Special Relativity and Maxwell's Equations

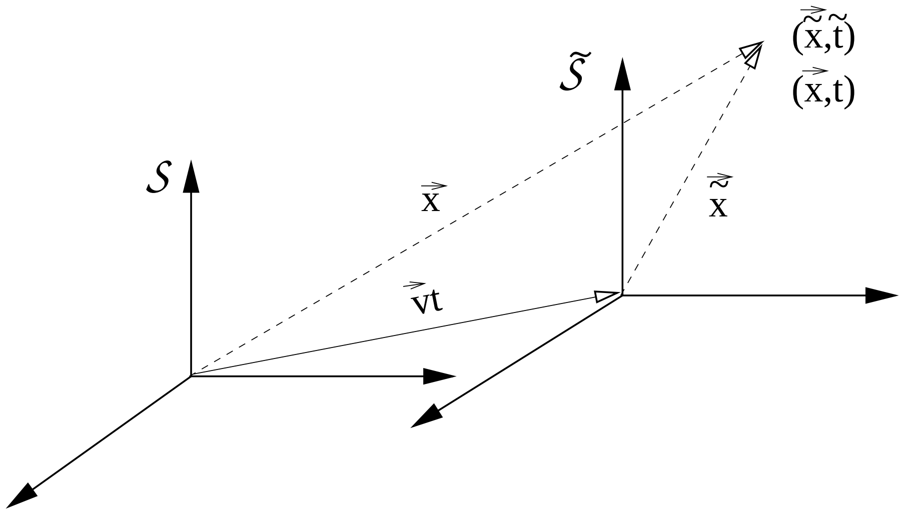
- Background (Relevance to Electrodynamics)
- Postulates of Special Relativity
- Digression: General Coordinate Transformations
- Lorentz Transformations and the Lorentz Group
- Lorentz Covariance of Maxwell's Equations
- Structure of the Lorentz Group
- Relativistic Addition of Velocities

(Lecture notes – Fawad Hassan, Oct 2008)

Special Relativity and Maxwell's Equations

1 Background

Consider two reference frames $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ such that $\widetilde{\mathcal{S}}$ moves away from $\mathcal{S}$ with constant velocity $\vec{v}$. Let two observers who are stationary in $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ respectively, measure the coordinates of an event (which occurs at some point in space and some instant in time) as $(\vec{x}, t)$ and $(\vec{\tilde{x}}, \tilde{t})$. If the velocity v is much smaller that the speed of light in vacuum c,



then, to a good degree of accuracy, the two coordinate measurements are related by

$$\vec{\tilde{x}} = \vec{x} - \vec{v}\,t\,, \qquad \tilde{t} = t$$

This is the *Galilean transformation* relating $\mathcal{S}$ and $\widetilde{\mathcal{S}}$. In particular, differentiating with respect to t one obtains the Galilean law of addition of velocities,

$$\vec{\tilde{u}} = \vec{u} - \vec{v}$$

where $\vec{\tilde{u}}$ and $\vec{u}$ are the velocities of the event point as measured in the two frames.

The concept of frames in relative motion with respect to one another naturally arises in electrodynamics and is very important. For example, an observer in the lab frame may measure magnetic and electric fields associated with a moving charge, while an observer moving along with the charge will only see an electric field. Thus compatibility with the notion of relative motion should be inbuilt in Maxwell's equations. However, it was well known, even before the advent of the special theory of relativity that Maxwell's equations were not consistent with Galilean transformations. That is to say, if we wrote the equations describing a given electromagnetic system in frame $\mathcal{S}$, then after a Galilean transformation to frame $\widetilde{\mathcal{S}}$ they would no longer look like Maxwell's equations.

In fact, using this observation, Lorentz had empirically determined a new set of coordinate transformations between $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ that preserved the form of Maxwell equations. These are the so called *Lorentz Transformations*. Poincaré had studied the mathematical properties of these transformations and had discovered their group structure, now called the Lorentz group. However, before Einstein's special theory of relativity the physical basis of these transformations were not clear.

2 Postulates of Special Relativity

Einstein's special theory of relativity rests on two postulates:

1. *The laws of physics have the same form in all* `inertial` *reference frames (An inertial frame is one on which no forces act. It is basically defined by Newton's first law of motion). This is called* `the principle of covariance`*.*
2. *The velocity of light in vacuum is a constant c which is the same in all inertial frames.*

These two very general statements have far reaching consequences. Before considering some of these in detail, let us remember that from our elementary courses on special relativity we already know that first of all, the Galilean transformation law is modified to Lorentz transformations. For simplicity let us assume that the relative motion between $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ is only in the x^1 direction ($v^1 = v, v^2 = 0, v^3 = 0$).Then the coordinates (x^i, t) and $(\tilde{x}^i, \tilde{t})$ of the same event as measured in the two frames $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ are related by

$$\tilde{t} = \frac{t - (v/c^2)x^1}{\sqrt{1 - v^2/c^2}}\,, \qquad \tilde{x}^1 = \frac{x^1 - vt}{\sqrt{1 - v^2/c^2}}\,, \qquad \tilde{x}^2 = x^2\,, \qquad \tilde{x}^3 = x^3$$

We will later give a formal and general derivation of Lorentz transformations. Here we will focus on some features of the above transformations. It is common to use the notation

$$x^0 = ct\,, \qquad \beta = \frac{v}{c}\,, \qquad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$$

in terms of which,

$$\tilde{x}^0 = \gamma\,(x^0 - \beta\,x^1)\,, \qquad \tilde{x}^1 = \gamma\,(x^1 - \beta\,x^0)\,, \qquad \tilde{x}^2 = x^2\,, \qquad \tilde{x}^3 = x^3$$

Since $0 \leq \beta \leq 1$ and $1 \leq \gamma \leq \infty$ they can be parametrized in terms of hyperbolic functions,

$$\beta = \tanh\xi\,, \qquad \gamma = \cosh\xi\,, \qquad \gamma\beta = \sinh\xi$$

where ξ is a hyperbolic "angle" known as *boost parameter* or *rapidity*. In this parameterization the transformation takes the form,

$$\begin{pmatrix} \tilde{x}^0 \\ \tilde{x}^1 \\ \tilde{x}^2 \\ \tilde{x}^3 \end{pmatrix} = \begin{pmatrix} \cosh\xi & -\sinh\xi & 0 & 0 \\ -\sinh\xi & \cosh\xi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

It is now easy to check that

$$(\tilde{x}^0)^2 - (\tilde{x}^1)^2 - (\tilde{x}^2)^2 - (\tilde{x}^3)^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$$

Comparing this with ordinary rotations that keep $(x^1)^2 + (x^2)^2 + (x^3)^2$ invariant, it is clear that a Lorentz transformation can be interpreted as a kind of rotation by an imaginary angle that involves the time direction. We will now discuss Lorentz transformations in a more formal way.

3 Digression: General Coordinate Transformations

Before discussing Lorentz transformations which is a special kind of coordinate transformations, it is instructive to review the general formalism of coordinate transformations. Let us denote the coordinates by x^μ where $\mu = 0, 1, 2, 3$ ($x^0 = ct, x^1, x^2, x^3$).

