Special Theory of Relativity

by

Indrajeet S. Patil*

T.Y.B.Sc., Fergusson College, Pune-411 004

Submitted to:

Indian Academy of Science's

Summer Research Fellowship Programme

Under Guidance of

Prof. Jnanadeva Maharana,

Institute of Physics,

Bhubaneshwar-751 005

June 1, 2008

*inderonline1988@gmail.com

"*There are two kinds of happiness or contentment for which mortals are adapted; the first we experience in* thinking *and the other in* feeling. *The first is the purest and most unmixed. Let a man once know what sort of being he is; how great the being which brought him into existence, how utterly transitory is everything in the material world, and let him realize this without passion in a quiet philosophical temper, and I maintain that he is happy; as happy indeed as it is possible him to be.*"

– **William Herschel**, from a letter to his brother

Contents

1 Galilean/Newtonian Relativity

1.1 Galilean transformations

Galileo's principle of relativity states that no mechanical experiment carried out entirely in one inertial[1] frame can tell the observer what the motion of that frame is with respect to any other inertial frame. See Resnick(Ref.[1]). To understand this in a better manner, let us consider two inertial frames of reference S and S' which have common origins at $t = t' = 0$. Let S' be moving uniformly away from S along positive-x direction with velocity $\vec{V}$, then Galilean transformations relating set of co-ordinates (t, x, y, z) to (t', x', y', z') are:

$$\boxed{x = x' + V_x t, \quad y = y', \quad z = z', \quad t = t'} \tag{1}$$

To get (t', x', y', z') in terms of (t, x, y, z), we just need to put $-V_x$ in above equations. In general, in vector notation, the above transformations(except for time) can be written as: $\vec{r} = \vec{r'} + \vec{V}t$.

In Newtonian mechanics, it is assumed that the length intervals and time intervals are absolute, i.e., that they are the same for all inertial observers of the same events. e.g. if meter sticks are of the same length when compared at rest with respect to one another, it is implicitly assumed that they are of the same length when compared in relative motion to one another. Similarly, if clocks are synchronized when at rest, it is assumed that they remain synchronized thereafter. Though these assumptions seem legitimate at low-speeds, for which Newtonian mechanics gives excellent results, , they begin to break down at high-speeds. Differentiating Equation(1) w.r.t. time,

$$\frac{dx}{dt} = \frac{dx'}{dt'} + V_x$$

$$\therefore v_x = v'_x + V_x, \quad v_y = v'_y, \quad v_z = v'_z \tag{2}$$

since $t = t', dt = dt'$. In vector notation, $\vec{v} = \vec{v'} + \vec{V}$. Differentiating once again,

$$a_x = a'_x, \quad a_y = a'_y, \quad a_z = a'_z$$

[1] A non-inertial observer will obviously be able to tell whether he/she is in motion because of the pseudo-forces.

$$\therefore \vec{a} = \vec{a'}$$

In classical physics, it is assumed that mass of particle is constant, independent of its motion with respect to an oberver. Thus,

$$\boxed{\vec{F} = \vec{F'}} \tag{3}$$

Thus, Newton's laws of motion are *invariant* under Galilean transformations. But what do we mean by that?

The quantities which remain unchanged under a given transformation are called invariants of that transformation, e.g acceleration is invariant under Galilean transformation. A statement of what the invariant quantities are is called a relativity principle; it says that for such quantities the reference frames are equivalent to one another, no one having an absolute or a priviliged status relative to the others.

Newton's laws of motion and the equations of motion of a particle would be exactly same in all inertal frames. Since in mechanics, all the conservation principles are upshots of Newton's laws, it follows that the laws of *mechanics* are the same in all inertial frames. Note that we do not include laws of electromagnetism here for the reasons that wil be discussed later.

