')^2 + ((x^1)')^2 + ((x^2)')^2 + ((x^3)')^2$$

By using $\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, then $I = \sum_{\mu,\nu} \eta_{\mu\nu} x^\mu x^\nu = \boxed{\eta_{\mu\nu} x^\mu x^\nu = x \cdot x}$

Carrying things a bit further, we define the covariant four-vector x_μ as follows:

$$\boxed{x_\mu \equiv \eta_{\mu\nu} x^\nu}$$

So $x_0 \equiv \sum_{\nu=0}^{3} \eta_{0\nu} x^\nu = \eta_{00} x^0 + \eta_{01} x^1 + \eta_{02} x^2 + \eta_{03} x^3$

$$= -x^0$$

After computations, $\boxed{x_0 = -x^0,\ x_1 = x^1,\ x_2 = x^2,\ x_3 = x^3}$

To emphasize the distinction we call the "original" four vector x^μ a contravariant. Thus, the invariant I can then be written in its cleanest form

$$\boxed{I = x_\mu x^\mu}$$

Example: $a \cdot b = \eta_{\mu\nu} a^\mu b^\nu = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} a^\mu b^\nu = \begin{matrix} \eta_{00} a^0 b^0 + \eta_{01} a^0 b^1 + \dots + \eta_{03} a^0 b^3 \\ + \eta_{10} a^1 b^0 + \dots + \eta_{13} a^1 b^3 \\ + \vdots + \dots + \vdots \\ + \eta_{30} a^3 b^0 + \dots + \eta_{33} a^3 b^3 \end{matrix}$

$= -a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3$ since $\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

$\left.\begin{array}{l} \text{If } a^2 > 0,\ a^\mu \text{ is called spacelike} \\ \text{If } a^2 < 0,\ a^\mu \text{ is called timelike} \\ \text{If } a^2 = 0,\ a^\mu \text{ is called lightlike} \end{array}\right\}$ where $a^2 \equiv a \cdot a$

Finally, we can go from covariant $a_\mu = g_{\mu\nu} a^\nu$ to contravariant by reversing the signs again:

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

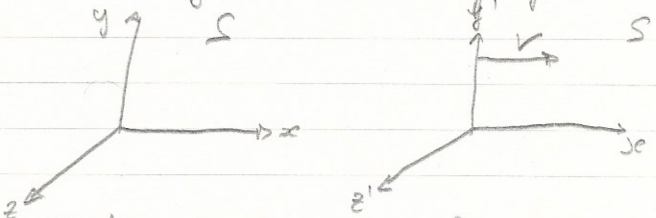
where $g^{\mu\nu}$ are technically the elements in the matrix g^{-1}.

MOLESKINE

IV. Relativistic Energy-Momentum Relationship:

According to special relativity, a moving clock is running slow relative to the "stationary" clocks on the ground. Specifically, while the "ground" time advances by an infinitesimal amount dt, the moving clock time (proper time) advances by the smaller amount $d\tau$:

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



Why? Suppose the clock at the origin in S' (moving) ticks off an interval $d\tau$; for simplicity, say it runs from $t'=0$ to $t'=d\tau$. In S, it begins when $t=0$, and ends when $t=\gamma(t'+\frac{v}{c^2}x') \rightarrow \gamma(d\tau+\frac{v}{c^2}0) = \gamma d\tau$ ✓

Although we can always get one from the other, in practice proper time is the best because τ is invariant! All observers can read the moving clock, and at any given moment they must all agree on what it says, even though their own clocks may differ from it and from one another.

In S, we have $\vec{v} = \frac{d\vec{x}}{dt}$. Something better is the proper velocity,

same $\vec{x}$, distance traveled measured in lab frame S.

$$\vec{p} = \frac{d\vec{x}}{d\tau} \quad \triangle$$

Giving, using the box above:

$$\vec{p} = \gamma\frac{d\vec{x}}{dt} = \gamma\vec{v}$$

$\vec{p}$ is much easier to work with, for if we want to go from S to S', both the numerator and denominator have to be changed for $\vec{v}$, while only the numerator for $\vec{p}$ since $d\tau$ is invariant.