A *general coordinate transformation* (GCT) between two coordinate systems x and $\tilde{x}$ means that $\tilde{x}^\mu$ are given as general functions of x^μ,

$$\tilde{x}^\mu = \tilde{x}^\mu(x^0, x^1, x^2, x^3).$$

The functional relationship can also be inverted and one can equally write,

$$x^\mu = x^\mu(\tilde{x}^0, \tilde{x}^1, \tilde{x}^2, \tilde{x}^3).$$

Now, using ordinary multi-variable differential calculus one can easily figure out how the differentials dx^μ and derivatives $\partial/\partial x^\mu$ transform,

$$d\tilde{x}^\mu = \sum_\nu \frac{\partial \tilde{x}^\mu}{\partial x^\nu} dx^\nu \,, \qquad \frac{\partial}{\partial \tilde{x}^\mu} = \sum_\nu \frac{\partial x^\nu}{\partial \tilde{x}^\mu} \frac{\partial}{\partial x^\nu}$$

These define the basic types of transformations under GCT: Any quantity that transforms like dx^μ is called a *contravariant* vector. Hence, $V^\mu(x)$ is a contravariant vector if in the $\tilde{x}$ coordinate system it is given by some $\widetilde{V}^\mu(\tilde{x})$ such that

$$\widetilde{V}^\mu(\tilde{x}) = \sum_\nu \frac{\partial \tilde{x}^\mu}{\partial x^\nu} V^\nu(x)$$

Similarly, any quantity transforming as $\partial/\partial x^\mu$ is called a *covariant* vector. Hence, a quantity $W_\mu(x)$ is a covariant vector if

$$\widetilde{W}_\mu(\tilde{x}) = \sum_\nu \frac{\partial x^\nu}{\partial \tilde{x}^\mu} W_\nu(x)$$

In general one can define a rank $m+n$ tensor with m contravariant and n covariant indices transforming as

$$\widetilde{T}^{\tilde{\mu}_1\cdots\tilde{\mu}_m}{}_{\tilde{\nu}_1\cdots\tilde{\nu}_n} = \frac{\partial \tilde{x}^{\tilde{\mu}_1}}{\partial x^{\mu_1}} \cdots \frac{\partial \tilde{x}^{\tilde{\mu}_n}}{\partial x^{\mu_n}} \frac{\partial x^{\nu_1}}{\partial \tilde{x}^{\tilde{\nu}_1}} \cdots \frac{\partial x^{\nu_n}}{\partial \tilde{x}^{\tilde{\nu}_n}} T^{\mu_1\cdots\mu_m}{}_{\nu_1\cdots\nu_n}$$

where summations over repeated indices are implied. It is also evident that in general, x^μ itself does not transform as a vector under general coordinate transformations. Only when the transformation is linear, x^μ can be regarded as a contravariant vector.

Note that $\partial\tilde{x}^\mu/\partial x^\nu$ and $\partial x^\rho/\partial\tilde{x}^\sigma$ can be regarded as elements of matrices M and N,

$$\frac{\partial \tilde{x}^\mu}{\partial x^\nu} = M^\mu{}_\nu\,, \qquad \frac{\partial x^\rho}{\partial \tilde{x}^\sigma} = N^\rho{}_\sigma$$

In general M and N are functions of the coordinates x^μ or $\tilde{x}^\mu$. Then in matrix notation, the transformations of contravariant and covariant vectors take the form [1]

$$\begin{array}{lcl} \widetilde{V}^\mu = M^\mu_\nu V^\nu & \qquad \text{or} \qquad & \widetilde{V} = M\,V \\ \widetilde{W}_\mu = N^\nu{}_\mu W_\nu & \qquad \text{or} \qquad & \widetilde{W} = N^T\,W \end{array} \tag{1}$$

where we have used $N^\nu_\mu = (N^T)_\mu{}^\nu$. Now, since

$$\sum_\nu \frac{\partial \tilde{x}^\mu}{\partial x^\nu}\frac{\partial x^\nu}{\partial \tilde{x}^\sigma} = \frac{\partial \tilde{x}^\mu}{\partial \tilde{x}^\sigma} = \delta^\mu_\sigma\,,$$

one concludes that

$$N = M^{-1} \tag{2}$$

As a result of this, covariant and contravariant vectors have the important property that their scalar product or *contraction*, defined as $\sum_\mu V^\mu W_\mu$ is invariant under general coordinate transformations:

$$\sum_\mu \widetilde{V}^\mu \widetilde{W}_\mu = \sum_{\nu,\lambda}\left(\sum_\mu \frac{\partial \tilde{x}^\mu}{\partial x^\nu}\frac{\partial x^\lambda}{\partial \tilde{x}^\mu}\right) V^\nu W_\lambda = \sum_{\nu,\lambda} \delta^\lambda_\nu\, V^\nu W_\lambda = \sum_\nu V^\nu W_\nu$$

It is now clear that given a contravariant vector V^μ and a covariant rank 2 tensor $W_{\mu\nu}$, one can construct a covariant vector $W_\nu = V^\mu W_{\mu\nu}$. However, this way of associating a

[1] To spell out the conventions: The index labeling matrix columns appears to the right of the index labeling rows. The up or down position of the index has noting to do with the matrix structure and is solely determined by the transformation property of the index, with the "up" position for cotravariant and "down" for covariant indices. Note that with this convention, matrix transposition changes the left-right position of the indices, keeping their up-down structure unchanged: $(N^T)_\nu{}^\mu = N^\mu_\nu$

covariant vector W_μ to a contravariant vector V^μ is not unique as it depends on the choice of $W_{\mu\nu}$. One can make this correspondence precise by choosing a unique rank 2 tensor for the purpose. Conventionally, one chooses the metric tensor $g_{\mu\nu}$ for the job. This is a symmetric tensor and its inverse is denoted by $g^{\mu\nu}$, so that $g_{\mu\nu}g^{\nu\lambda} = \delta_\mu^\lambda$. Then one can write

$$V^\mu\, g_{\mu\nu} = V_\nu\,, \qquad V_\mu\, g^{\mu\nu} = V^\nu$$

where summations over repeated indices are implied. This allows us to associate to any contravariant vector V^μ a unique covariant vector V_μ and *vice versa*. In particular, it allows us to define the *scalar product* of a vector with itself, or its *length* as

$$V^\mu V_\mu = V^\mu g_{\mu\nu} V^\nu$$

This clearly is invariant under general coordinate transformations. If we arrange the components of V^μ ($\mu = 0, 1, 2, 3$) in a column vector and regard $g_{\mu\nu}$ as the components of a square matrix g, then $V_\mu = (gV)_\mu$ and in matrix notation

$$V^\mu V_\mu = V^\mu g_{\mu\nu} V^\nu = V^T g V$$

An example of $g_{\mu\nu}$ is the metric of flat space-time $\eta_{\mu\nu}$ which naturally arises in the context of special relativity (as we will see below) and is given by

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

3.1 General Linear Coordinate Transformations

These are a subset of general coordinate transformations in which $\tilde{x}^\mu$ are linear functions of x^μ,

$$\tilde{x}^\mu = A^\mu_{\ \nu}\, x^\nu\,, \qquad x^\mu = (A^{-1})^\mu_{\ \nu}\, \tilde{x}^\nu$$

where $A^\mu_{\ \nu}$ are independent of x^μ. In matrix notation,

$$\tilde{x} = A\, x\,, \qquad x = A^{-1}\, \tilde{x}$$

The feature of linear transformations that is of interest to us here is that, since A does not depend on x, the transformation treats all space-time points on equal footing.

4 Lorentz Transformations and the Lorentz Group

Let us now get back to the postulates of special relativity and see how they can be used to determine the relation between x and $\tilde{x}$.

We now regard x^μ and $\tilde{x}^\mu$ as being the coordinates of the same event E as measured by observers in the frames $\mathcal{S}$ and $\widetilde{\mathcal{S}}$, respectively. Our aim is to find the relationship between x^μ and $\tilde{x}^\mu$ in agreement with the postulates of special relativity. First we observe that by the first postulate, the transformation between $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ should not depend on the location of the space-time point because any such dependence would imply that all space-time points are not equivalent. This could then be used to identify some preferred points and frames in contradiction with the principle of relativity that treats all points and frames equally. Hence the only possibility is for x^μ and $\tilde{x}^\mu$ to be related by a linear coordinate transformation,

$$\tilde{x}^\mu = L^\mu_{\ \nu}\, x^\nu \,, \qquad \text{or} \qquad \tilde{x} = L\, x$$

Since L is a constant matrix, one has $d\tilde{x}^\mu = L^\mu_{\ \nu}\, dx^\nu$. Hence x^μ itself transforms as a contravariant vector. From now on we will be careful about the position of the index μ as it specifies the transformation properties of the object. Clearly for two frames in relative motion, L should depend on the relative velocity of the frames.