1.2 Geometry of spacetime in Classical Physics

Newton's law of motion says that a free particle moves in a straight line path in some inertial frame if it is initially moving in that frame or stays at rest if it was initially at rest. But, straight line is the shortest distance between two points (x_1, y_1, z_1) and (x_2, y_2, z_2) in 3-dimensional Euclidean space. Thus, Newtonian mechanics implicitly assumes that the geometry of space is $\mathbb{R}^3$. Also, it assumes that time flows smoothly for all the observers at the same rate. Thus, the geometry of space time is 4-dimensional flat Euclidean($\mathbb{R}^4$). Any geometry is specified by the line element. The line element of 4-dimensional Euclidean flat space is,

$$\boxed{ds^2 = dx^2 + dy^2 + dz^2 + dt^2} \tag{4}$$

Conventionally, $(ds)^2$ is written as ds^2 and so on.

Galilean relativity requires that this line element preserves its form under transformations like dispacement or rotation, e.g. Let's consider rotation in the tz-axis, then new coordinates are given by,

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ \cos\varphi & \sin\varphi & 0 & 0 \\ -\sin\varphi & \cos\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$$

Then, it can be easily seen that,

$$ds^2 = dx^2 + dy^2 + dz^2 + dt^2 = dx'^2 + dy'^2 + dz'^2 + dt'^2 = ds'^2$$

Thus, the form of the line element is invariant under rotations.

1.3 Galilean transformations and electromagnetism

In section(1.1), we talked about invariance of laws of mechanics under Galilean transformations. Now, let us see if laws of electromagnetism are also invariant under the same transformations.

Maxwell's theory of electromagnetism gives us the electromagnetic wave equation,

$$\frac{\partial^2 \vec{E}}{\partial x^2} + \frac{\partial^2 \vec{E}}{\partial y^2} + \frac{\partial^2 \vec{E}}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = 0$$

in which the constant $c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}$ is identified with the speed of propagation of planave in vacuum, i.e. with the speed of light. The natural question to ask is "*Speed of light is c, but in which frame?*" because if we use velocity addition law of classical mechanics, then the speed of light can take in value in the range $(c - V)$ to $(c + V)$. In other words, speed of light is not invariant under Galilean transformations. Thus, the fact that Galilean transformations apply to Newtonian mechanics but not to Maxwell's theory forces us to choose one of the following three alternatives:

1. The relativity principle holds *only* for mechanics and not for electrodynamics, which means that there exists a preferred frame of reference, in which the spped of light is c and an optical experiment performed

in a given frame of reference can give information about the state of motion of the frame.

2. Relativity principle holds *both* for mechanics and electrodynamics, but the laws of electrodynamics are incorrect. In this case, we should be able to find deviations from Maxwell's theory and classical electrodynamics should be reformulated in such a way that it will be invariant under the Galilean transformations.

3. Relativity principle holds *both* for mechanics and electrodynamics, but the laws of mechanics are incorrect. In this case, we should be able to find out deviations from Newton's laws. We then need to find out new laws of mechanics and new transformation laws, compatible with this new mechanics and electrodynamics.

Experiments performed found out neither deviations from classical electrodynamics nor any evidence for the existence of preferred frame of reference, ruling out first two options.

Thus, the natural starting point to find out new mechanical laws was to assume that classical electrodynamics is correct and laws of physics are invariant under both mechanics and electrodynamics, which inescapably leads to the conclusion that speed of light is same for all the observers. And as we will see shortly, this one of the assumptions of the special relativity and leads[2] to the new transformation laws, known as Lorentz transformations.

[2] The usual treatment is based on the invariance of the speed of light. However, this must not necessarily be the starting point: indeed what is really at stake is the locality of interactions: one supposes that the influence that one particle, say, exerts on another can not be transmitted instantaneously. Hence, there exists a theoretical maximal speed of information transmission which must be invariant, and it turns out that this speed coincides with the speed of light in the vacuum. In an 1964 paper, Erik Christopher Zeeman showed that a, in a mathematical sense, weaker condition, the causality preserving property, is enough to assure that the coordinate transformations be the Lorentz-transformations. Refer[5].

2 Postulates of Special Relativity

2.1 Postulates

Special relativity is based on two postulates, viz.