In fact,
$$p^\mu = \frac{dx^\mu}{d\tau}, \quad p^0 = \frac{dx^0}{d\tau} = \frac{d(ct)}{\frac{dt}{\gamma}} = \gamma c.$$

Giving
$$\boxed{p_\mu = \gamma(c, v_x, v_y, v_z)}$$

Incidentally, $\eta_\mu\eta^\mu$ should be invariant: $p_\mu p^\mu = \gamma^2(-c^2+v_x^2+v_y^2+v_z^2) = \underbrace{\gamma^2}_{1/\gamma^2}c^2\left(\frac{v^2}{c^2}-1\right) = -c^2$ ✓

If we define momentum as $m\vec{\eta}$, then conservation of momentum is consistent with the principle of relativity (if it holds in one inertial frame, it holds in all).

$$\boxed{\vec{p} \equiv m\vec{\eta}}$$, proper velocity is a four vector, so

$$\boxed{p^\mu = m\eta^\mu}$$

The spatial components are $\boxed{\vec{p} = \gamma m\vec{v} = \frac{m\vec{v}}{\sqrt{1-v^2/c^2}}}$

The "temporal" component is $p^0 = \gamma mc = \frac{mc}{\sqrt{1-v^2/c^2}}$

For reasons that will appear in a moment, we define the relativist energy, E, as

$$E = \gamma mc^2 = \frac{mc^2}{\sqrt{1-v^2/c^2}}$$

Thus, $\boxed{p^0 = \frac{E}{c}}$, so $p_\mu = \left(\frac{E}{c}, p^1, p^2, p^3\right)$

Which gives : $p_\mu p^\mu = -\frac{E^2}{c^2} + |\vec{p}|^2 = m^2 \eta_\mu \eta^\mu = -m^2c^2$

$$\frac{E^2}{c^2} = m^2c^2 + |\vec{p}|^2 \Longleftrightarrow \boxed{E^2 = (mc^2)^2 + (|\vec{p}|c)^2}$$

In classical, there is no such thing as a massless particle; its momentum and its kinetic energy would both be zero, and it could sustain no force since $F = ma$, and hence (by Newton's third law) it could not exert a force on anything else. You might think it would be the same in relativity, but

$\vec{p} = \frac{m\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}}$, $E = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}$ reveals a loophole. If we have that massless particles always travels at the speed of light, we are back in business, since then we "have" $\frac{0}{0}$, so indeterminate.

But, we can use the box above. If $m = 0$ & $v = c$, we have

$$\boxed{E = |\vec{p}|c}$$

MOLESKINE

IV. Special Relativity and Lagrangians:

Consider the general idea of Hamilton's principle: nature chooses a path the extremizes the action. Physically, this means that nature always chooses the path that minimizes the energy that needs to be expended in traversing it.

let $c=1$ (convention).

We have just seen that $p_\mu = (E/c, p_x, p_y, p_z)$. The energy and momentum of a particle are functions of the particle's location and velocity through spacetime. So if we want to consider a path the particle travels in spacetime and then take a variation of that path to find the extremum in "energy-momentum space", we would set up our action as

$$\int_{\text{path}} p_\mu \, dx^\mu .$$

This integral represents the total energy-momentum along a particular path in spacetime

$$= \int \left(-p^0 dx^0 + p_1 dx^1 + p_2 dx^2 + p_3 dx^3\right)$$

we are considering a non-relativistic limit so we can do $E = \mathcal{H}$.

$$= \int \left(-E\,dt + \vec{p}\cdot d\vec{x}\right) = \int \left(\vec{p}\cdot d\vec{x} - \mathcal{H}\,dt\right)$$

$$= \int \left(\vec{p}\cdot\frac{d\vec{x}}{dt} - \mathcal{H}\right)dt$$

$$= \int \left(\vec{p}\cdot\dot{\vec{x}} - \mathcal{H}\right)dt$$

But $\mathcal{H} = \sum_i p_i \dot{x}_i - L$ so $L = \vec{p}\cdot\dot{\vec{x}} - \mathcal{H}$

Thus
$$\boxed{\int L\,dt = \int p_\mu \, dx^\mu}$$

We can think of the Lagrangian in the action as being a sort of non-relativistic limit of an action defined by integrating energy-momentum over spacetime.