We now set out to specify the matrix $L^\mu_{\ \nu}$ using the second postulate of relativity. This can be done by considering the following set up: Assume (for simplicity) that the origins of the frames $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ coincide at some instant of time and choose this instant as the origin the time coordinate in both frames ($t = \tilde{t} = 0$). From this instant onward, the frames move apart with velocity $\vec{v}$. Consider a light signal emitted from the origin of the $\mathcal{S}$ frame ($x^1 = x^2 = x^3 = 0$) at time $x^0 = ct = 0$. The light will propagate outward on a spherical wavefront with speed c. After a time t, the distance of a point on the wavefront from the centre is $\sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2} = c\,t = x^0$. In other words,

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = 0$$

Now consider the same process from the point of view of the observer in $\widetilde{\mathcal{S}}$. According to the second postulate, the observer in $\widetilde{\mathcal{S}}$ also sees a light signal emitted from the origin ($\tilde{x}^1 = \tilde{x}^2 = \tilde{x}^3 = 0$) which now expands with speed c in a spherical wavefront in $\widetilde{\mathcal{S}}$. When the observer in $\mathcal{S}$ makes her/his measurements, the observer in $\widetilde{\mathcal{S}}$ also measures the distance from the origin to the wavefront as $\sqrt{(\tilde{x}^1)^2 + (\tilde{x}^2)^2 + (\tilde{x}^3)^2} = c\,\tilde{t} = \tilde{x}^0$, or

$$(\tilde{x}^0)^2 - (\tilde{x}^1)^2 - (\tilde{x}^2)^2 - (\tilde{x}^3)^2 = 0$$

Thus the constancy of the speed of light implies that if the interval $(x^0)^2 - \sum_{i=1}^3 (x^i)^2$ vanishes in one frame, its Lorentz transform $(\tilde{x}^0)^2 - \sum_{i=1}^3 (\tilde{x}^i)^2$ also vanishes in other frames.

The quantity $(x^0)^2 - \sum_{i=1}^3 (x^i)^2$ is the space-time interval between two events. These we have chosen as the emission of light at $x^\mu = 0$ and its measurement at $\tilde{x}^\mu \neq 0$. We can also consider other intervals not necessarily corresponding to the travel of light (for example, the motion of an electron). Then in general, $(x^0)^2 - \sum_{i=1}^3 (x^i)^2 \neq 0$. However, the fact that the vanishing of the interval in one frame leads to its vanishing in other frames implies that, in general, the two measurements of the same interval are related by

$$(x^0)^2 - \sum_{i=1}^{3}(x^i)^2 = \alpha(\vec{v})\left[(\tilde{x}^0)^2 - \sum_{i=1}^{3}(\tilde{x}^i)^2\right]$$

We can also give an exactly equivalent description of the above process by interchanging the roles of $\mathcal{S}$ and $\widetilde{\mathcal{S}}$, now regarding $\mathcal{S}$ to be moving away from $\widetilde{\mathcal{S}}$ with velocity $-\vec{v}$. This will lead to

$$(\tilde{x}^0)^2 - \sum_{i=1}^{3}(\tilde{x}^i)^2 = \alpha(-\vec{v})\left[(x^0)^2 - \sum_{i=1}^{3}(x^i)^2\right]$$

Comparing these two equations, one gets,

$$\alpha(-\vec{v}) = \alpha^{-1}(\vec{v})$$

Now, notice that the sign of $\vec{v}$ can be changed either by sending $t \to -t$ or $\vec{x} \to -\vec{x}$. However, the interval involves $(x^0)^2$ and $(x^i)^2$ and is insensitive to this change which implies that it is also insensitive to the sign of $\vec{v}$ and hence $\alpha(-\vec{v}) = \alpha(\vec{v})$. One then has $\alpha = \alpha^{-1}$ which gives $\alpha = 1$ ($\alpha = -1$ is physically unacceptable as it does not lead to the expected result for $\vec{v} = 0$).

Therefore, we conclude that the transformation $\tilde{x} = L\,x$ is constrained such that

$$(\tilde{x}^0)^2 - \sum_{i=1}^{3}(\tilde{x}^i)^2 = (x^0)^2 - \sum_{i=1}^{3}(x^i)^2$$

The invariance of the interval translates into a restriction on $L^\mu{}_\nu$ which serves as the definition of Lorentz transformations. It is convenient to express this restriction on $L^\mu{}_\nu$ in terms of a matrix $\eta_{\mu\nu}$,

$$\eta = (\eta_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \eta^{-1} = (\eta^{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

This allows us to write

$$(x^0)^2 - \sum_{i=1}^{3}(x^i)^2 = x^\mu \eta_{\mu\nu} x^\nu = x^T \eta x$$

Thus in matrix notation the Lorentz transformation is defined by

$$\tilde{x} = L\,x\,, \qquad \tilde{x}^T \eta \tilde{x} = x^T (L^T \eta L) x = x^T \eta x$$

In order for this to hold for any x, the matrix L must satisfy [2],

$$L^T \eta L = \eta$$

The set of all matrices L satisfying this condition form the *Lorentz group* of transformations, $O(1,3)$. To summarize, Lorentz transformations consist of all those linear transformations L that keep η invariant.

[2]In components, $(L^T)_\mu{}^{\mu'} \eta_{\mu'\nu'} L^{\nu'}{}_\nu = \eta_{\mu\nu}$ or $(L)^{\mu'}{}_\mu L^{\nu'}{}_\nu \eta_{\mu'\nu'} = \eta_{\mu\nu}$ (since $(L^T)_\mu{}^{\mu'} = (L)^{\mu'}{}_\mu$)

4.1 Some properties of L

Some properties of L can be directly obtained from its defining equation. For example, multiplying both sides by η^{-1} one has $\eta^{-1}L^T\eta L = \mathbf{1}$ so that

$$L^{-1} = \eta^{-1}L^T\eta$$

Therefore it is very easy to compute the inverse of a Lorentz transformation. Furthermore, one can see that $(\det L)^2 = 1$ or

$$\det L = \pm 1$$

The significance of this will be discussed later.

4.2 Covariant and Contravariant Lorentz vectors

Since L is a constant matrix, from the transformation of x^μ is follows that,

$$d\tilde{x}^\mu = L^\mu_{\ \nu}\, dx^\nu$$

It is also easy to work out the transformation of $\partial_\mu = \partial/\partial x^\mu$ using $x = L^{-1}\tilde{x}$,

$$\tilde{\partial}_\mu \equiv \frac{\partial}{\partial\tilde{x}^\mu} = \frac{\partial x^\nu}{\partial\tilde{x}^\mu}\frac{\partial}{\partial x^\nu} = (L^{-1})^\nu_{\ \mu}\frac{\partial}{\partial x^\nu} \equiv (L^{-1})^\nu_{\ \mu}\,\partial_\nu$$

or, in matrix notation,

$$\tilde{\partial} = (L^{-1})^T\partial$$

where ∂ is a column vector constructed from the components of ∂_μ. Thus, the 4-vectors dx^μ and ∂_μ transform differently under Lorentz transformations, one with L and the other with $(L^{-1})^T$. Vectors transforming like dx^μ are called contravariant vectors and those transforming like ∂_μ are called covariant vectors. Thus under a Lorentz transformation, general contravariant vectors V^μ and covariant vectors W_μ transform as

$$\widetilde{V}^\mu = L^\mu_{\ \nu}\, V^\nu\,, \qquad \widetilde{W}_\mu = (L^{-1T})_\mu^{\ \nu}\, W_\nu = (\eta L\eta^{-1})_\mu^{\ \nu}\, W_\nu$$

or in matrix notation,

$$\widetilde{V} = L\,V\,, \qquad \widetilde{W} = (L^{-1})^T\,W = (\eta L\eta^{-1})\,W$$

so that the contraction $V^TW = V^\mu W_\mu$ is invariant (this is the generalization of the scalar or "dot" product to the 4 dimensional space-time). These equations for the Lorentz transformation are consistent with equations (1) and (2) for general coordinate transformations. Note that the notation ∂_μ is not only compact, but it also makes the covariant nature of the partial derivatives manifest.