Principle of relativity : The *laws of physics* (and not just mechanics) are the same in all inertial frames of reference.

Principle of universality of speed of light : The speed of light is same for all the observers, regardless of the motion of the light source relative to the observer.

The best definition of an inertial observer is given in Schutz(Ref.[1]): The coordinate system is said to be inertial if,

1. The distance between two spatial points is independent of time.
2. The clocks that sit at every point ticking off the time coordinate are synchronized and all run at the same rate.
3. The geometry of space at any instant of time is Euclidean.

This definition does not mention whether the observer accelerates or not because it turns out that only an unaccelerated observer can keep his clocks synchronized.

Note that Einstein's principle of relativity is an extension of Galileo's principle of relaivity since it says that *any* kind of physical experiment done entirely within the system can not tell if the frame is intrinsically moving or is at rest.

In addition to these postulates, special relativity assumes isotropy and homogeneity of space, which means that no point of space or no direction is special. In other words, a measurement of length or time interval is not dependent on where or in what direction it was carried out in a given frame of reference. Also, it assumes homogeniety of time which means that no instant of time is special, a measurement of length or time interval carried out does not depend on when it was carried out. The immediate implicatiom of this assumptions is that the transformation laws relating the coordinates of any two frames *can't be non-linear*.

e.g. Suppose that the transformation is non-linear, say of the form $x = kx'^2$, where k is some constant. Then the length transformation will be of the form $x_2 - x_1 = k({x'_2}^2 - {x'_1}^2)$. Consider a rod having a unit length in frame S'. Suppose that the endpoints of rod has coordinates, $x'_2 = 3$ and $x'_1 = 1$, then the corresponding length in frame S will be $x_2 - x_1 = 8k$, while if the endpoints of rod has coordinates, $x'_2 = 4$ and $x'_1 = 2$, then the corresponding length in frame S will be $x_2 - x_1 = 12k$. Thus, the length of the rod depends on where the measurement is carried out, clearly in disagreement with the homogeneity assumption.

2.2 Relativity of simultaneity

Let us try to find out what implication does the second postulate has. Consider a frame of reference in which at each spatial point, we have a clock to measure time. The time of an event is measured by the clock whose location coincides with that of an event. All the clocks are synchronized using Einstein's synchronization procedure, which can be described as follows:

Consider two clocks at a distance of L from each other. Let a light source, say a bulb, is in between them, equidistant from each clock. Let the bulb flashes at some instant of time, which we call $t = 0$. Then the light wavefront takes time $t = L/2c$ to reach to both the clocks. Hence, when the light reaches the clock, the clocks are set to read time as $t = L/2c$. In this manner, all the clocks of the frame are synchronized. This synchronization method is adopted by all the inertial observers.

We now consider two frames of reference S and S' in relative motion. For the sake of simplicity, we assume that S' is moving with a relative velocity of V_x along positive x-direction. At $t = t' = 0$, let us assume that the origins of both the frames were coincident. Consider that there are two light sources in frame S' which emit light at a regular intervals. We have two observers $\mathcal{S}$ and $\mathcal{S}'$ in frames S and S' respectively residing at the midpoint between the two light sources initially. Then, we would like to consider simultaneity of two events:flashing of bulb on the right, which we call $\mathcal{A}$, and flashing of bulb on the left, which we call $\mathcal{B}$.

1. In frame S' where $\mathcal{A}$ and $\mathcal{B}$ sources are at rest, suppose that the signal reaches at the same time at $\mathcal{S}$, so that he concludes that both light sources emitted signals simultaneously.

2. In frame S, both $\mathcal{A}$ and $\mathcal{B}$ are moving along positive x-direction. Thus, observer $\mathcal{S}$ would conclude that the signal emitted from $\mathcal{B}$ will require less time to reach $\mathcal{S}'$ because it is travelling in the direction opposite to the direction of motion of the frame. On the other hand, the signal emitted from $\mathcal{A}$ will require more time to reach $\mathcal{S}'$ because it is travelling in the direction of motion of the frame. Thus, he would conclude that the signal from $\mathcal{A}$ was emitted before the signal emitted from $\mathcal{B}$.

