1 Introduction

First, we shall introduction the two assumptions from Einstein as below:

- The laws of physics are preserved for all observers in inertial frame of reference relative to one another. (law of relativity)
- The speed of light in vacuum is the same for all observers regardless of their motion.

In relativity, we know that there is no more difference between space and time, and we call them altogether as *spacetime*. And we define the spacetime difference between two points is

$$(\Delta s)^2 = (space\ interval)^2 - c^2(time\ interval)^2$$

and if we write out small displacement, we'll get

$$(ds)^2 = (dx^2 + dy^2 + dz^2) - (cdt)^2$$

The value of ds will determine the properties of the interval. For instance, when the object is moving in the speed that is slower than that of the light, then $(ds)^2 < 0$ and thus we call the interval is *timelike*; and when $(ds)^2 > 0$ we would call the interval *spacelike*; when $(ds)^2 = 0$, we call it *lightlike*. What is the difference and physical meaning behind this seperation?

First, if the distance between two events is *timelike*, that is, $(ds)^2 = (cdt)^2 - (dx^2 + dy^2 + dz^2) > 0$, dt cannot be 0, physically, we say that these two events cannot happen **at the same time** but can happen **at the same place**.

Second, if the distance between two events is *spacelike*, that is, $(ds)^2 = (cdt)^2 - (dx^2 + dy^2 + dz^2) < 0$, dt can be 0 this time and there must be space interval between these events, in other words, these two events can only happen **at the same time** but not **at the same place**.

Lastly, if the distance betwween two events is *lightlike*, that is, $(ds)^2 = (cdt)^2 - (dx^2 + dy^2 + dz^2) = 0$, we say light can travel between these two events.

Since the spacetime interval is the geometric properties of events or objects, its value should be preserved in all inertial frame, in other words, $(ds)^2$ is an invariant quantity called **invariant spacetime interval**. Because in relativity, the concept of time is no longer be seen as an independent varialbe that together with coordinates describe the motion of objects. However, we're eager to find a new concept that makes things easier, so we define a new concept about time, that is, **proper time**. The meaning of the proper time is that it can be seen as the motion of object has been measureed in **rest frame**, in other words, we can always choose an inertial frame that the object is at rest in it, and then we call the time in this rest time **proper time**. Now assume an object moving with velocity **v**, and the proper time is denoted as τ

$$(ds)^2 = (ds')^2$$
$$-c^2(d\tau)^2 = (vdt)^2 - c^2(dt)^2$$

And then we'll have

$$d\tau = dt\sqrt{1 - \frac{v^2}{c^2}} = \frac{1}{\gamma}dt$$

2 Vectors and the Metric Tensor

In relativity, we use a special notaion to express the operation in this study. For instance, we would view time and position coordinates as same thing, and denote them as x^μ, where $x^0 = ct$ is the time coordinate, and x^1, x^2, x^3 are the space coordinates. And if we consider time together with our position coordinates, then we shall define two different **vectors** that are useful to us.

- covariant vector: $\mathbf{x}^\mu =< x^0, x^1, x^2, x^3 >=< ct, x, y, z >$
- contravariant vector: $\mathbf{x}_\mu =< -x^0, x^1, x^2, x^3 >=< -ct, x, y, z >$

The transform between them can be achieved by introducing a special tensor called **Minkowski tensor**, one can adapt +- - - or -+++ convention as one prefer, here we pick the first one.

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

$$\mathbf{x}_\mu = \eta_{\mu\nu}\mathbf{x}^\nu$$

Conversely, the inverse relationship is as below:

$$\eta^{\mu\nu} = (\eta_{\mu\nu})^{-1}$$

$$\mathbf{x}^\mu = \eta^{\mu\nu}\mathbf{x}_\nu$$

3 Action

First, we consider the motion of **free particle** so that the potential is not in our discussion. In relativity, time t should be considered in the generalized coordinates, that is, time t become the $n+1$ th coordinate q_{n+1}. The action should be scalar (invariant) and the natural choice of the invariant parameter may be the **proper time of the particle** τ. However, we know that the action should involve generalized coordinate and generalized velocities, and the component of the generalized velocities would obey the following relation,