So far we have seen that the space-time coordinate x^μ (and its differential dx^μ) transform as contravariant 4-vectors under Lorentz transformations. One can easily associate a covariant vector to x^μ as,

$$x_\mu = \eta_{\mu\nu}\, x^\nu = (\eta\, x)_\mu$$

(we say that "the index on x has been lowered by using η"). That this really transforms as a covariant vector can be easily verified using the transformation of x^μ,

$$(\eta\,\tilde{x}) = (\eta\,L\,x) = (\eta\,L\,\eta^{-1})(\eta x) = (L^{-1})^T\,(\eta x)$$

which is consistent with equations (1) and (2). Note that $x^\mu \eta_{\mu\nu} x^\nu = x^\mu x_\mu$ which is invariant under the transformations.

The same construction applies to any contravariant vector V^μ. Similarly, a covariant vector W_μ can be converted to a contravariant one. Explicitly

$$V_\mu = \eta_{\mu\nu} V^\nu\,, \qquad W^\mu = \eta^{\mu\nu} W_\nu$$

Thus the matrix $\eta_{\mu\nu}$ and its inverse $\eta^{\mu\nu}$ can be used to convert covariant and contravariant vectors of the Lorentz group into each other. Note that because of the simple structure of the matrix η (also called the *Minkowski metric*), the components of V^μ and those of the associated covariant vector $V_\mu = \eta_{\mu\nu}\,V^\nu$ are related in a simple way,

$$\begin{pmatrix} V_0 \\ V_1 \\ V_2 \\ V_3 \end{pmatrix} = \begin{pmatrix} V^0 \\ -V^1 \\ -V^2 \\ -V^3 \end{pmatrix}$$

In words, raising or lowering of an index keeps the time-component of the 4-vector unchanged, $V^0 = V_0$, while reversing the sign of the spatial components, $V^i = -V_i$ $(i = 1, 2, 3)$.

5 Lorentz Covariance of Maxwell's Equations

5.1 The principle of covariance

This principle, which is the first postulate of special relativity, states that the laws of physics should have the same form in all intertial reference frames. Since physical quantities in different inertial frames are related by Lorentz transformations, this is equivalent to saying that the laws of physics should be "covariant" (that is, should retain their form) under Lorentz transformations [3]

The covariance of laws of physics (for example, that of Maxwell's equations) is sometimes not manifest. However, from our experiance with the space-time 4-vector x^μ and partial-derivative 4-vector ∂_μ it is clear that if we could express all physical quantities and laws of physics in terms of Lorentz 4-vectors and 4-tensors, then their compatibility

[3] Note that the term "covariant" used in this context has a different meaning than in the phrase "covariant vector". Thus a covariant law (or equation) is one that has the same form in all inertail frames while a covariant vector is one that transforms in a specific way

(covariance) with Lorentz transformations will become manifest. Besides the space-time coordinates, in electrodynamics one deals with the fields $\phi, \vec{A}, \vec{E}, \vec{B}$ and the charge and current densities $\rho, \vec{J}$. We will now see how these can be expressed in terms of Lorentz 4-vectors and 4-tensors, thereby making their behaviour under Lorentz transformations manifest. We can then rewrite the Maxwell's equations in manifestly Lorentz covariant form.

5.2 Charge and Current densities $\rho\,,\ \vec{J}$

The conservation of charge which is expressed in terms of the continuity equation,

$$\frac{\partial \rho}{\partial t} + \frac{\partial J^i}{\partial x^i} = 0$$

is a basic concept in physics and should hold true in all inertial frames. This means that if a different observer (say in frame $\widetilde{S}$) measures charge and current densities $\widetilde{\rho}$ and $\widetilde{J}^i$, then he/she should find that

$$\frac{\partial \widetilde{\rho}}{\partial \tilde{t}} + \frac{\partial \widetilde{J}^i}{\partial \tilde{x}^i} = 0$$

The above two equations should be related by Lorentz transformations and should imply each other.

We already know that $\frac{1}{c}\frac{\partial}{\partial t}$ and $\frac{\partial}{\partial x^i}$ combine into a covariant 4-vector $\frac{\partial}{\partial x^\mu} = \partial_\mu$. Let us now combine $c\rho$ and J^i $(i = 1, 2, 3)$ into a 4-component object J^μ $(\mu = 0, 1, 2, 3)$ with $J^0 = c\,\rho$. The continuity equation now takes the form

$$\partial_\mu\, J^\mu = 0$$

(where a summation over μ is implied). Since ∂_μ transforms as a covariant 4-vector, one concludes that J^μ has to transforms as a contravariant 4-vector so that the product remains invariant under Lorentz transformations, $\partial_\mu\, J^\mu = \tilde{\partial}_\mu\, \widetilde{J}^\mu$.

Under ordinary spatial rotations $SO(3)$, the charge density ρ transforms as a scalar and the current density J^i as a vector. However, under Lorentz transformations $SO(1,3)$, $c\rho$ behaves as the time-component of a 4-vector and can therefore mix with its space-components,

$$\widetilde{J}^\mu(\tilde{x}) = L^\mu{}_\nu\, J^\nu(x) = L^\mu{}_\nu\, J^\nu(L^{-1}\tilde{x})$$

The last step above makes it manifest that the final answer should be written in terms of variables $\tilde{x}^\mu$ which are the natural variables for the observer in $\widetilde{\mathcal{S}}$. The mixing between ρ and J^i under Lorentz transformations is obvious intuitively since a charge density will apear as a current to a moving observer.

5.3 Electric and Magnetic potentials ϕ, $\vec{A}$

In our study of electrostatics and magnetostatics we saw that ρ and $\vec{J}/c$ act as sources for the electric potential ϕ and the magnetic vector potential $\vec{A}$,

$$\phi(x) = \int d^3x' \frac{\rho(x')}{|x - x'|}\,, \qquad A^i(x) = \frac{1}{c}\int d^3x' \frac{J^i(x')}{|x - x'|}\,,$$

Hence, in the same way that $c\rho$ and J^i combine into a 4-vector J^μ, one can infer that ϕ and A^i should also correspond to components of a 4-vector, say, A^μ with components

$$A^{\mu=0} = \phi\,, \qquad A^{\mu=i} = A^i$$

Taking the transformations of d^3x' and $|x - x'|$ into account does not change the situation. Hence, in matrix notation,

$$\tilde{A}(\tilde{x}) = LA(L^{-1}\tilde{x})$$

Note that in this notation, gauge transformations of Φ and $\vec{A}$ take a simple form,

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \Lambda$$

where Λ is a scalar function of space and time. The Lorentz gauge condition on ϕ and A^i takes the form

$$\partial_\mu A^\mu = 0$$

We also know that in this gauge the potentials satisfy the equations

$$\nabla^2 \vec{A} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\vec{A} = -\frac{4\pi}{c}\vec{J}\,, \qquad \nabla^2\phi - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\phi = -4\pi\rho$$

The differential operator appearing in the above equations is Lorentz invariant since,

$$\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} = \partial_i\partial_i - \partial_0\partial_0 = -\eta^{\mu\nu}\partial_\mu\partial_\nu = -\partial^\mu\partial_\mu \equiv -\Box$$

The equations then have a manifestly Lorentz covariant form,

$$\Box A^\mu = \frac{4\pi}{c}\,J^\mu\,, \qquad \partial_\mu A^\mu = 0$$

5.4 Electric and Magnetic Fields $\vec{E}$ and $\vec{B}$

To write the electric and magnetic fields $\vec{E}$ and $\vec{B}$ in a form that makes their transformation under the Lorentz group explicit, we use their definitions in tersm of ϕ and A^i. For the electric field we have $\vec{E} = -\vec{\nabla}\phi - (1/c)\partial\vec{A}/\partial t$ which becomes,

$$E^i = -\partial_i\phi - \frac{1}{c}\frac{\partial A^i}{\partial t} = -\partial_i A^0 - \partial_0 A^i$$