Hence, we see that the events which are simultaneous in one frame are not simultaneous in other, i.e. simultaneity is a relative concept.
Note that the absolute status of simultaneity in classical physics can still survive if we assume that the speed of light depends upon the speed of source. In which case, we can argue that the signal emitted from $\mathcal{A}$ has a speed of $(c + V_x)$, while the signal emitted from $\mathcal{B}$ has a speed of $(c - V_x)$ from the point of view of observer $\mathcal{S}$. But, this clearly contradicts second postulate of relativity. This is why the constancy of speed,of light is so crucial.
Also, we know that the length measurement of a moving rod needs to be carried out simultaneously. Now as the concept of simultaneity is relative, so is the concept of length measurement. This shows how the constancy of speed of light leads to interdependence of spatio-temporal mesurements.

3 Line Element of Minkowski Spacetime

3.1 Line element

Consider two frames of reference S and S' with coordinate axes (t, x, y, z) and (t', x', y', z') respectively. Let's consider two events:sending out of the light signal, propagating with velocity of light[3] and reception of the signal. Let (t_1, x_1, y_1, z_1) and (t_2, x_2, y_2, z_2) be the coordinates of the two events in S. Then, as $c = 1$, we have-

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - (t_2 - t_1)^2 = 0$$

Similarly, in frame S',

$$(x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2 - (t'_2 - t'_1)^2 = 0$$

If (t_1, x_1, y_1, z_1) and (t_2, x_2, y_2, z_2) are coordinates of two events, the quantity

$$\Delta s = [(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - (t_2 - t_1)^2]^{\frac{1}{2}}$$

is called as event interval. For infinitesimally close events,

$$\boxed{ds^2 = dx^2 + dy^2 + dz^2 - dt^2}$$

We posit that this is the line element of the Minkowski spacetime. Note that the *geometry of spacetime* is not a flat four-dimensional Euclidean geometry because of the term"$-dt^2$". It is sometimes called as psuedo-Euclidean.Also, note that the separation between events in Euclidean space can be zero only for coincident events, while in Minkowski space it can be zero even for non-coincident events. We would later derive Lorentz transformations on the basis of this line element.

3.2 Types of event-intervals

Let's consider some event O having coordinates $(0, 0, 0, 0)$ in some frame of reference. Then the square of the event interval between some other event P and O can have either positive, negative or zero value.

[3]Throughout this report, we are going to assume that $c = 1$, i.e. we will be working with natural units.

3.3 Proper time

Now, let's consider a particle moving with a velocity $\vec{v}(t)$ (which can be uniform or non-uniform) in a frame of reference. In infinitesimal time interval dt as measured in a frame of reference, the particle moves a distance of $\sqrt{dx^2 + dy^2 + dz^2}$. But, in particle's own frame of reference, it is obviously at rest, i.e. $dx' = dy' = dz' = 0$. Then, from invariance of event intervals, we have-

$$dx^2 + dy^2 + dz^2 - dt^2 = dx'^2 + dy'^2 + dz'^2 - d\tau^2 = -d\tau^2 \tag{5}$$

where $d\tau$ is the time measured by the particle in the clock that it carries along with itself.

$$\therefore \boxed{ds^2 \equiv -d\tau^2} \tag{6}$$

Here, τ is called as *proper time*, which is time measured by the clock between two events that occur at the same place as that of the clock. In special relativity, coordinate time(t) is reckoned relative only to inertial observers, whereas proper time can be measured by accelerated observers too. From Equation(5),

$$d\tau^2 = dt^2\big(1 - \frac{dx^2 + dy^2 + dz^2}{dt^2}\big) = dt^2(1 - \vec{v^2}(t))$$

$$\because \vec{v^2}(t) = \frac{dx^2 + dy^2 + dz^2}{dt^2}.$$

$$\therefore \boxed{\tau_2 - \tau_1 \equiv \int_{t_1}^{t_2} \sqrt{(1 - \vec{v^2}(t))}dt = \int_{t_1}^{t_2} \frac{dt}{\gamma}} \tag{7}$$

where $\gamma = 1/(1 - v^2(t))^{1/2}$ is known as Lorentz factor. From this equaion, it can be easily seen that the proper time of a moving clock is always less than the corresponding interval in rest frame. In other words, moving clocks run slow! This phenomenon is called as time dilation.