V. Physically Allowable Transformations:

Consider some transformation matrix $R = R^{\mu}_{\nu}$ which acts on a vector a^{μ} as

$$a^{\mu} \rightarrow a^{\mu\prime} = R^{\mu}_{\nu} a^{\nu}$$

We are interested in the set of all transformations R that don't change the dot product, i.e.,

$$\begin{aligned} M_{\mu\nu} a^{\mu} b^{\nu} = a \cdot b = a' \cdot b' &= M_{\mu\nu} (a^{\mu})' (b^{\nu})' \\ &= M_{\mu\nu} (R^{\mu}_{\rho} a^{\rho})(R^{\nu}_{\sigma} b^{\sigma}) \\ &= (M_{\mu\nu} R^{\mu}_{\rho} R^{\nu}_{\sigma}) a^{\rho} b^{\sigma} \end{aligned}$$

This will hold if we impose

$$M_{\mu\nu} R^{\mu}_{\rho} R^{\nu}_{\sigma} = M_{\rho\sigma}$$

Which gives in the Minkowski metric:

$$\boxed{\eta_{\mu\nu} \Lambda^{\mu}_{\rho} \Lambda^{\nu}_{\sigma} = \eta_{\rho\sigma}}$$

This means that $\det(\eta_{\mu\nu} \Lambda^{\mu}_{\rho} \Lambda^{\nu}_{\sigma}) = \det(\eta_{\rho\sigma})$, but $\det(AB) = \det(A)\det(B)$,

$$\Leftrightarrow \quad \eta \Lambda^2 = \eta \Rightarrow \Lambda = \pm 1, \text{ where } \Lambda = \det(\Lambda^{\mu}_{\nu}) \text{ and } \eta = \det(\eta_{\mu\nu}).$$

$$\boxed{\det(\Lambda^{\mu}_{\nu}) = \pm 1}$$

What about the determinant -1 transformations? Like $(\Lambda_P)^{\mu}_{\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

Its effects on an arbitrary vector will be

$$\Lambda_P \cdot (t, \vec{x})^T = (t, -\vec{x})^T$$

So if (x, y, z) is right-handed, then $(-x, -y, -z)$ is left-handed. So the transformation changes the "handedness" of the system. ⚠
For this reason, Λ_P is called a Parity transformation.

MOLESKINE

Recall that $\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\gamma\beta & 0 \\ -\gamma\beta & \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}$, with $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ and $\beta = \frac{v}{c}$

so $\det(\Lambda^\mu{}_\nu) = 1 \begin{vmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{vmatrix} = \gamma^2 - \gamma^2\beta^2 = \gamma^2(1-\beta^2) = \gamma^2(1-\frac{v^2}{c^2})$

$= 1$

We call transformations that have $\det = 1$ the <u>Proper Lorentz transformations</u>.

Because they preserve $\eta_{\mu\nu}\Lambda^\mu{}_\rho\Lambda^\nu{}_\lambda = \eta_{\rho\lambda}$ (*), transformations with $\det = -1$ are still legitimate Lorentz transformations, but because you can't perform with them any combination of the proper transformations, such transformations are called <u>Improper</u>.

Only proper transformations are physical in the sense that they describe possible reference frames. However, as we will see, for a given physical theory to be relativistically invariant it must be invariant under all valid Lorentz transformations (need to satisfy (*)).

We can classify Lorentz transformations another way: since $\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

$$\eta_{00} = \Lambda^\mu{}_0\Lambda^\nu{}_0\eta_{\mu\nu} \iff -1 = -\Lambda^0{}_0\Lambda^0{}_0 + \sum_{i=1}^{3}\Lambda^i{}_0\Lambda^i{}_0$$

$$\iff (\Lambda^0{}_0)^2 = 1 + \underbrace{\sum_{i=1}^{3}(\Lambda^i{}_0)^2}_{\text{always} \geq 0}$$

So $(\Lambda^0{}_0)^2 \geq 1$, giving $\Lambda^0{}_0 \geq 1$, $\Lambda^0{}_0 \leq -1$.

We call any Lorentz transformation satisfying $\Lambda^0{}_0 \geq 1$ Orthochronous (chronous: time), and $\Lambda^0{}_0 \leq -1$ Non-Orthochronous.

It is not possible to perform a non-orthochronous transformation as any combination of orthochronous transformations and therefore these transformations are not physical. However, they satisfy (*) and therefore a relativistically invariant theory must account for them.

Example of a non-orthochronous transformation is $(\Lambda_T)^\mu{}_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, $(\Lambda_T)^0{}_0 = -1$, which takes $(t, \vec{x})^T \to (-t, \vec{x})^T$ and is called the Time Reversal transformation.

MOLESKINE