Here we have used the fact that ϕ and $\vec{A}$ are components of a 4-vector which is naturally contravariant (upper index) while the partial derivatives are naturally covariant (lower index). So far it is not fully clear that the electric field should be written with an upper index. But now, noting that for any 4-vector V^μ, $V^0 = V_0$ while $V^i = -V_i$, this can be rewritten as

$$E^i = \partial^i A^0 - \partial^0 A^i \tag{3}$$

Similarly, for the magnetic field $\vec{B} = \vec{\nabla} \times \vec{A}$,

$$B^i = \sum_{j,k=1}^{3} \epsilon^{ij}{}_k \, \partial_j \, A^k = \sum_{j,k=1}^{3} \epsilon^{ijk} \, \partial_j \, A_k$$

where ϵ^{ijk} is the completely antisymmetric tensor in 3 dimentions. In our conventions, $\epsilon_{123} = +1$. Clearly,

$$\epsilon^{ijk} = -\epsilon^{ij}{}_k = -\epsilon_{ijk}$$

Using antisymmetry of the ϵ-tensor (and supressing the sum),

$$B^i = \frac{1}{2}\epsilon^i{}_{jk} \left(\partial^j \, A^k - \partial^k \, A^j\right) = \frac{1}{2}\epsilon^i{}_{jk} \, F^{jk} = \frac{1}{2}\epsilon^{ijk} \, F_{jk} \tag{4}$$

Based on equations (3) and (4), it is natural to define a rank-2 antisymmetric 4-tensor,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,, \qquad \text{or} \qquad F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$$

In terms of this the electric field is given by

$$E^i = F^{i0} = -F^{0i}$$

As for the magnetic field, one has

$$B^i = \frac{1}{2} \, \epsilon^{ijk} \, F_{jk}$$

which can be inverted as

$$\begin{aligned}
\sum_{k=1}^{3} \epsilon_{mnk} B^k &= \frac{1}{2} \sum_{i,j,k=1}^{3} \epsilon_{mnk} \epsilon^k{}_{ij} \, F^{ij} = -\frac{1}{2} \sum_{i,j,k=1}^{3} \epsilon_{mnk} \, \epsilon_{ijk} \, F^{ij} \\
&= -\frac{1}{2} \sum_{i,j=1}^{3} \left(\delta_{mi}\delta_{nj} - \delta_{mj}\delta_{ni}\right) F^{ij} = -\frac{1}{2} \left(F_{mn} - F_{nm}\right) = -F_{mn}
\end{aligned}$$

In short,

$$F^{ij} = -\epsilon^{ij}{}_k B^k$$

This shows that all elements of the matrix $F^{\mu\nu}$ are determined by the components of the electric and magnetic fields (note that the diagonal elements of the matrix are zero because of its antisymmetry). Explicitly, one has,

$$F = \begin{pmatrix} F^{00} & F^{01} & F^{02} & F^{03} \\ F^{10} & F^{11} & F^{12} & F^{13} \\ F^{20} & F^{21} & F^{22} & F^{23} \\ F^{30} & F^{31} & F^{32} & F^{33} \end{pmatrix} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}$$

$F^{\mu\nu}$ is called the electromagnetic fieldstrength tensor and provides a Lorentz covariant description of electric and magnetic fields.

It is clear that from the point of view of the Lorentz group E^i and B^i are components of an antisymmetric rank-2 Lorentz tensor in 3+1 dimensions. Their Lorentz transformation properties can therefore be easily extracted from the transformation of $F^{\mu\nu}$. Hence under a Lorentz transformation $\widetilde{x}^\mu = L^\mu{}_\nu\, x^\nu$, we have

$$\widetilde{F}^{\widetilde{\mu}\widetilde{\nu}}(\widetilde{x}) = L^{\widetilde{\mu}}{}_\mu\, L^{\widetilde{\nu}}{}_\nu\, F^{\mu\nu}(x)$$

It is very convenient to rewrite this in matrix notation. Since $(L^T)_{\widetilde{\nu}}{}^\nu = L^\nu{}_{\widetilde{\nu}}$, one has

$$\widetilde{F}^{\widetilde{\mu}\widetilde{\nu}}(\widetilde{x}) = L^{\widetilde{\mu}}{}_\mu\; F^{\mu\nu}(x)\,(L^T)_{\widetilde{\nu}}{}^\nu$$

This makes the matrix structure of the transformation explicit so that in terms of matrices, one has

$$\widetilde{F}(\widetilde{x}) = L\;F(x)\,L^T$$

After the matrix multiplication, all x^μ should be re-expressed in terms of $\tilde{x}$ through $x^\mu = (L^{-1})^\mu{}_\nu \tilde{x}^\nu$

Thus, given L, one can always find how $\vec{E}$ and $\vec{B}$ transform. For example, consider frames S and $\widetilde{S}$ in relative motion with velocity $\vec{v}$ only in the x^1 direction. Then as we have mentioned earlier L is given by

$$L = \begin{pmatrix} C & -S & 0 & 0 \\ -S & C & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where $C = \cosh\xi$ and $S = \sinh\xi$, ξ being the *rapidity* or *boost parameter*. Plugging this into the transformation equation for F and carrying out the matrix multiplications, one can read off the transformation of E^i and B^i as

$$\begin{aligned} \widetilde{E}^1 &= E^1 & \widetilde{B}^1 &= B^1 \\ \widetilde{E}^2 &= E^2\cosh\xi - B^3\sinh\xi & \widetilde{B}^2 &= B^2\cosh\xi + E^3\sinh\xi \\ \widetilde{E}^3 &= E^3\cosh\xi + B^2\sinh\xi & \widetilde{B}^3 &= B^3\cosh\xi - E^2\sinh\xi \end{aligned}$$

In terms of $\beta = v/c = \tanh\xi$ and $\gamma = 1/\sqrt{1-\beta^2}$ these are,

$$\begin{aligned} \widetilde{E}^1 &= E^1 & \widetilde{B}^1 &= B^1 \\ \widetilde{E}^2 &= \gamma(E^2 - \beta B^3) & \widetilde{B}^2 &= \gamma(B^2 + \beta E^3) \\ \widetilde{E}^3 &= \gamma(E^3 + \beta B^2) & \widetilde{B}^3 &= \gamma(B^3 - \beta E^2) \end{aligned}$$

5.5 Covariant form of Maxwell's equations

We are now finally in a position to write the Maxwell equations in a manifestly covariant form. Let us first consider the two equations with source terms,

$$(i)\;\vec{\nabla}\cdot\vec{E} = 4\pi\rho\,, \qquad (ii)\;\vec{\nabla}\times\vec{B} - \frac{1}{c}\frac{\partial}{\partial t}\vec{E} = \frac{4\pi}{c}\vec{J}$$

Note that (i) is $\partial_i F^{i0} = (4\pi/c) J^0$. Since $F^{00} = 0$ this can also be written as,

$$\partial_\mu F^{\mu 0} = \frac{4\pi}{c}\, J^0$$

As for (ii) note that,

$$(\vec{\nabla} \times \vec{B})^i = \partial_j (\epsilon^{ij}{}_k\, B^k) = \partial_j\, F^{ji}\,, \qquad \text{and} \qquad -\frac{1}{c}\frac{\partial}{\partial t} E^i = \partial_0\, F^{0i}$$

so it can also be written as

$$\partial_\mu\, F^{\mu i} = \frac{4\pi}{c}\, J^i$$

It is now clear that (i) and (ii) can be combined into a single equation

$$\partial_\mu\, F^{\mu\nu} = \frac{4\pi}{c}\, J^\nu \tag{5}$$

which is manifestly Lorentz covariant.

Let us now look at the remaining two equations,

$$(iii)\ \vec{\nabla} \cdot \vec{B} = 0\,, \qquad\qquad (iv)\ \vec{\nabla} \times \vec{E} + \frac{1}{c}\frac{\partial \vec{B}}{\partial t} = 0$$

To write these in covariant form, it is convenient to first define a tensor $*F$ (which is said to be *dual* to F) as

$$(*F)_{\mu\nu} = \frac{1}{2} \sum_{\rho,\sigma=0}^{3} \epsilon_{\mu\nu\rho\sigma}\, F^{\rho\sigma}$$

Here, $\epsilon_{\mu\nu\rho\sigma}$ is the completely antisymmetric tensor in 4-dimentions. In our conventions,

$$\epsilon^{0123} = 1\,, \qquad \Rightarrow \qquad \epsilon_{0123} = -1$$

It is related to ϵ_{ijk} in 3-dimensions as

$$\epsilon_{0ijk} = -\epsilon_{ijk}$$

It is now easy to verify that Maxwell equations (iii) and (iv) can be combined into a Lorentz covariant equation,

$$\partial_\mu\, (*F)^{\mu\nu} = 0 \tag{6}$$

We show that this does contain (iii) and (iv): First consider the $\nu = 0$ component of the equation. Since $(*F)^{00} = 0$, this is $\partial_i\, (*F)^{i0} = 0$. But,

$$(*F)^{i0} = \frac{1}{2}\epsilon^{i0\rho\sigma}\, F_{\rho\sigma} = -\frac{1}{2}\epsilon^{0i\rho\sigma}\, F_{\rho\sigma} = -\frac{1}{2}\epsilon^{0ijk}\, F_{jk} = \frac{1}{2}\epsilon^{ijk}\, F_{jk} = B^i$$

Hence, $\partial_i\, (*F)^{i0} = \partial_i B^i = \vec{\nabla} \cdot \vec{B} = 0$.

