Relativity. Introduction

The reason for studying the relativity theory in this course is that the Maxwell theory of electromagnetism is historically first (and the simplest) relativistic theory. The relativity theory was tested in numerous experiments and was always confirmed. The increase of the life time of particles set in motion with the speed approaching the speed of light is one such example.

I. MAIN ASSUMPTIONS

The special relativity theory is based on two assumptions.

1. The laws of nature and the results of all experiments performed in a given frame of reference are independent of the translational motion of the system as a whole. There is an infinite set of such physically equivalent reference frames. These reference frames are referred to as inertial. This assumptions is also known as the principle of relativity. And as such appears in the Newtonian mechanics as formulated first by Galileo.
2. The speed of light is universal. In every inertial reference frame speed of light,

$$c = 3 \times 10^8 m/sec \tag{1}$$

is finite and independent of the motion of the light source. Any change in the physical condition propagates at a speed that cannot exceed c, (1). The universal constant speed allows the unique conversion of $[m]$ into $[sec]$ units and makes space and time interrelated into what we will call space-time.

We comment that the mathematical formulation of the principle of relativity is based on realization that it describes a certain *symmetry*. By symmetry we normally mean that things do not change under certain operation. In the latter case the operation is referred to as the symmetry operation. The relativity principle is the statement that the laws formulated as equations of motion for instance remain the same for all inertial observers. Certainly the motion itself look different for different observers. Nevertheless the equation that govern the motion are the same. This constancy of the laws is a symmetry which we will call Lorentz symmetry (for historical reasons). And the relativity principle will be often referred to as Lorentz symmetry.

A. Some consequences

(a) One obvious consequence is the relativity of simultaneity. Let the light source be equally distant from detectors A and B in K' see Fig.1. And let K' move relative to K with some speed $\boldsymbol{v}$. After the light source flashes it reaches the detectors A and B at the same time in K'. But in K the light reaches A earlier than B because the speed of light is the same in K and K'. It follows that the time cannot be absolute and the simple Galilean rule, $t = t'$ has to be abandoned.

(b) As a second consequence we note that the Galilean transformation is inconsistent with the principle of relativity. Specifically, the Galilean law of velocity transformation is in contradiction with the universality of the speed of light.

For the case of the axes of K and K' being parallel, and the relative velocity, $\boldsymbol{v} \parallel \hat{x} \parallel \hat{x}'$, see Fig.2 the Galilean transformation reads

$$x' = x - vt, y' = y, z = z' , \qquad t = t' \tag{2}$$

(2) assumes the K and K' coincide at $t = t' = 0$ which is totally immaterial for the argument. The velocity u' in K' is related to the velocity u in K as

$$u' = \frac{dx'}{dt'} = \frac{d}{dt}(x - vt) = u - v \tag{3}$$

For $u' = c$, $u = c + v$ in contradiction to the second postulate. As a result the Galilean transformation require revision as well as the Newtonian mechanics that is based on it.

(c) The Maxwell theory has to be Lorentz invariant in order to produce waves propagating with the universal velocity (1). This is a consistency check that will preoccupy us for few lectures.

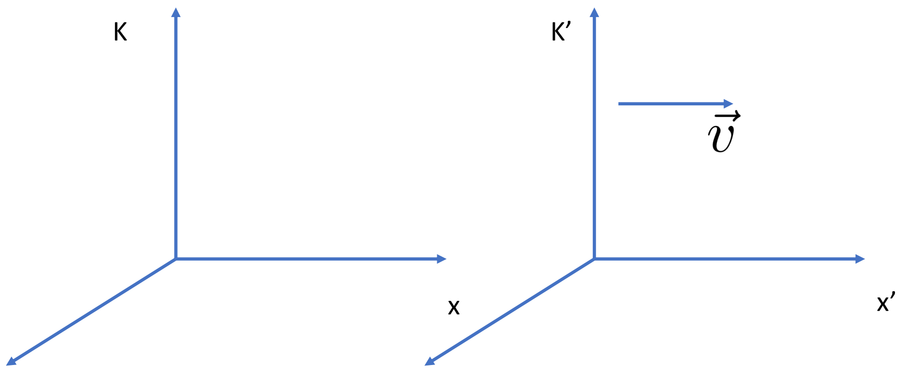

FIG. 1: K' moves relative to K

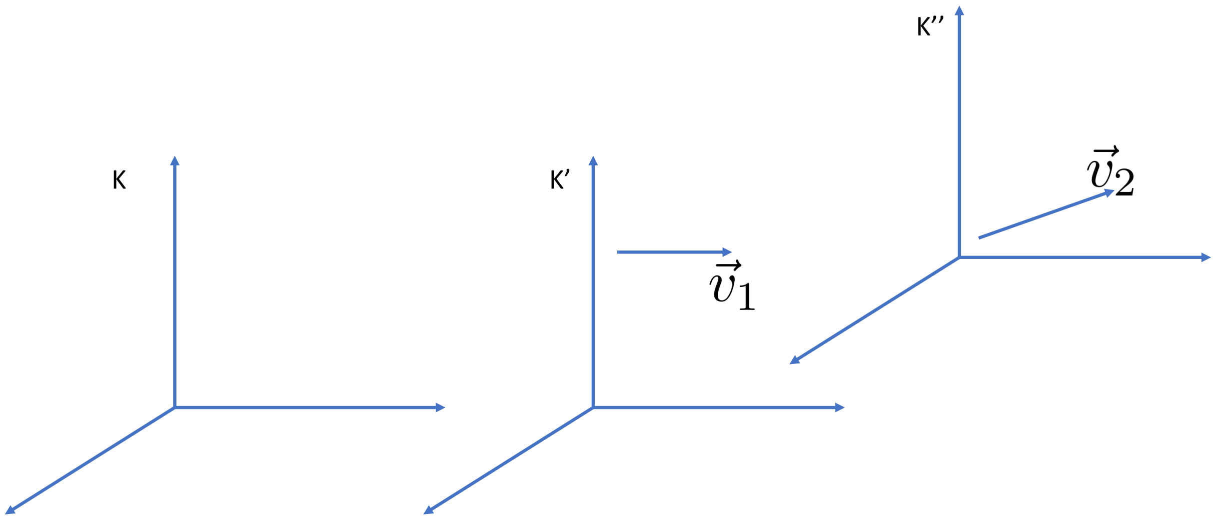

FIG. 2: K' and K" move relative to K

II. LORENTZ TRANSFORMATION

Lets's derive the trasnsformation of coordinates (e actually mean coordinates and time, or space-time coordinates) between the two inertial reference frames, K and K', where K' moves relative to K with velocity $\boldsymbol{v}$. It is conveniant to make a choice $\boldsymbol{v} \parallel \hat{x} \parallel \hat{x}'$, $\hat{y} \parallel \hat{y}'$ and $\hat{z} \parallel \hat{z}'$ as this is inessential and can easily be relaxed. Assume that K and K' coinside at $t = t' = 0$, and at this moment the flash of light occurs at the origin. One can think of a spherical wave outgoing from the origin $x = x' = y = y' = z = z' = 0$ at $t = t' = 0$. In our notations if the event is detected by K its space-time coordinates are (t, x, y, z). The **same** event in K' would occur at (t', x', y', z'). Our goal is to relate the two sets of the space-time coordinates.