$$u \cdot u = u^\nu u_\nu = c^2$$

which is the constraint that makes the generalized velocities dependent. Therefore, we must choose another invariant parameter. Suppose we pick an arbitrary invariant parameter θ, and the generalized velocities is defined as below

$$x'^\nu \equiv \frac{dx^\nu}{d\theta}$$

Then Hamilton principle can be written as

$$\delta S = \int_{\theta_1}^{\theta_2} \Lambda(x^\nu, x'^\nu)\, d\theta$$

And the Lagrange's equation would become

$$\frac{d}{d\theta}\left(\frac{\partial \Lambda}{\partial x'^{\mu}}\right) - \frac{\partial \Lambda}{\partial x^{\mu}} = 0$$

Where the Lagrange fuction Λ is the world scalar(invariant). We should try to make connection between the invariant Lagrange function Λ and the previous Lagrange fuction L. One way is to change the orignal parameter time t to the invariant parameter.

$$\dot{x}^{\mu} = \frac{dx^{\mu}}{dt} = c\frac{dx^{\mu}}{d\theta}\frac{d\theta}{dct} = c\frac{x'^{\mu}}{x'^{0}}$$

$$\delta S = \int_{t_1}^{t_2} L(x, \dot{x}, t)\, dt = \int_{\theta_1}^{\theta_2} L(x^{\mu}, c\frac{x'^{\mu}}{x'^{0}})\, \frac{x'^{0}}{c}\, d\theta$$

Compare with the action we have mentioned previously in this section, we should have

$$\Lambda(x^{\mu}, x'^{\mu}) = \frac{x'^{0}}{c} L(x^{\mu}, c\frac{x'^{\mu}}{x'^{0}})$$

We could discuss much deeper into this transformation. First, we know that the light speed is not so important and can be set to 1 if we change the units. x'^{0} is the derivative of time t with respect to the invariant parameter θ. Since we haven't find out the physical meaning of this invariant parameter, we're not able to discuss the physical meaning of the new Lagrangian here; however, we know that this equation is caused by considering time t as the $n+1$ th of the gerneralized parameter rather than an independent parameter. Yet, there are some mathemetic properties and consequence that we'd be interested in. First, consider scalar transformation, that is, $\Lambda(x^{\mu}, x'^{\mu}) \rightarrow \Lambda(x^{\mu}, \alpha x'^{\mu})$. And then the most beautiful thing happens, by the transformation equation above

$$\Lambda(x^{\mu}, \alpha x'^{\mu}) = \frac{\alpha x'^{0}}{c} L(x^{\mu}, c\frac{\alpha x'^{\mu}}{\alpha x'^{0}}) = \alpha\frac{x'^{0}}{c} L(x^{\mu}, c\frac{x'^{\mu}}{x'^{0}})$$

Which is exacly equal to the orignal relativistic Lagrangian multiplied by a scalar, that is, $\alpha\Lambda(x^\mu, x'^\mu)$. From this powerful properties, we can see that the relativistic Lagrangian is a homogeneous function of the generalized velocities in the first degree. It follows from Euler's theorem on homogeneous functions

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

And because of this properties, we'd have the following equation holds

$$\left[\frac{d}{d\theta} \left(\frac{\partial \Lambda}{\partial x'^\mu} \right) - \frac{\partial \Lambda}{\partial x^\mu} \right] x'^\mu = 0$$

The proof is as below

4 Solution of EM-SR

If the light is passing through the space, and be observed by two inertial frames S and S'.

$$\begin{aligned} \eta_{\mu\nu} dx^\mu dx^\nu &= (dx)^2 + (dy)^2 + (dz)^2 - (dt)^2 \\ &= \eta_{\mu\nu} dx'^\mu dx'^\nu \end{aligned}$$

Then, we have,

$$x^\mu = < x^0, x^1, x^2, x^3 >$$

$$x'^\mu = < x'^0, x'^1, x'^2, x'^3 >$$

and also Lorentz transformation have the form,

$$x'^{\mu} = \Lambda^{\mu}{}_{\rho} x^{\rho}$$

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

To prove that $\eta = \Lambda^T \eta \Lambda$, we could use the properties of **invariant**, that is, **Minkowski norm**.