As $\int d\tau$ has its minimum value if it it is taken along the straight line joining the two points on spacetime diagram, $\int ds$ has a maximum value for the same path. Note the difference between the geometries: $\int ds$ has minimum value for straight line path in Euclidean geometry, while maximum value for the same path in psuedo-Euclidean geometry of spacetime.

4 Lorentz transformations

4.1 Derivation based on invariance of Minkowski line element

It is evident that the line element of the Minkowski spacetime is not invariant under Galilean transformations. We would like to find general transformation laws which will keep this line element invariant.

To obtain this transformation, we make use of the fact that any rotation in four-dimensional spacetime (which is Minkowski space[4]) leaves the quantity $x^2 - t^2$ invariant, see Landau and Lifshitz(Ref.[3]).

Thus, let's consider a rotation in the tx-plane; under which y and z coordinates do not change. The relation between old coordinates (x, t) and new coordinates (x', t') is given by,

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cosh\theta & \sinh\theta & 0 & 0 \\ \sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} \tag{8}$$

Now let S be the inertial frame of reference and S' be moving uniformly with respect to S with a velocity[5] V_x along x-axis. We focus on the motion of origin of S' for which $x' = 0$. Thus above equations reduce to-

$$x = t' \sinh\theta \quad \text{and} \quad t = t' \cosh\theta$$
$$\Rightarrow \quad V_x = \frac{x}{t} = \tanh\theta$$

Using identity $\cosh^2 - \sinh^2 = 1$, we obtain,

$$\sinh\theta = \gamma V \quad \text{and} \quad \cosh\theta = \gamma \tag{9}$$

Substituting these expressions in Equation(9), we have:

$$\boxed{x = \gamma(x' + V_x t'), \quad y = y', \quad z = z', \quad t = \gamma(t' + V_x x')} \tag{10}$$

[4]Note that space here is used in a more general sense than just physical space.

[5]Note that we have not written $V(t)$ as in equation(7) because though special relativity can deal with non-uniform motion of *particles* in a given reference frame, it can not deal with non-uniform motion of *reference frames*.

In matrix form, these equations can be written as:

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma V_x & 0 & 0 & \gamma \\ \gamma & 0 & 0 & \gamma V_x \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix}$$

These are the required transformation laws, known as Lorentz transformations. These transformations reduce to Galilean transformations when $V \ll c$. Formal way of doing this is letting $c \to \infty$ which is nothing but $\gamma \to 1$, remembering that in usual units, $\gamma = 1/(1 - V_x^2/c^2)^{1/2}$.
Taking differentials of the above equations, we get-

$$dx = \gamma(dx' + V_x dt'), \quad dy = dy', \quad dz = dz', \quad dt = \gamma(dt' + V_x dx') \tag{11}$$

It can be easily seen that $dx'^2 - dt'^2 = dx^2 - dt^2$.

4.2 Implications of Lorentz transformations

Let's study some of the implications of the Lorentz transformations:

Length contraction : Condider a rod of length dx at rest in frame S. Its length in frame S' is the distance between its two ends measured *simultaneously*, i.e. $dt' = 0$, which from Equation(12) gives,

$$\boxed{dx = \gamma dx'} \tag{12}$$

As $\gamma > 1$ for non-zero velocity, length of rod is maximum in the frame where it is at rest. In all other frames its length contracts by a factor of γ. This phenomenon is called as length contraction. Note that the rod here is immaterial, what it essentially tells us is that the *space* itself contracts, which leads to contraction of material objects.