If our event is the detection of light its coordinates in K, (t, x, y, z) satisfy

$$c^2t^2 - (x^2 + y^2 + z^2) = 0 \tag{4}$$

According to the universlity of spped of light we also have in K',

$$c^2t'^2 - (x'^2 + y'^2 + z'^2) = 0 \tag{5}$$

We are looking for the linear trasndormation since all the points in space-time must be equivalent, so we can write from the conditions (4) and (5) that for arbitrary (t, x, y, z)

$$c^2t'^2 - (x'^2 + y'^2 + z'^2) = \lambda[v](c^2t^2 - x^2 - y^2 - z^2) \tag{6}$$

because we have to satisfy (5) whenever (4) holds. In (6) $\lambda[v]$ is the numerical coefficient that depends on the absolute value of the relative velocity, $v = |\boldsymbol{v}|$. If the relative velocity $\boldsymbol{v}$ happens to not be collinear with the axes, one can always rotate the axes without changing the combinations, $x^2 + y^2 + z^2$ and $x'^2 + y'^2 + z'^2$.

Now consider the three reference frames K,K',K" that coincide at $t = t' = t'' = 0$. Such that K' (K") moves relative to K with velocities $\boldsymbol{v}_1$ ($\boldsymbol{v}_2$). Then we have

$$c^2t'^2 - (x'^2 + y'^2 + z'^2) = \lambda[v_1](c^2t^2 - x^2 - y^2 - z^2)$$
$$c^2t''^2 - (x''^2 + y''^2 + z''^2) = \lambda[v_2](c^2t^2 - x^2 - y^2 - z^2) \tag{7}$$

Also we have

$$c^2t''^2 - (x''^2 + y''^2 + z''^2) = \lambda[v_{12}](c^2t'^2 - (x'^2 + y'^2 + z'^2)) \tag{8}$$

where $v_{12} = |\boldsymbol{v}_{12}|$ and $\boldsymbol{v}_{12}$ is the relative velocity of the K' and K". Devide the two equations (7) and combine it with (8)

$$\frac{c^2t''^2 - (x''^2 + y''^2 + z''^2)}{c^2t'^2 - (x'^2 + y'^2 + z'^2)} = \frac{\lambda[v_2]}{\lambda[v_1]} = \lambda[v_{12}] \tag{9}$$

It is not important what v_{12} exactly equals to. What matters is that in the relation deduced from (9),

$$\frac{\lambda[v_2]}{\lambda[v_1]} = \lambda[v_{12}] \tag{10}$$

v_{12} depends on the relative orientation of $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$., while v_1 and v_2 does not. As a result we must conclude that

$$\lambda[v_1] = \lambda_0 = const \tag{11}$$

and from (10) the constant in (11) is

$$\lambda_0 = \frac{\lambda_0}{\lambda_0} = 1 \tag{12}$$

Hence form (6) we have simply,

$$c^2t'^2 - (x'^2 + y'^2 + z'^2) = (c^2t^2 - x^2 - y^2 - z^2) \tag{13}$$

The quantity in (13) is interval between the two events: one with coordinates $(0, 0, 0, 0)$ in both K and K' and the other with coordinates (t, x, y, z) in K and with coordinates (t', x', y', z') in K'. We see from (13) that the interval is frame independent. Such quantities are Lorentz invariant by definition and play a fundamental role.

For the choice of the axes as in Fig. we first note that $y = y'$ and $z = z'$. Indeed imagine the two identical rods placed along the y and y' axes in K and K', respectively. And consider the time instant when the axes coincide. If say $y' < y$ because K' is moving then as from the point of view of K' it is K that is moving we would also have $y < y'$ which is a contradiction. So we have

$$y = y', z = z' \tag{14}$$

Because of (14) we have from (13),

$$c^2t'^2 - x'^2 = c^2t^2 - x^2 \tag{15}$$

The one relation (15) leaves one parameter to parametrize the transformation. Wirte it as

$$\begin{bmatrix} ct \\ x \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} ct' \\ x' \end{bmatrix}, \tag{16}$$

and lets identify a_{ij} coefficients. Consider the event that in K' haas coordinates $(0, x', 0, 0)$. Then from (15) and (16) we have

$$c^2t^2 - x^2 = (x')^2(a_{12}^2 - a_{22}^2) = -(x')^2 \Rightarrow a_{22}^2 - a_{11}^2 = 1 \tag{17}$$

Similarly, considering the event with K' coordinates $(t', 0, 0, 0)$ we obtain from (15) and (16)

$$c^2t^2 - x^2 = (ct')^2(a_{11}^2 - a_{21}^2) = (ct')^2 \Rightarrow a_{11}^2 - a_{21}^2 = 1 \tag{18}$$

Consider finally the event with K' coordinate $(ct', x' = \pm ct', 0, 0)$ such that $(ct')^2 - (x')^2 = 0$. Then from (15) and (16) we get

$$c^2t^2 - x^2 = 0 = (ct')^2[(a_{11} \pm a_{12})^2 - (a_{21} \pm a_{22})^2] \Rightarrow 0 = (a_{11}^2 - a_{21}^2) - (a_{22}^2 - a_{12}^2) \pm 2(a_{11}a_{12} - a_{21}a_{22}) \tag{19}$$

Using (17) (18) in (19) gives

$$a_{11}a_{12} - a_{21}a_{22} = 0 \tag{20}$$

Constrains (17) and (18) can be resolved by writing,

$$a_{22} = \cosh\eta, a_{12} = \sinh\eta\,, \qquad a_{11} = \cosh\zeta, a_{21} = \sinh\zeta \tag{21}$$

Now (20) gives

$$0 = \cosh\zeta\sinh\eta - \sinh\zeta\cosh\eta = \sinh(\eta - \zeta) \Rightarrow \eta = \zeta \tag{22}$$

As a result we have for the transformation matrix in (16)

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \cosh\zeta & \sinh\zeta \\ \sinh\zeta & \cosh\zeta \end{bmatrix} \tag{23}$$

Check:

$$(ct)^2 - x^2 = (ct'\cosh\zeta + x'\sinh\zeta)^2 - (ct'\sinh\zeta + x'\cosh\zeta)^2 = (ct')^2 - (x')^2 \tag{24}$$

as expected.