Let us now consider the $\nu = j$ components of the equation,

$$\partial_\mu\, (*F)^{\mu j} = \partial_i\, (*F)^{ij} + \partial_0\, (*F)^{0j} = 0$$

As we saw above, $(*F)^{0j} = -B^j$ and also,

$$(*F)^{ij} = \frac{1}{2}\epsilon^{ij\rho\sigma} F_{\rho\sigma} = \frac{1}{2}(\epsilon^{ijk0} F_{k0} + \epsilon^{ij0k} F_{0k}) = \epsilon^{ijk0} F_{k0} = \epsilon^{ijk} F_{k0} = \epsilon^{ij}{}_k E^k$$

so that $\partial_i(*F)^{ij} = -\epsilon^{ji}{}_k \partial_i E^k = -(\vec{\nabla} \times \vec{E})^j$. Hence, $\partial_\mu (*F)^{\mu j} = 0$ implies equation (iv). Equation (6) can also be written as

$$\partial_\mu F_{\nu\rho} + \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} = 0 \tag{7}$$

6 Structure of the Lorentz Group

In the following we discuss the structure of the Lorentz group and the determination of the Lorentz transformation matrix L in terms of the relative velocity of the frames.

To summarize, any vector V can be written either in a covariant form V_μ or a contravariant form V^μ, and $V^\mu = \eta^{\mu\nu} V_\nu$, $V_\mu = \eta_{\mu\nu} V^\nu$. In terms of the components, the relation is $V^0 = V_0$ and $V^i = -V_i$ for $i = 1, 2, 3$. The scalar product $V^\mu W_\mu$ is invariant under Lorentz transformations. In components, this product has the form,

$$V^\mu W_\mu = V^0 W_0 + V^i W_i = V^0 W^0 - V^i W^i$$

Let us now look at the structure of the Lorentz group. Spelling out explicitly the space and time components of $L^\mu_{\ \nu}$ one has the matrix

$$L = \begin{pmatrix} L^0_{\ 0} & L^0_{\ j} \\ L^i_{\ 0} & L^i_{\ j} \end{pmatrix} \tag{8}$$

Let us first consider the case $L^0_{\ j} = L^i_{\ 0} = 0$, $L^0_{\ 0} = 1$. Then,

$$\widetilde{x}^\mu = L^\mu_{\ \nu} x^\nu \quad \Rightarrow \quad \widetilde{x}^0 = x^0\,, \quad \widetilde{x}^i = L^i_{\ j} x^j$$

and $L^i_{\ j}$ keeps the length of the vector x unchanged. But this is how we had defined the elements of the rotation group $O(3)$. So, in this special case, $L^i_{\ j} = O^i_{\ j}$. In other words, the group of spatial rotations (including parity) is a subgroup of the Lorentz group; $O(3) \subset O(1,3)$. However, when $L^0_{\ j}$ and $L^i_{\ 0}$ are non-zero, then $L^i_{\ j} \neq O^i_{\ j}$, although even in this case the $L^i_{\ j}$ components correspond to rotations accompanied by going to a moving frame. In the next section we will show that $L^0_{\ j}$ and $L^i_{\ 0}$ are related to the relative velocity between two frames.

There is a further classification of the transformations implemented by L. We have seen that $\det L = \pm 1$. The Lorentz group thus has 4 disconnected components which are identified by the signs of $det\, L$ and $L^0_{\ 0}$: (i) $\det L = +1$, $L^0_{\ 0} > 0$ corresponds to *proper Lorentz transformations*. The other components are (ii) $det\, L = +1$, $L^0_{\ 0} < 0$ (iii) $det\, L = -1$, $L^0_{\ 0} > 0$ and (iv) $det\, L = -1$, $L^0_{\ 0} < 0$.

As a simple example, consider a Lorentz transformation that affects only x^0 and x^1, keeping x^2 and x^3 unchanged. In this case L has the general form,

$$L = \begin{pmatrix} L^0_{\ 0} & L^0_{\ 1} & 0 & 0 \\ L^1_{\ 0} & L^1_{\ 1} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{9}$$

The condition $L^T \eta L = \eta$ gives,

$$\begin{pmatrix} (L^0_{\ 0})^2 - (L^1_{\ 0})^2 & L^0_{\ 0} L^0_{\ 1} - L^1_{\ 0} L^1_{\ 1} \\ L^0_{\ 0} L^1_{\ 0} - L^1_{\ 0} L^1_{\ 1} & (L^0_{\ 1})^2 - (L^1_{\ 1})^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{10}$$

The off-diagonal components of this relation lead to $L^0_{\ 1} = L^1_{\ 0}$ and $L^0_{\ 0} = L^1_{\ 1}$. Then either of the diagonal components imposes the constraint $(L^0_{\ 0})^2 - (L^1_{\ 0})^2 = 1$. If we restrict ourselves to proper Lorentz transformations, then $L^0_{\ 0} > 0$ and the constraint has the solutions,

$$L^0_{\ 0} = \cosh \xi \,, \qquad L^1_{\ 0} = \pm \sinh \xi$$

This form of L has already appeared in our discussion. The sign of $\sinh \xi$ depends on whether one identifies the relative velocity as $v = \tanh \xi$ or as $v = -\tanh \xi$. This will be discussed in more generality bellow. Note that if we do not restrict to $L^0_{\ 0} > 0$, then we could have $L^0_{\ 0} = \pm \cosh \xi$.

6.1 Identification of Relative Velocity of the Frames

We have been able to define Lorentz transformations L without making any reference to the relative velocity between the two frames $\mathcal{S}$ and $\widetilde{\mathcal{S}}$. However, from hindsight we know that $L^\mu_{\ \nu}$ should somehow depend on this velocity. This dependence can be determined by noting that $\vec{v}$ is the velocity of the origin of $\widetilde{\mathcal{S}}$ (defined by $\widetilde{x}^i = 0$) in the frame $\mathcal{S}$. Alternatively, the origin on $\mathcal{S}$ (defined by $x^i = 0$) moves in $\widetilde{\mathcal{S}}$ with velocity $-\vec{v}$. We use these statements to relate $L^\mu_{\ \nu}$ to $\vec{v}$.