Time dilation : Consider a clock which is at rest, thus $dx' = 0$, in frame S' measures time interval dt' between two events. Then, from Equation(12), we get-

$$\boxed{dt = \gamma dt'} \tag{13}$$

which is in agreement with Equation(7).

Law of velocity addition : Let's consider a particle which is moving with a velocity of $\vec{v}(t)$ in frame S and $\vec{v'}(t)$ in frame S'. Then,

$$\vec{v}(t) = \frac{d\vec{r}}{dt} \quad \text{and} \quad \vec{v'}(t) = \frac{d\vec{r'}}{dt}$$

In Equation(12), dividing the whole equation by dt, we get-

$$\boxed{v_x = \frac{v'_x + V_x}{1 + v'_x V_x}, \quad v_y = \frac{v'_y + V_x}{\gamma(1 + v'_y V_x)}, \quad v_z = \frac{v'_z + V_x}{\gamma(1 + v'_z V_x)}} \tag{14}$$

which are the required velocity addition laws based on the Lorentz transformations.

Limiting speed : Note from Equation(10) that if $V > c$, then x and t become imaginary which essentially means that such kind of motion is impossible. Any material particle can move with a speed less than that of light, i.e. it will always follow a timelike worldline on spacetime diagram.

5 Mathematics of Four-vectors

In $\mathbb{R}^2$, whenever any pair of coordinates (x, y) in the coordinate system is transformed by rotation of the axis through angle φ into coordinate system with coordinates given by,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

then we define x and y as the components of some vecter, say $\vec{r}$.
Similarly, a set of four quantities (t, x, y, z) are said to form a four-dimensional vector or *four-vector* if length of that vector does not change under any rotations of the four-dimensional coordinate system, in particular under Lorentz transformations.

• *Notation*: We will denote four-vector with a bold face letter (e.g. $\mathbf{r}$), while a three-vector with an arrow overhead (e.g. $\vec{r}$).

We denote components of any vector by x^μ where $\mu = 0, 1, 2, 3$; so that

$$x^0 \equiv t, \quad x^1 \equiv x, \quad x^2 \equiv y, \quad x^3 \equiv z$$

Thus, radius four-vector is given by,

$$\mathbf{r} = x^0\hat{\mathbf{e}_0} + x^1\hat{\mathbf{e}_1} + x^2\hat{\mathbf{e}_2} + x^3\hat{\mathbf{e}_3} = \sum_{\mu=0}^{3} x^\mu \hat{\mathbf{e}_\mu}$$

where $(\hat{\mathbf{e}_0}, \hat{\mathbf{e}_1}, \hat{\mathbf{e}_2}, \hat{\mathbf{e}_3})$ forms an orthonormal basis for flat[6] Minkowski space, i.e. spacetime. Here, μ is called as a dummy or summation index because any repeated index indicates summation, i.e.

$$\mathbf{r} = \sum_{\mu=0}^{3} x^\mu \hat{\mathbf{e}_\mu} = \sum_{\nu=0}^{3} x^\nu \hat{\mathbf{e}_\nu} = \dots$$

The components of four-vector can be specified in one of the following ways,

$$r^\mu = (x^0, x^1, x^2, x^3) = (x^0, \vec{r})$$

[6]Though geometry of spacetime is non-Euclidean, it is still flat in special relativity.

The length of radius vector (which is nothing but event interval between events $(0,0,0,0)$ and (x^0, x^1, x^2, x^3)) is given by,

$$|\mathbf{r}|^2 = -(x^0)^2 + (x^1)^2 + (x_2)^2 + (x_3)^2$$

For convenience, we introduce two types of components of four-vectors, viz.
(i) *contravariant* component; denoted by x^μ and
(ii) *covariant* component; denoted by x_μ.
These are related by,

$$x_0 = -x^0, \quad x_1 = x^1, \quad x_2 = x^2, \quad x_3 = x^3 \tag{15}$$

With this convention, magnitude of any four-vector appears in the form

$$\sum_{\mu=0}^{3} x^\mu x_\mu = \sum_{\mu=0}^{3} x_\mu x^\mu = x^0 x_0 + x^1 x_1 + x^2 x_2 + x^3 x_3 \tag{16}$$