Alternative and more general way to determine the possible matrices of Lorentz transformations would be as follows. Write the defining requirement (15) in the matrix form,

$$\begin{bmatrix} ct & x \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} ct \\ x \end{bmatrix} = \begin{bmatrix} ct' & x' \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} ct' \\ x' \end{bmatrix} \tag{25}$$

Then introducing the notation

$$\Lambda = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \Rightarrow \begin{bmatrix} ct \\ x \end{bmatrix} = \Lambda \begin{bmatrix} ct' \\ x' \end{bmatrix}\,, \quad \begin{bmatrix} ct & x \end{bmatrix} = \begin{bmatrix} ct' & x' \end{bmatrix} \Lambda^{tr} \tag{26}$$

As a result the requirement (15) takes the form,

$$\Lambda^{tr} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{27}$$

The matrices on both sides of (27) are symmetric, therefore the 4 matrix elements of Λ are constrained by 3 conditions. And the number of free parameters parametrizing the boost along $\hat{x}$ is $4 - 3 = 1$. These three conditions contained in one matrix equation, (27) are exactly Eqs. (17), (18) and (20) (Check it!)

A. Fixing the rapidity parameter, ζ and introduction of β and γ parameters.

To fix ζ note that the origin of K' has coordinates $(0, 0, 0, 0)$ in K' and must have coordinates $(ct, x = vt, 0, 0)$ in K. Then (16) gives

$$\begin{bmatrix} ct \\ x = vt \end{bmatrix} = \begin{bmatrix} \cosh\zeta & \sinh\zeta \\ \sinh\zeta & \cosh\zeta \end{bmatrix} \begin{bmatrix} ct' \\ 0 \end{bmatrix} = ct' \begin{bmatrix} \sinh\zeta \\ \cosh\zeta \end{bmatrix}$$

$$\Rightarrow \frac{vt}{ct} = \frac{ct'\sinh\zeta}{ct'\cosh\zeta} \Rightarrow \frac{v}{c} = \tanh\zeta \Rightarrow \sinh\zeta = \frac{v/c}{\sqrt{1-(v/c)^2}}, \cosh\zeta = \frac{1}{\sqrt{1-(v/c)^2}} \tag{28}$$

In terms of standard notations,

$$\boldsymbol{\beta} = \frac{\boldsymbol{v}}{c}\,, \gamma = (1-\beta^2)^{-1/2} \tag{29}$$

and unifying space and time coordinates,

$$x^0 = ct, x^1 = x, x^2 = y, x^3 = z \tag{30}$$

(Do not confuse with dimension label with the power exponent) we have $\sinh\zeta = \beta\gamma$ and $\cosh\zeta = \beta$, and we have

$$\begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^{'0} \\ x^{'1} \\ x^{'2} \\ x^{'3} \end{bmatrix} \tag{31}$$

In some cases the form,

$$\begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} \cosh\zeta & \sinh\zeta & 0 & 0 \\ \sinh\zeta & \cos\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^{'0} \\ x^{'1} \\ x^{'2} \\ x^{'3} \end{bmatrix} \tag{32}$$

is more convenient. The inverse transformation is obtained from (31) by changing $\boldsymbol{v} \to -\boldsymbol{v}$ or $\zeta \to -\zeta$. Indeed,

$$\begin{bmatrix} \cosh\zeta & \sinh\zeta \\ \sinh\zeta & \cosh\zeta \end{bmatrix}^{-1} = \begin{bmatrix} \cosh\zeta & -\sinh\zeta \\ -\sinh\zeta & \cosh\zeta \end{bmatrix} \tag{33}$$

Or alternatively,

$$\begin{bmatrix} x^{'0} \\ x^{'1} \\ x^{'2} \\ x^{'3} \end{bmatrix} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \tag{34}$$

And alternative form,

$$\begin{bmatrix} x^{'0} \\ x^{'1} \\ x^{'2} \\ x^{'3} \end{bmatrix} = \begin{bmatrix} \cosh\zeta & -\sinh\zeta & 0 & 0 \\ -\sinh\zeta & \cos\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \tag{35}$$

B. Example: Addition of velocities: the simplest configuration

Let the observer K' move relative to the observer K with the velocity $\boldsymbol{v}_1 = \hat{x}v_1$, and let the observer K" move with respect to K' with velocity $\boldsymbol{v}_2 = \hat{x}v_2$. What is the velocity, v_3 of the observer K" relative to K? Assume all the axes collinear.
Solution Denote $\tanh\zeta_{1,2} = v_{1,2}/c$. With the notation $c_{1,2} = \cosh\zeta_{1,2}$, $s_{1,2} = \sinh\zeta_{1,2}$ we get from (32)

$$\begin{bmatrix} x^0 \\ x^1 \end{bmatrix} = \begin{bmatrix} c_1 & s_1 \\ s_1 & c_1 \end{bmatrix} \begin{bmatrix} x^{'0} \\ x^{'1} \end{bmatrix} ; \begin{bmatrix} x^{'0} \\ x^{'1} \end{bmatrix} = \begin{bmatrix} c_2 & s_2 \\ s_2 & c_2 \end{bmatrix} \begin{bmatrix} x^{''0} \\ x^{''1} \end{bmatrix} \tag{36}$$

As a result we have using the known trigonometric identities for the cosh and sinh of the sum of the two arguments,

$$\begin{bmatrix} x^0 \\ x^1 \end{bmatrix} = \begin{bmatrix} c_1 & s_1 \\ s_1 & c_1 \end{bmatrix} \begin{bmatrix} c_2 & s_2 \\ s_2 & c_2 \end{bmatrix} \begin{bmatrix} x^{''0} \\ x^{''1} \end{bmatrix} = \begin{bmatrix} \cosh(\zeta_1+\zeta_2) & \sinh(\zeta_1+\zeta_2) \\ \sinh(\zeta_1+\zeta_2) & \cosh(\zeta_1+\zeta_2) \end{bmatrix} \begin{bmatrix} x^{''0} \\ x^{''1} \end{bmatrix} \tag{37}$$

Writing $\tanh\zeta_3 = v_3/c$ we see that $\zeta_3 = \zeta_1 + \zeta_2$ (even though the velocities are not additive, the "rapidities" are). We then have,

$$\frac{v_3}{c} = \tanh\zeta_3 = \tanh(\zeta_1+\zeta_2) = \frac{\tanh\zeta_1 + \tanh\zeta_2}{1+\tanh\zeta_1\tanh\zeta_2} = \frac{v_1/c + v_2/c}{1+v_1v_2/c^2} \Rightarrow v_3 = \frac{v_1+v_2}{1+v_1v_2/c^2} \tag{38}$$

which is the relativistic law of addition of velocities for collinear motion. It indeed consistent with the independence of the speed of light on the motion of the observer.