A point with coordinates x^μ in $\mathcal{S}$ will have coordinates $\widetilde{x}^\mu = L^\mu_{\ \nu} x^\nu$ in $\widetilde{\mathcal{S}}$. In particular, the spatial origin of $\mathcal{S}$ has coordinates $x^\mu_S = (x^0, 0, 0, 0)$ as measured by observers in $\mathcal{S}$ (The subscript S signifies that the coordinates refer to the origin of frame $\mathcal{S}$). The coordinates of this point as measured by observers in $\widetilde{\mathcal{S}}$ are

$$\widetilde{x}^\mu_S = L^\mu_{\ \nu}\, x^\nu_S = L^\mu_{\ 0}\, x^0 \tag{11}$$

In components, this becomes

$$\widetilde{x}^0_S = L^0_{\ 0}\, x^0 \,, \qquad \widetilde{x}^i_S = L^i_{\ 0}\, x^0 \tag{12}$$

Note that a subscript S is not needed on the time variable x^0 since all points in $\mathcal{S}$, including the origin, have the same time <u>in $\mathcal{S}$ clocks</u>. However, clocks in $\widetilde{\mathcal{S}}$ register different times corresponding to different points in $\mathcal{S}$ (*i.e.,* simultaneous events in $\mathcal{S}$ are not simultaneous in $\widetilde{\mathcal{S}}$) and hence the subscript on $\widetilde{x}^0$. Now the velocity of the origin of $\mathcal{S}$ in $\widetilde{\mathcal{S}}$ is given, in our conventions, by

$$-v^i = \frac{\partial \widetilde{x}^i_S}{\partial \widetilde{t}_S} = c\frac{\partial \widetilde{x}^i_S}{\partial \widetilde{x}^0_S} = c\frac{\partial \widetilde{x}^i_S}{\partial x^0}\frac{\partial x^0}{\partial \widetilde{x}^0_S} \tag{13}$$

Using (12) this gives

$$-\frac{v^i}{c} = \frac{L^i{}_0}{L^0_{\ 0}} \tag{14}$$

On the other hand,

$$(L^T\eta L)_{00} = (L^T)_0{}^\mu \eta_{\mu\nu} L^\nu{}_0 = L^\mu{}_0 L^\nu{}_0 \eta_{\mu\nu} = (L^0{}_0)^2 - \sum_{i=1}^3 (L^i{}_0)^2$$

$$= \quad (L^0{}_0)^2 \left[1 - \sum_{i=1}^3 \left(\frac{L^i{}_0}{L^0{}_0}\right)^2\right] = (L^0{}_0)^2 \left(1 - \frac{v^2}{c^2}\right) = 1 \tag{15}$$

where we have used (14) and $v^2 = \sum_{i=1}^3 v^i v^i$. Finally we get

$$L^0{}_0 = \frac{1}{\sqrt{1 - v^2/c^2}}\,, \qquad L^i{}_0 = \frac{-v^i/c}{\sqrt{1 - v^2/c^2}} \tag{16}$$

In terms of the parameters $\vec{\beta} = \vec{v}/c$ and $\gamma = (1-\beta^2)^{-1/2}$, these become

$$L^0{}_0 = \gamma\,, \qquad L^i{}_0 = -\gamma\,\beta^i \tag{17}$$

We can also determine $L^0_{\ i}$ in terms of $\vec{v}$ by interchanging the roles of $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ in the previous argument. Now we focus on the origin of $\widetilde{\mathcal{S}}$ which is moving with velocity $\vec{v}$ in the frame $\mathcal{S}$. A general point $\widetilde{x}^\mu$ in $\widetilde{\mathcal{S}}$ has $\mathcal{S}$ frame coordinates given by

$$x = L^{-1}\,\widetilde{x} \qquad \text{or} \qquad x^\mu = (L^{-1})^\mu{}_\nu\,\widetilde{x}^\nu \tag{18}$$

where, $L^{-1} = \eta^{-1} L^T \eta$ or, in components,

$$(L^{-1})^\mu{}_\nu = \eta^{\mu\rho}(L^T)_\rho{}^\sigma \eta_{\sigma\nu} = \eta^{\mu\rho} L^\sigma{}_\rho \eta_{\sigma\nu}$$

In particular, this gives,

$$(L^{-1})^0{}_0 = L^0{}_0 = \gamma\,, \qquad (L^{-1})^i{}_0 = -L^0{}_i \tag{19}$$

Now the origin of $\widetilde{\mathcal{S}}$ is given by $\widetilde{x}^\mu_{\widetilde{\mathcal{S}}} = (\widetilde{x}^0, \widetilde{x}^i_{\widetilde{\mathcal{S}}} = 0)$. Hence the coordinates of the origin of $\widetilde{\mathcal{S}}$ in $\mathcal{S}$ become,

$$x^\mu_{\widetilde{\mathcal{S}}} = (L^{-1})^\mu{}_\nu\,\widetilde{x}^\nu_{\widetilde{\mathcal{S}}} = (L^{-1})^\mu{}_0\,\widetilde{x}^0 \tag{20}$$

In our conventions, $\vec{v}$ denotes the velocity of the origin of $\widetilde{\mathcal{S}}$ in $\mathcal{S}$. Hence we have

$$\frac{v^i}{c} = \frac{\partial x^i_{\widetilde{\mathcal{S}}}}{\partial x^0_{\widetilde{\mathcal{S}}}} = \frac{(L^{-1})^i{}_0}{(L^{-1})^0{}_0} = -\frac{L^0{}_i}{L^0{}_0} \tag{21}$$

Using the fact that $v^i = \eta^{ij} v_j = -v_i$, one obtains $L^0{}_i$ as

$$L^0{}_i = \gamma \frac{v_i}{c} = \gamma\, \beta_i \tag{22}$$

In the following we determine $L^i{}_j$ for the case that the coordinate axes in $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ are parallel (the more general case can be obtained from this by applying rotations to the final expression, as will be discussed later). The condition $L^T \eta L = \eta$ constrains the elements $L^i{}_j$ in terms of $L^0{}_0$, $L^i{}_0$ and $L^0{}_i$ as

$$(L^T \eta L)_{0i} = 0 \quad \Rightarrow \quad L^0{}_0 L^0{}_i - \sum_{k=1}^{3} L^k{}_0 L^k{}_i = 0$$

$$(L^T \eta L)_{ij} = -\delta_{ij} \quad \Rightarrow \quad L^0{}_i L^0{}_j - \sum_{k=1}^{3} L^k{}_i L^k{}_j = -\delta_{ij}$$

Using the expressions for $L^0{}_0$, $L^i{}_0$ and $L^0{}_i$, the constraints on $L^i{}_j$ become

$$\sum_{k=1}^{3} \beta_k \left(L^k{}_i - \gamma \delta^k_i \right) = 0 \tag{23}$$

$$\sum_{k=1}^{3} L^k{}_i L^k{}_j = \delta_{ij} + \gamma^2 \beta_i \beta_j \tag{24}$$

In these formulae, δ_{ij} and δ^i_j are equal to $+1$ for $i = j$ and equal to 0 otherwise.

As a check, note that when $\vec{v} = 0$, then $\sum_{k=1}^{3} L^k{}_i L^k{}_j = \sum_{k=1}^{3} (L^T)_i{}^k L^k{}_j = \delta_{ij}$ which means that $L^i{}_j$ are elements of the rotation group in 3-dimensions [4]. Below we will write down a solution for $L^i{}_j$ for the case when the coordinate axes in $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ are parallel to each other. This solution must satisfy the requirement that in the limit $\vec{v} = 0$, $L^i{}_j = \delta_i j$.

The constraint (23) implies that for each value of i, the vector $\sum_{k=1}^{3} \left(L^k{}_i - \gamma \delta^k_i \right) \hat{\mathbf{1}}_k$ is perpendicular to $\vec{\beta} = \sum_k \beta^k \hat{\mathbf{1}}_k$, where $\hat{\mathbf{1}}_k$ is a unit vector in the k^{th} direction. A basic such object is $\sum_k (\delta^k_i + \beta^k \beta_i / \beta^2) \hat{\mathbf{1}}_k$ since

$$\sum_k \beta_k (\delta^k_i + \beta^k \beta_i / \beta^2) = \beta_i + (-\beta^2) \beta_i / \beta^2 = 0$$

Hence, as a general solution of (23) we can write

$$L^k{}_i - \gamma \delta^k{}_i = A(\vec{v})\, (\delta^k{}_i + \frac{\beta^k \beta_i}{\beta^2}) \tag{25}$$

[4] Remember that 3-dimensional rotations are implemented by a 3×3 matrix R satisfying $R^T R = \mathbf{1}$

where is $A(\vec{v})$ is a yet to be determined function of velocity. From here we evaluate $L^i{}_j$ and substitute in (24). This gives $A(\vec{v}) = -\gamma \pm 1$. Finally, the requirement that for $\vec{v} = 0$, $L^i{}_j = \delta^i_j$ fixes $A(\vec{v}) = 1 - \gamma$ and hence,

$$L^i{}_j = \delta^i{}_j + \frac{1-\gamma}{\beta^2}\,\beta^i\,\beta_j \tag{26}$$

Let us now summarise our results: When the coordinate axes in frames $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ are parallel to each other, the Lorentz transformation between the two frames is given by the matrix

$$L = \begin{pmatrix} L^0_{\ 0} & L^0_{\ j} \\ L^i_{\ 0} & L^i_{\ j} \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta_j \\ -\gamma\beta^i & \delta^i{}_j + \dfrac{1-\gamma}{\beta^2}\,\beta^i\,\beta_j \end{pmatrix} \tag{27}$$

where,

$$\beta^i = \frac{v^i}{c}\,, \quad \beta^i = -\beta_i\,, \quad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$$