Now, let's define a scalar product for four-vectors: Consider two four-vectors

$$\mathbf{a} = \sum_{\mu=0}^{3} a^\mu \hat{\mathbf{e}_\mu} \quad \text{and} \quad \mathbf{b} = \sum_{\nu=0}^{3} b^\nu \hat{\mathbf{e}}_\nu$$

Then their scalar product is given by,

$$\mathbf{a}.\mathbf{b} = (\sum_{\mu=0}^{3} a^\mu \hat{\mathbf{e}_\mu}).(\sum_{\nu=0}^{3} b^\nu \hat{\mathbf{e}_\nu}) = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} (\hat{\mathbf{e}_\mu}.\hat{\mathbf{e}_\nu}) a^\mu b^\nu$$

We introduce a special notation for the scalar product of the basis vectors,

$$\boxed{\eta_{\mu\nu} \equiv \hat{\mathbf{e}_\mu}.\hat{\mathbf{e}_\nu}} \tag{17}$$

where $\eta_{\mu\nu}$ is called as *metric* of flat spacetime. Thus, the scalar product becomes:

$$\mathbf{a}.\mathbf{b} = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} \eta_{\mu\nu} a^\mu b^\nu \tag{18}$$

The line lement of flat spacetime can then be expressed as scalar product in the following manner:

$$ds^2 = \mathbf{dx}.\mathbf{dx} = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} \eta_{\mu\nu} dx^\mu dx^\nu \tag{19}$$

But, we know the line element of flat spacetime:

$$ds^2 = \overbrace{-(dx^0)^2}^{\text{time component}} + \underbrace{(dx^1)^2 + (dx_2)^2 + (dx_3)^2}_{\text{space components}} \tag{20}$$

Comparing Equations(19, 20), we get

$$\eta_{\mu\nu} = \underbrace{\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\text{Minkowski metric}} = \text{diag}(-1, 1, 1, 1) \tag{21}$$

In similar manner, comparing Equations(), we get

$$\eta_{\mu\nu} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\text{Euclidean metric}} = \text{diag}(1, 1, 1, 1) \tag{22}$$

Then, in general, scalar product in flat spacetime is given by,

$$\boxed{\mathbf{a}.\mathbf{b} = -a^0a^0 + a^1a^1 + a^2a^2 + a^3a^3 = -a^0a^0 + \vec{a}.\vec{b}} \tag{23}$$

We note that because of the way scalar product is defined, it is same in all inertial frames.

6 Special relativistic kinematics and dynamics

Any material particle follows a timelike worldline through spacetime. This curve can be parametrized in terms of some parameter, say α. Then for each value of α, the four functions $x^\mu(\alpha)$ determine a point along the curve. Though a number of parameters can be used, a natural one is the proper time τ. Thus, a world line in described by four equations: $x^\mu = x^\mu(\tau)$.

Then *four-velocity* of a particle ($\mathbf{u}$) is defined as-

$$\boxed{u^\mu \equiv \frac{dx^\mu}{d\tau}} \tag{24}$$

It can be shown easily that: $u^\mu = (\gamma, \gamma\vec{v})$, where $\vec{v}$ is the velocity of particle in frame.

References

[1] R. Resnick, *“Introduction to Special Relativity”*, John Wiley and Sons, 1968.

[2] B. F. Schutz, *“A First Course in General Relativity”*, Cambridge University Press, 1985.

[3] J. B. Hartle, *”Gravity: An Introduction to Einstein’s General Relativity”*, Pearson Education, 2003.

[4] E. M. Lifshitz and L. D. Landau, *”The Classical Theory of Fields: Vol.2”*, Butterworth-Heinemann, 1980.

[5] http://en.wikipedia.org/wiki/Lorentz transformation