C. Example: Lorentz contraction

Consider again K' moving relative to K with the velocity $\boldsymbol{v} = v\hat{x}$ and the rod of the length ℓ_{rest} is at rest in K'. What would be the length of the rod, ℓ in K?
Solution Let the ends of the rod be located at $x_{2,1}$ at time t in K, such that $x_2 - x_1 = \ell$. Then in K'

$$x'_{2,1} = \gamma x_{2,1} - \gamma vt\,, \tag{39}$$

and we have

$$\ell_{rest} = x'_2 - x'_1 = \gamma(x_2 - x_1) = \gamma\ell \Rightarrow \ell = \gamma^{-1}\ell_{rst} \tag{40}$$

The Lorentz contraction has a simple geometrical interpretation using Fig.3 Case 2.

D. Example: Time delay

Consider two events that occur at times t'_2 and t'_1 at x' in K' , $t'_2 > t'_1$ (particle created and annihilated) What is the time interval in K if K' moves relative to K at velocity $\boldsymbol{v}$?
Solution In K,

$$ct_{2,1} = \gamma ct'_{1,2} + \gamma\beta x' \Rightarrow t_2 - t_1 = \gamma(t'_2 - t'_1) \tag{41}$$

which is the standard time delay in relativity which can be interpreted geometrically in Fig.3 Case 1.

E. Non-relativistic limit, $v \ll c$

In this limit $\gamma \to 1$, $\beta \to 0$, but $\beta x^0 \to vt$ is finite. Therefore from (31) we get $t = t'$ and $x = vt' + x' = vt + x$ which is the Galilean transformation law as expected.

III. INTERVAL AND CAUSALITY

We have seen that the simultaneity is not Lorentz invariant concept. Similarly, the future and past in general are not. Nevertheless the causality is Lorentz invariant concept. Furthermore, for the events related casually the concepts of past and future are absolute.

Consider two events, A and B with coordinates $(ct_A, \boldsymbol{x}_A)$ and $(ct_B, \boldsymbol{x}_B)$ in K. Define the Lorentz invariant interval between A and B

$$s^2_{AB} = c^2(t_A - t_B)^2 - |\boldsymbol{x}_A - \boldsymbol{x}_B|^2 \tag{42}$$

The interval is also independent on the choice of the origin. The shift in time and/or space of the axes of K will not affect it. In addition, usual spatial rotations of the coordinate axis. Indeed, the time is not affected under such rotation and it may be convenient to align the $\hat{x}$ axis with the vector $\boldsymbol{x}_A - \boldsymbol{x}_B$.

There are 3 types of intervals.

1. $s^2_{AB} > 0$ – time like separation between A and B. There exists K' where $\boldsymbol{x}_A = \boldsymbol{x}_B$. Furthermore, the two events in this case can be connected casually and the sign of $t_A - t_B$ is Lorentz invariant. (The absolute future and past for casually connected events)
2. $s^2_{AB} < 0$ – space-like separation between A and B. There exists K" where $t''_A = t''_B$.
3. $s^2_{AB} = 0$ – light-like separated events can be connected by the light signal only. In this case as in the case 1. the time interval has a sign independent on the observer.

Since the interval is Lorentz invariant the type of the interval is invariant as well.

Let us show the above statements.

In the case 1. Let us show the existence of the frame, where the two events happen at the same location. This is in fact clear as we can take the velocity of this frame to be just $\boldsymbol{v} = (\boldsymbol{x_B} - \boldsymbol{x}_A)/(t_B - t_A)$ if say $t_B > t_A$. But let's see it formally. Aligning the $\hat{x} \parallel \boldsymbol{x}_A - \boldsymbol{x}_B$ and considering the boost to K' with velocity $\boldsymbol{x} \parallel \boldsymbol{v}$ we use the condition

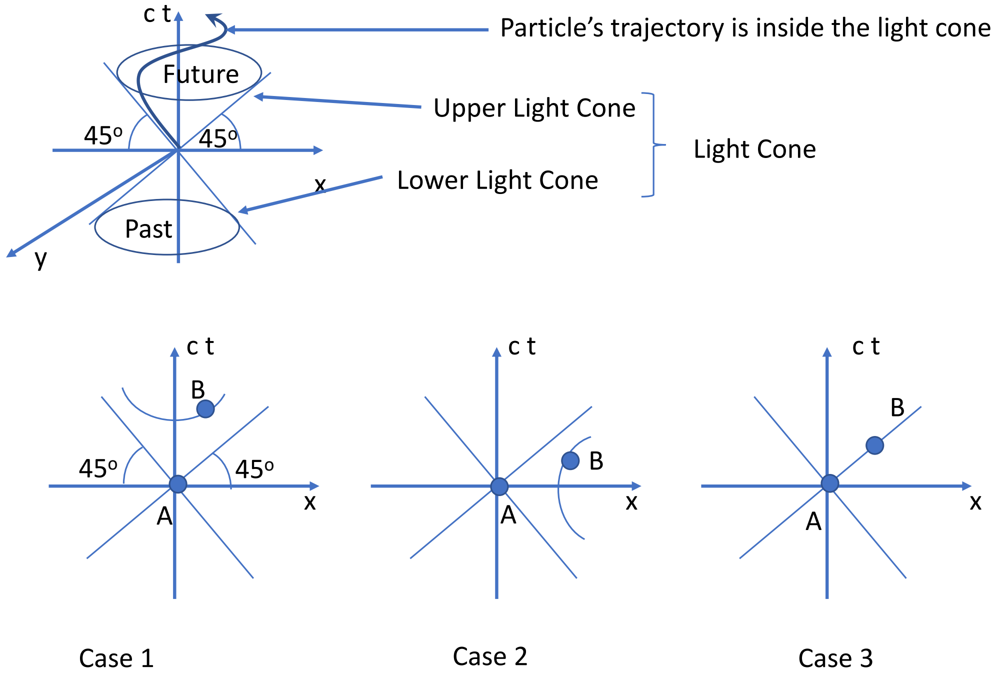

FIG. 3: Light Cone

$s_{AB}^2 = c^2(t_A - t_B)^2 - (x_A^1 - x_B^1)^2 > 0$ to write $x_A^1 - x_B^1 = L\sinh\eta$ and $c(t_A - t_B) = \pm L\cosh\eta$ for $(t_A - t_B) > 0$ and $(t_A - t_B) < 0$ respectively for some η. $L = \sqrt{s_{AB}^2}$ is a constant with dimensionality of the length. Parametrizing the transformation to K' by (35) (recall $\tanh\zeta = v/c = \beta$) we have

$$x_A^{'1} - x_B^{'1} = L(\mp\sinh\zeta\cosh\eta + \cosh\zeta\sinh\eta) = L\sinh(\eta\mp\zeta) \tag{43}$$

which vanishes for $\zeta = \pm\eta$. It is also seen that the sign of $x_A^{'1} - x_B^{'1}$ may be different in different coordinate frames.