The coordinates transform according to $\widetilde{x}^\mu = L^\mu{}_\nu x^\nu$. Explicitly,

$$\widetilde{x}^0 = \gamma\left(x^0 + \beta_j\, x^j\right)\,, \qquad \widetilde{x}^i = \gamma\beta^i x^0 + \frac{1-\gamma}{\beta^2}(\beta_j x^j)\,\beta^i \tag{28}$$

Alternatively, using $\sum_{j=1}^3 \beta_j x^j = -\sum_{j=1}^3 \beta^j x^j = -\vec{\beta}\cdot\vec{x}$ (and $x^0 = ct$), this can be rewritten in vector notation,

$$\widetilde{t} = \gamma\left(t - \frac{1}{c}\,\vec{\beta}\cdot\vec{x}\right)\,, \qquad \vec{\widetilde{x}} = \vec{x} - \left(c\gamma t + \frac{1-\gamma}{\beta^2}\,\vec{\beta}\cdot\vec{x}\right)\vec{\beta} \tag{29}$$

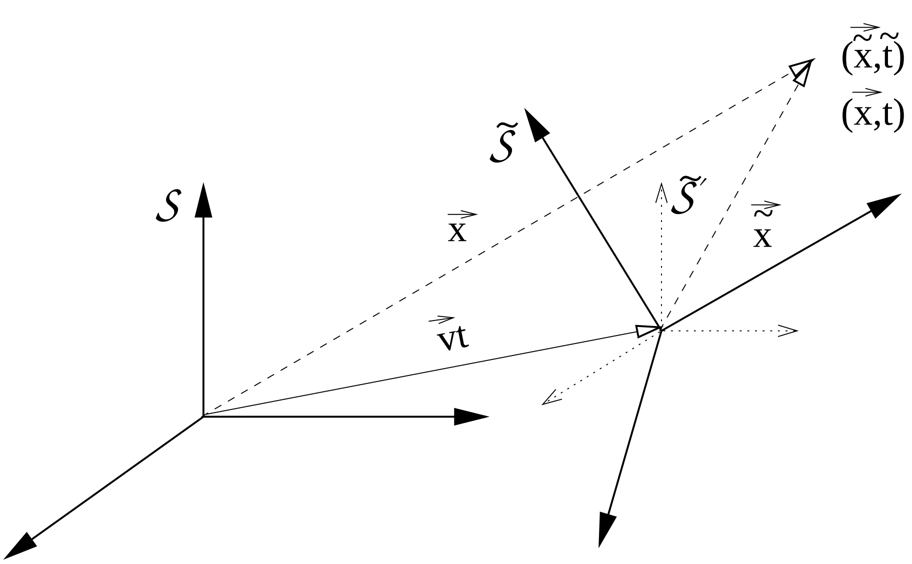
Usually, one considers the special case when the relative motion between $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ is only in the x^1 direction, *i.e.,* $v^1 = v, v^2 = 0, v^3 = 0$. Then the coordinate transformation becomes

$$\widetilde{t} = \gamma\left(t - \frac{v}{c^2}x^1\right)\,, \qquad \widetilde{x}^1 = \gamma\left(x^1 - vt\right)\,, \qquad \widetilde{x}^2 = x^2\,, \qquad \widetilde{x}^3 = x^3 \tag{30}$$

So far we have considered the special case of Lorentz transformations where the coordinate axes in $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ are parallel. Let us now consider the general case where the coordinate axes of $\widetilde{\mathcal{S}}$ are rotated with respect to that of $\mathcal{S}$ by a 3×3 rotation matrix R ($R^T R = 1$) (see figure below). In this case, we can first choose a new coordinate system $\widetilde{\mathcal{S}}\,'$ the origin of which coincides with the origin of $\widetilde{\mathcal{S}}$ but the axes of which are parallel to that of $\mathcal{S}$. Clearly $\widetilde{\mathcal{S}}\,'$ and $\widetilde{\mathcal{S}}$ are related by the rotation R which can be embedded into a Lorentz transformation L_R,

$$L_R = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}$$

Now to transform a vector from $\mathcal{S}$ to $\widetilde{\mathcal{S}}$, we first make a Lorentz transformation from $\mathcal{S}$ to $\widetilde{\mathcal{S}}\,'$ ($\tilde{x}' = Lx$) and then rotate $\tilde{x}'$ into $\tilde{x}$ using L_R: $\tilde{x} = L_R\tilde{x}' = L_R L x$. Thus the



general transformation can be parameterized as

$$L_{gen} = L_R\, L = \begin{pmatrix} L^0_{\ 0} & L^0_{\ j} \\ R^k_{\ i} L^i_{\ 0} & R^k_{\ i} L^i_{\ j} \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta_j \\ -\gamma R^k_{\ i}\beta^i & R^k_{\ j} + \dfrac{1-\gamma}{\beta^2} R^k_{\ i}\beta^i\,\beta_j \end{pmatrix} \tag{31}$$

Note that L_{gen} has 6 independent parameters, 3 corresponding to the components of the relative velocity $\vec{v}$ and 3 more from the rotation matrix $R^i_{\ j}$ consistent with the defining equatin for L.

7 Relativistic Addition of Velocities

Assume that observers in frames $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ measure the velocity of a moving point as $\vec{u}$ and $\vec{\tilde{u}}$ respectively. We want to find the relation between the velocities when the coordinates in the two frames are related by a Lorentz transformation $\tilde{x}^\mu = \sum_{\nu=0}^3 L^\mu_{\ \nu}\, x^\nu$. Over a small time interval the observers measure displacements $\Delta x^i = u^i \Delta t$ and $\Delta \tilde{x}^i = u^i \Delta \tilde{t}$ for $i = 1, 2, 3$. Also in our notation, $\Delta x^0 = c\Delta t$ and $\Delta \tilde{x}^0 = c\Delta \tilde{t}$. Since the Lorentz transformation is linear, the intervals are related by $\Delta \tilde{x}^\mu = \sum_{\nu=0}^3 L^\mu_{\ \nu}\, \Delta x^\nu$. Using these equations, one can easily see that

$$\frac{\tilde{u}^i}{c} = \frac{\Delta \tilde{x}^i}{\Delta \tilde{x}^0} = \frac{L^i_{\ \nu}\, \Delta x^\nu}{L^0_{\ \nu}\, \Delta x^\nu} = \frac{L^i_{\ j}\, \Delta x^j + L^i_{\ 0}\, \Delta x^0}{L^0_{\ j}\, \Delta x^j + L^0_{\ 0}\, \Delta x^0}$$

which on using $\Delta x^i = u^i \Delta t$ and $\Delta x^0 = c\Delta t$ leads to the desired result

$$\frac{\tilde{u}^i}{c} = \frac{L^i_{\ j}\, u^j + L^i_{\ 0}\, c}{L^0_{\ j}\, u^j + L^0_{\ 0}\, c} \tag{32}$$

This is the Lorentzian law of addition of velocities. In particular, when the frame $\widetilde{\mathcal{S}}$ moves away from $\mathcal{S}$ with velocity v in the x^1 direction, then the above relation between $\tilde{u}^i$ and

u^i reduces to the familiar form of the law of addition of velocities,

$$\tilde{u}^1 = \frac{u^1 - v}{1 - \dfrac{vu^1}{c^2}}\,, \qquad \tilde{u}^2 = \frac{u^2}{\gamma\left[1 - \dfrac{vu^1}{c^2}\right]}\,, \qquad \tilde{u}^3 = \frac{u^3}{\gamma\left[1 - \dfrac{vu^1}{c^2}\right]}$$