$$\|x\|^2 = \|x'\|^2$$

And then we have,

$$\|x\|^2 = \eta_{\mu\nu} x^{\mu} x^{\nu} = \eta_{\rho\gamma} x'^{\rho} x'^{\gamma}$$

$$= \eta_{\rho\gamma} (\Lambda^{\rho}{}_{\mu} x^{\mu})(\Lambda^{\gamma}{}_{\nu} x^{\nu})$$
$$= \eta_{\rho\gamma} \Lambda^{\rho}{}_{\mu} \Lambda^{\gamma}{}_{\nu} x^{\mu} x^{\nu}$$

for all x^{μ}, x^{ν}, that is, we'll have,

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

If we consider the transpose of $\Lambda^{\rho}{}_{\mu}$, and we know that changing the order of the back and forth indexes of Lorentz matrix is just the same as the ordinary matrix, even though the back and forth indexes are not at the same horizontal line in Lorentz matrix.

$$A \to A^T$$

$$a_{ij} \to a_{ji}$$

$$\Lambda \to \Lambda^T$$

$$\Lambda^{\rho}{}_{\mu} \to \Lambda_{\mu}^{\rho}$$

Then we'll get,

$$\eta_{\mu\nu} = \left(\Lambda^T\right)_\mu^{\ \rho} \eta_{\rho\gamma} \Lambda^\gamma_{\ \nu}$$

$$\eta = \Lambda^T \eta \Lambda$$

5 Lorentz Group

In this section, we'll going to prove that Lorentz transformation forms a group called **Lorentz group**. First, we list the definition of group,

- **Closure**, the result of the product of two elements also belongs to the group.

$$\Lambda_1, \Lambda_2 \in G, \ \ s.t. \ \Lambda_1 \Lambda_2 \in G$$

- **Associativity**, the order of doing product wouldn't affect the result.

$$\Lambda_1 \left(\Lambda_2 \Lambda_3\right) = \left(\Lambda_1 \Lambda_2\right) \Lambda_3$$

- **Identity**, exists a element called **identity** such that other elements' product with identity will equal to themselves.

$$I \in G, \ \ s.t. \ \Lambda I = I \Lambda = \Lambda$$

- **Inverse**, for all elements in group, there is an **inverse** such that the product of element and its inverse will equal to identity.

$$\Lambda^{-1} \in G, \ \ s.t. \ \Lambda \Lambda^{-1} = I$$

Then we should prove the existence of Lorentz group $G = \{\Lambda_1, \Lambda_2, ...\}$. Note that we have a very useful equation $\Lambda^T \eta \Lambda = \eta$. First, **Closure**.

$$\Lambda_1^T \eta \Lambda_1 = \eta, \ \ \Lambda_2^T \eta \Lambda_2 = \eta$$

consider

$$\left(\Lambda_1 \Lambda_2\right)^T \eta \ \Lambda_1 \Lambda_2 = \Lambda_2^T \Lambda_1^T \eta \Lambda_1 \Lambda_2$$

$$= \eta$$

Then we can see that $\Lambda_1\Lambda_2 \in G$.

Second, **Associativity**, we can see that Lorentz transformation can be expressed in matrix form. Also, the associativity can be applied on matrix product, so we automatically have the associativity in Lorentz transformation.

Third, **Identity**. We merely have to prove that **identity** I in Lorentz group, that is,

$$I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$I^T \eta I = \eta$$

Hence, $I \in G$.

Last, **Inverse**. To prove that Lorentz transformation has inverse, we merely need to prove that its determinant is not 0 just like we have to do when we want to prove the matrice have inverse. Moreover, we'll better find the explict form of inverse. Again, we should use the equation $\Lambda\eta\Lambda = \eta$

$$\det\left(\Lambda^T \eta \Lambda\right) = \det(\eta)$$

we have,

$$\det\left(\Lambda^T\right) = \det\left(\Lambda\right)$$

Then,

$$\det\left(\Lambda\right)^2 \det(\eta) = \det(\eta)$$

And we have η is an invertible matrix so that $\det(\eta) \neq 0$, or one can compute it explictly. Then,

$$\det\left(\Lambda\right) = \pm 1$$

Hence, **Inverse** holds, and Lorentz transformation form Lorentz group.