Let us now assume for definiteness, $(t_A - t_B) > 0$. From (35) then

$$c(t'_A - t'_B) = x_A^{'0} - x_B^{'0} = L(\cosh\zeta\cosh\eta - \sinh\zeta\sinh\eta) = L\cosh(\zeta-\eta) > 0\,, \qquad \forall\zeta \tag{44}$$

In the other case, $(t_A - t_B) < 0$ we have

$$c(t'_A - t'_B) = x_A^{'0} - x_B^{'0} = L(-\cosh\zeta\cosh\eta - \sinh\zeta\sinh\eta) = -L\cosh(\zeta+\eta) < 0\,, \qquad \forall\zeta \tag{45}$$

which proves that the absolute character of the past and future for the events separated by the time-like interval at least for the case of the boost along $\hat{x} \parallel \boldsymbol{x}_A - \boldsymbol{x}_B$. In general situation $\boldsymbol{\beta} \nparallel \boldsymbol{x}_A - \boldsymbol{x}_B$ the transformation $x_A^{'0} - x_B^{'0} = \gamma[(x_A^0 - x_B^0) - \beta(x_A^1 - x_B^1)]$ obtained for $\boldsymbol{\beta} \parallel \boldsymbol{x}_A - \boldsymbol{x}_B$ is replaced with $x_A^{'0} - x_B^{'0} = \gamma[(x_A^0 - x_B^0) - \boldsymbol{\beta}(\boldsymbol{x}_A^1 - \boldsymbol{x}_B^1)]$. (Show it! or see (A7)) In result the previous arguments contained in (44) and (45) go through with $\zeta = \tanh\beta_\parallel$, $\beta_\parallel = \boldsymbol{\beta}\cdot\hat{x}$ instead of $\zeta = \tanh\beta$.

In the case 2. we similarly write, $x_A^0 - x_B^0 = L\sinh\eta$ and $x_A^1 - x_B^1 = \pm L\cosh\eta$ depending on the sign of $x_A^1 - x_B^1$. From (35),

$$x_A^{'0} - x_B^{'0} = L(\sinh\eta\cosh\zeta \mp \cosh\eta\sinh\zeta) = L\sinh(\eta\mp\zeta) \tag{46}$$

Clearly depending on ζ, $x_A^{'0} - x_B^{'0}$ can have arbitrary sign or vanish.

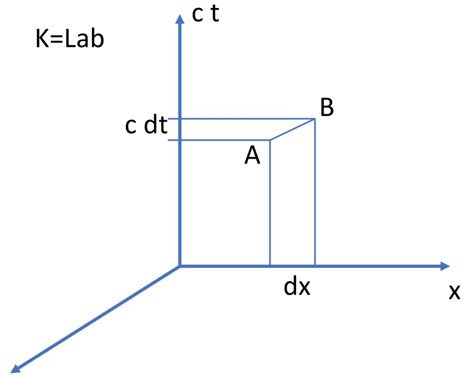

FIG. 4: The two events, A and B are infinitesimally close. A (B) is the particle is at $x_{A(B)}$ at time $t_{A(B)}$

In the case 3. we cannot apply the parametrization as above. Instead we proceed directly, as in this case $x^1_A - x^1_B = \pm(x^0_A - x^0_B)$. We then have straight from (35),

$$x'^0_A - x'^0_B = (x^0_A - x^0_B)(\cosh\zeta \mp \sinh\zeta) \tag{47}$$

$$\cosh\zeta \mp \sinh\zeta > 0 \Rightarrow \mathrm{sgn}(x'^0_A - x'^0_B) = \mathrm{sgn}(x^0_A - x^0_B)\,, \qquad \forall\zeta \tag{48}$$

A. Geometrical meaning of the causal relationship between the events

The absolute character of the future and the past has a transparent geometrical meaning. The events that are in future relative to a given (reference) event placed at origin are found within the upper part of the light cone. The Lorentz transformation is a continues transformation leaving the interval the same. All the points with the same interval relative to a reference event occupy a either one of the two hyperboloids. Once the event is found in the upper hyperboloid cannot be transformed into the lower part by continouse transformation because the two hyperboloids are disconnected. As a result the future (past) is an absolute concept.

In contrast or the space-like intervals the two hyperboloids cross the $t = 0$ plane. Therefore, for such events the future (past) depends on the observer. This “relativity” of the future(past) does not lead to a contradiction with the causality principle since the two events separated by the space-like interval can not be related by any physical signal because such a signal would have to propagate with speed exceeding c in contradiction to the postulates of relativity.

On the light cone the future and past are also absolute which is expected as the tao events might be connected by the light.

IV. PROPER TIME

Let A and B be the two events connected by the particle's trajectory. And let the velocity of the particle be $\boldsymbol{u}$ in the lab frame. Then the interval between the two infinitesimally close events A and B is

$$ds^2 = c^2dt^2 - (d\boldsymbol{x})^2 = c^2dt^2(1-\beta^2), \qquad \beta = u/c \tag{49}$$

In the frame attched to the particle, the particle does not move so the time elapsed between A and B in the frame of the particle $d\tau$ satisfies,

$$c^2d\tau^2 = ds^2 = c^2dt^2(1-\beta^2) \tag{50}$$

The time τ is the proper time by a particle. By (49) it is Lorentz invariant concept just like interval. From (50)

$$d\tau = dt\sqrt{1-\beta^2(t)} = \frac{dt}{\gamma(t)} \tag{51}$$

Integrating (51) we have for the time in the lab when the particle changes its velocity,

$$t_2 - t_1 = \int_{\tau_1}^{\tau_2} \frac{d\tau}{\sqrt{1-\beta^2(\tau)}} = \int_{\tau_1}^{\tau_2} d\tau' \gamma(\tau') \tag{52}$$

As $\gamma(\tau) \geq 1$ the time in the lab is always longer unless the particle is at rest. For relativistic particles $v/c \lesssim 1$, $\gamma \gg 1$ and $t_2 - t_1 \gg \tau_2 - \tau_1$. This phenomenon is routinely observed in particle accelerators. For instance the life time of the accelerated particle is (much) longer then the life tie of the same particle at rest.

V. 4-VECTORS: PRELIMINARY DISCUSSION

In modern point of view the event is itself 4-vector and is understood as geometrical object. Namely the event exists regardless of the reference frame. Its component however (x^0, x^1, x^2, x^3) do depend on the frame. In particular for the Lorentz transformation to another frame the two sets of components are related via, (31). In this sense the Lorentz transformation is a passive one: the vectors (events) stay the same, but their components change.

We now define the 4-vector A as an object with components (A^0, A^1, A^2, A^3) that transforms exactly as components of the event vector, (31) or (A3). By writing $(A^1, A^2, A^3) = \boldsymbol{A}$, we have by definiton

$$\begin{aligned} A^{'0} &= \gamma(A^0 - \beta A_\parallel) \\ A'_\parallel &= \gamma(A_\parallel - \beta A^0) \\ \boldsymbol{A}'_\perp &= \boldsymbol{A}_\perp \end{aligned} \tag{53}$$

The object A defined by its component is by definition a geometrical object in the same way as event vector is geometrical. The same way as event vector is the same for all observers also the 4-vector A is the same for all observers, only its component change according to (53).

As the Lorentz transformation (53) is linear if A and B are 4-vectors the same is true for any linear combination $c_1 A + c_2 B$. By construction, the Lorentz transformation leaves the squared length of the vector invariant,

$$(A^{'0})^2 - \boldsymbol{A}'^2 = (A^0)^2 - \boldsymbol{A}^2 \tag{54}$$

For two 4-vectors A and B we have also the invariance of the inner product,

$$A^{'0} B^{'0} - \boldsymbol{A}' \cdot \boldsymbol{B}' = A^0 B^0 - \boldsymbol{A} \cdot \boldsymbol{B} \tag{55}$$

(55) follows from applying (54) to the 4-vector $A + B$ (Show it!).

A. 4-vectors: modern perspective

Following the modern point of view [**?**] we associate with a given "event" the invariant object, $\boldsymbol{e}_\mu x^\mu$ where the set of basis vectors $\boldsymbol{e}_\mu$ is determined by the observer or more precisely by the frame of reference. Here and everywhere the repeated indices are summed over. The objectiveness of the event is expressed by the following identity,

$$\boldsymbol{e}_\mu x^\mu = \boldsymbol{e}'_\mu x'^\mu \tag{56}$$

In (56) the (un)primed quantities refer to the reference frame (K)K'. (56) means that in relativity we understand the Lorentz transformation as a "passive" one. (Any symmetry operation can be regarded as the transformation of the object without touching the coordinate axes (active) or the coordinate axis without touching the object in the opposite sense (passive)). If the coordinate transform according to the transformation

$$x'^\mu = \Lambda^\mu{}_\nu(\boldsymbol{v}) x^\nu \,, \tag{57}$$

where the Lorentz transformation matrix Λ is fixed by the velocity $\boldsymbol{v}$ of K' relative to K, the transformation of the basis vectors is then fixed in order to satisfy (56),

$$\boldsymbol{e}'_\mu = \Lambda^\nu{}_\mu(-\boldsymbol{v}) \boldsymbol{e}_\nu = [\Lambda^{-1}]^\nu{}_\mu(\boldsymbol{v}) \boldsymbol{e}_\nu \,, \tag{58}$$

where we inverted the sing of the velocity in order to define the matrix in (58) as the inverse to the matrix in (57):

$$\Lambda^\alpha{}_\nu(\boldsymbol{v}) \Lambda^\nu{}_\beta(-\boldsymbol{v}) = \delta^\alpha_\beta \tag{59}$$

(56) is then satisfied by construction,

$$\boldsymbol{e}'_\mu x'^\mu = \Lambda^\alpha{}_\mu(-\boldsymbol{v})\boldsymbol{e}_\alpha \Lambda^\mu{}_\beta(\boldsymbol{v})x^\beta = \boldsymbol{e}_\alpha x^\alpha \tag{60}$$

For another important instance of invariance is obtained by considering the invariant function $\phi(x^\mu)$ such as the phase in the wave. By definition the transformation law,

$$K : \phi(x^\mu) \Rightarrow K' : \phi'(x'^\mu) = \phi(x^\mu) \tag{61}$$

for the coordinates x^μ and x'^μ related by (57) We will also call such a function *(may be or may not be Lorentz) scalar*. Consider the differential,

$$d\phi(x^\mu) = \phi(x^\mu + dx^\mu) - \phi(x^\mu) = \phi'(x'^\mu + dx'^\mu) - \phi'(x'^\mu) = d\phi'(x'^\mu) \tag{62}$$

based on the definition (61). From (62) the differential of the scalar function is itself a scalar:

$$d\phi(x^\mu) = d\phi'(x'^\mu) \tag{63}$$

On the other hand we can write,

$$\frac{\partial\phi}{\partial x^\mu}dx^\mu = \frac{\partial\phi'}{\partial x'^\mu}dx'^\mu \tag{64}$$

Comparing it with (56) and (58) we get the transformation law for gradients,

$$\frac{\partial\phi'}{\partial x'^\mu} = \Lambda^\nu{}_\mu(-\boldsymbol{v})\frac{\partial\phi}{\partial x^\mu} \tag{65}$$

Lets check (65) directly, by noticing that from (57)

$$\Lambda^\mu{}_\nu(\boldsymbol{v}) = \frac{\partial x'^\mu}{\partial x^\nu} \tag{66}$$

and then by the chain rule,

$$\frac{\partial\phi'(x'^\mu)}{\partial x'^\mu} = \frac{\partial\phi(x^\mu)}{\partial x'^\mu} = \frac{\partial\phi(x^\mu)}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial x'^\mu} \tag{67}$$

We wish to demonstrate that

$$\frac{\partial x^\alpha}{\partial x'^\mu} = \Lambda^\alpha{}_\mu(-\boldsymbol{v}) \tag{68}$$

Indeed,

$$\frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x'^\mu}{\partial x^\mu} = \delta^\alpha_\beta \tag{69}$$

and as there is only one inverse to invertible matrix we have shown (68).
DEFINITION The objects that transform as x^μ, namely according to the matrix Λ are called (in old terminology that we will follow) the **contravariant** vectors.
The objects that transform like $\boldsymbol{e}_\mu$ or equivalently as $\frac{\partial\phi}{\partial x^\mu}$, namely according to the *inverse* matrix Λ^{-1} are called (in old terminology that we will follow) the **covariant** vectors.
These terminology does not really fit to the presented description. Nevertheless we will follow it as is done in most of the textbooks used.

B. Metric spaces

Up until now we did **NOT** specify what the transformation matrix Λ is. And in fact it can be anything if nothing else is given. In physical problems however, often the physics provides a further constrain which is the so called **metric**.

1. *Special relativity*

In the special relativity we have the invariant,

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (x'^0)^2 - (x'^1)^2 - (x'^2)^2 - (x'^3)^2 \tag{70}$$

Which tells us that the objects,

$$x_\mu = g_{\mu\nu} x^\nu \tag{71}$$

with the metric,

$$g = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{72}$$

must transform covariantly because

$$x_\mu x^\mu = x'_\mu x'^\mu \tag{73}$$

shows that x_μ transform exactly as $\boldsymbol{e}_\mu$. This imposes the constrain on the possible transformation matrices,

$$x'_\mu = \Lambda^\alpha{}_\mu(-\boldsymbol{v}) x_\alpha \Rightarrow g_{\mu\nu} x'^\nu = \Lambda^\alpha{}_\mu(-\boldsymbol{v}) g_{\alpha\beta} x^\beta \Rightarrow g_{\mu\nu} \Lambda^\nu{}_\gamma(\boldsymbol{v}) x^\gamma = \Lambda^\alpha{}_\mu(-\boldsymbol{v}) g_{\alpha\beta} x^\beta \tag{74}$$

As a result,

$$g_{\mu\nu} \Lambda^\nu{}_\beta(\boldsymbol{v}) = \Lambda^\alpha{}_\mu(-\boldsymbol{v}) g_{\alpha\beta} \tag{75}$$

Multiplying it by $\Lambda^\mu{}_\delta(\boldsymbol{v})$ we finally get,

$$\Lambda^\mu{}_\delta(\boldsymbol{v}) g_{\mu\nu} \Lambda^\nu{}_\beta(\boldsymbol{v}) = \Lambda^\mu{}_\delta(\boldsymbol{v}) \Lambda^\alpha{}_\mu(-\boldsymbol{v}) g_{\alpha\beta} = g_{\delta\beta} \tag{76}$$

which is equivalent to the defining relation inthe matrix form,

$$\Lambda^{trans}(\boldsymbol{v}) g \Lambda(\boldsymbol{v}) = g \tag{77}$$

2. *3D rotations*

In the case of ordinary rotations in three-dimensions we have the condition of conservation of the length of the vector,

$$(x^1)^2 + (x^2)^2 + (x^3)^2 = (x'^1)^2 + (x'^2)^2 + (x'^3)^2 \tag{78}$$

which allows us to identify the covariant component as simply equal to their contra-variant counterparts,

$$x_j = x^j \tag{79}$$

which is the reason why in the purely Euclidian space there is no distinction between the contra- and co-variant components. It then follows that the transformation of the conta- and co-variant components is the same in this case. The comparison of (57) and (58) gives,

$$[\Lambda^{-1}]^\nu{}_\mu = [\Lambda]^\nu{}_\mu \tag{80}$$

or on other words,

$$\Lambda^{-1} = \Lambda^{tr} \tag{81}$$

which tells us that the transformation matrices are orthogonal as is also well known.

3. Spinors space with SU(2) symmetry

Here we consider the space of spinors which describe the quantum mechanical spin 1/2 particle such as electron,

$$\begin{bmatrix}\psi^1\\ \psi^2\end{bmatrix} = \begin{bmatrix}\psi_\uparrow\\ \psi_\downarrow\end{bmatrix}. \tag{82}$$

In the present discussion the logic is reversed. The transformations acting on the spinors are prescribed and we are asking if we could turn the spinor space into a metric space. In other words we are looking for the linear transformation of index lowering

$$\psi_\mu = g_{\mu\nu}\psi^\nu \tag{83}$$

such that the combination

$$\psi_\mu\psi^\mu = g_{\mu\nu}\psi^\mu\psi^\nu \tag{84}$$

is invariant under all the specified transformations.

By construction, spinors transform under spatial rotations according to

$$\begin{bmatrix}\psi'^1\\ \psi'^2\end{bmatrix} = \hat{U}\begin{bmatrix}\psi^1\\ \psi^2\end{bmatrix}, \hat{U} = \begin{bmatrix}a & b\\ c & d\end{bmatrix} \tag{85}$$

with the matrix $\hat{U}$ is unitary and has determinant one, namely it belongs to SU(2) group, App.C. The inverse and Hermitian conjugated matrix read,

$$\hat{U}^{-1} = \begin{bmatrix}d & -b\\ -c & a\end{bmatrix}, \qquad \hat{U}^\dagger = \begin{bmatrix}a^* & c^*\\ b^* & d^*\end{bmatrix} \tag{86}$$

and then unitarity, $\hat{U}^\dagger = \hat{U}^{-1}$ gives

$$\hat{U} = \begin{bmatrix}a & b\\ -b^* & a^*\end{bmatrix} \tag{87}$$

We are given the invariance of the combination

$$\psi^{*1}\psi^1 + \psi^{*2}\psi^2 \tag{88}$$

We should find the combinations of $\psi^{1,2}$ that transform in the same way as $\psi^{*1,2}$ For the complex conjugated spinor the transformation matrix,

$$\begin{bmatrix}\psi'^{*1}\\ \psi'^{*2}\end{bmatrix} = \begin{bmatrix}a^* & b^*\\ -b & a\end{bmatrix}\begin{bmatrix}\psi^{*1}\\ \psi^{*2}\end{bmatrix} \tag{89}$$

Consider the combination,

$$\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix} \equiv \begin{bmatrix}\psi^2\\ -\psi^1\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}\psi^1\\ \psi^2\end{bmatrix} \tag{90}$$

The transformation law of this object is

$$\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix} \to \begin{bmatrix}\psi_1'\\ \psi_2'\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}\psi'^1\\ \psi'^2\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\hat{U}\begin{bmatrix}\psi^1\\ \psi^2\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\hat{U}\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix} \tag{91}$$

Now we notice that,

$$\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\hat{U}\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix} = \begin{bmatrix}a^* & b^*\\ -b & a\end{bmatrix} \tag{92}$$

In other words we have found the metric,

$$g = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix} \tag{93}$$

In particular for two spinors we get a scalar,

$$\psi^1\phi^2 - \psi^2\phi^1 \tag{94}$$

which in the theory of angular momentum is known as "singlet".

Appendix A: Lorentz transformation for the velocity $\boldsymbol{\beta} \nparallel \hat{x}$

We now consider the K' moving with arbitrary velocity $\boldsymbol{v}$ relative to K. We still can assume the axes of K are parallel to the axes of K'. Decompose the vector $\boldsymbol{x}$ into the components parallel and perpendicular to $\boldsymbol{v}$,

$$\boldsymbol{x} = \boldsymbol{x}_\parallel + \boldsymbol{x}_\perp, \qquad \boldsymbol{x}_\parallel = \frac{\boldsymbol{v}}{v}(\boldsymbol{x} \cdot \frac{\boldsymbol{v}}{v}), \quad \boldsymbol{x}_\perp = \boldsymbol{x} - \boldsymbol{x}_\parallel = \boldsymbol{x} - \frac{\boldsymbol{v}}{v}(\boldsymbol{x} \cdot \frac{\boldsymbol{v}}{v}) \tag{A1}$$

The Lorentz transformation then gives

$$\begin{aligned} x^{'0} &= \gamma(x^0 - \beta x_\parallel) \\ \boldsymbol{x}' &= \gamma(\boldsymbol{x}_\parallel - \boldsymbol{\beta} x^0) + \boldsymbol{x}_\perp \end{aligned} \tag{A2}$$

Sometimes we write (A2) as

$$\begin{aligned} x^{'0} &= \gamma(x^0 - \beta x_\parallel) \\ x'_\parallel &= \gamma(x_\parallel - \beta x^0) \\ \boldsymbol{x}'_\perp &= \boldsymbol{x}_\perp \end{aligned} \tag{A3}$$

Substituting (A1) in (A2) we get,

$$\begin{aligned} x^{'0} &= \gamma(x^0 - \beta(\boldsymbol{x} \cdot \frac{\boldsymbol{v}}{v})) = \gamma(x^0 - (\boldsymbol{x} \cdot \boldsymbol{\beta})) \\ \boldsymbol{x}' &= \gamma(\frac{\boldsymbol{v}}{v}(\boldsymbol{x} \cdot \frac{\boldsymbol{v}}{v}) - \boldsymbol{\beta} x^0) + \boldsymbol{x} - \frac{\boldsymbol{v}}{v}(\boldsymbol{x} \cdot \frac{\boldsymbol{v}}{v}) \end{aligned} \tag{A4}$$

Writing

$$\frac{\boldsymbol{v}}{v}(\boldsymbol{x} \cdot \frac{\boldsymbol{v}}{v}) = \frac{\boldsymbol{\beta}}{\beta^2}(\boldsymbol{x} \cdot \boldsymbol{\beta}) \tag{A5}$$

we have

$$\boldsymbol{x}' = \gamma(\frac{\boldsymbol{\beta}}{\beta^2}(\boldsymbol{x} \cdot \boldsymbol{\beta}) - \boldsymbol{\beta} x^0) + \boldsymbol{x} - \frac{\boldsymbol{\beta}}{\beta^2}(\boldsymbol{x} \cdot \boldsymbol{\beta}) = \boldsymbol{x} + \frac{\gamma - 1}{\beta^2}\boldsymbol{\beta}(\boldsymbol{x} \cdot \boldsymbol{\beta}) - \gamma \boldsymbol{\beta} x^0 \tag{A6}$$

In summary, combining (A4) and (A6)

$$\begin{aligned} x^{'0} &= \gamma(x^0 - (\boldsymbol{x} \cdot \boldsymbol{\beta})) \\ \boldsymbol{x}' &= \boldsymbol{x} + \frac{\gamma - 1}{\beta^2}\boldsymbol{\beta}(\boldsymbol{x} \cdot \boldsymbol{\beta}) - \gamma \boldsymbol{\beta} x^0 \end{aligned} \tag{A7}$$

Appendix B: Relativistic Doppler shift

Reversing our previous arguments we can show that if $A^0 B^0 - \boldsymbol{A} \cdot \boldsymbol{B}$ is a scalar and A is a 4-vector then B is also 4 vector.

Consider the wave with the phase $\phi = \omega t - \boldsymbol{k} \cdot \boldsymbol{x}$ which is a scalar. Indeed all the observers will detect for instance a maximum at the same 4-vector event , so the phase of 2π will be the same for all observers. The same is for minima and by linearity of the transformation for all phases. It follows that the 4 quantities, $(k^0 = \omega/c, \boldsymbol{k})$ form a 4-vector. In other words they transform as (A3) or (53), which means,

$$\begin{aligned} k^{'0} &= \gamma(k^0 - \beta k_\parallel) \\ k'_\parallel &= \gamma(k_\parallel - \beta k^0) \\ \boldsymbol{k}'_\perp &= \boldsymbol{k}_\perp \end{aligned} \tag{B1}$$

For light, $|\boldsymbol{k}| = k_0$, and $|\boldsymbol{k}'| = k'_0$ We have

$$\boldsymbol{\beta} \cdot \boldsymbol{k} = \beta|\boldsymbol{k}|\cos\theta = \beta k_0 \cos\theta \Rightarrow k'_0 = \gamma(k_0 - k_0 \beta \cos\theta) \Rightarrow \omega' = \gamma\omega(1 - \beta\cos\theta) \tag{B2}$$

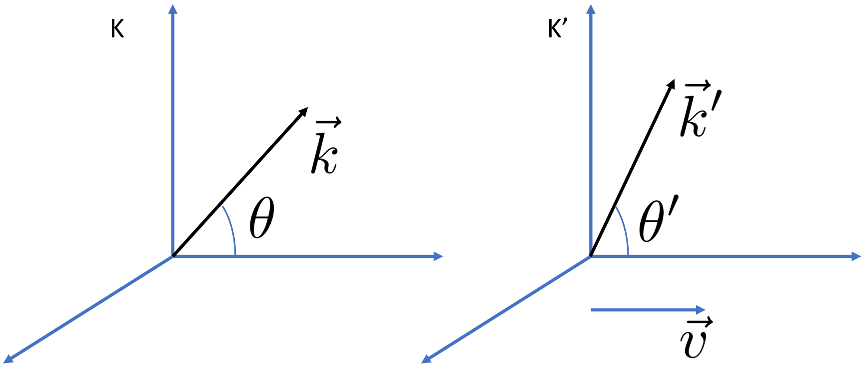

FIG. 5: Doppler shift

with γ replaced by 1 it gives the classical result for the Doppler shift. But as $\gamma > 1$ the effect exists even for $\theta = \pi/2$. As was confirmed experimentally, $\omega' > \omega$ at $\theta = \pi/2$.

Let's discuss the light aberration, which is the change in the direction of light propagation as a result of motion of the observer.

$$\tan\theta' = \frac{k'_\perp}{k'_{||}} = \frac{k_\perp}{\gamma(k_{||} - \beta k_0)} \tag{B3}$$

Dividing the numerator and denominator of (B3) by k_0 and using the relations, $\sin\theta = k_\perp/|\boldsymbol{k}| = k_\perp/k_0$, $\cos\theta = k_{||}/|\boldsymbol{k}| = k_{||}/k_0$ we rewrite (B3) as

$$\tan\theta' = \frac{k'_\perp}{k'_{||}} = \frac{\sin\theta}{\gamma(\cos\theta - \beta)} \tag{B4}$$

And again (B4) goes over to the non-relativistic result when γ is set to 1 as expected.

Appendix C: SU(2) group

The unitary property is needed by physical reasons of invariance of the probability to find the particle at a given point $|\psi_1|^2 + |\psi_2|^2$. The requirement of unit determinant follows from the transformation properties of the combination $\psi^1\phi^2 - \psi^2\phi^1$ which transform in a simple way, $\psi'^1\phi'^2 - \psi'^2\phi'^1 = \det\hat{U}(\psi^1\phi^2 - \psi^2\phi^1)$. Consistency with having a bilinear scalar in the theory gives $\det\hat{U} = 1$.