We have spent quite a bit of time on the notions that lead to the Lorentz boost transformations. In this supplemental material I shall go along with the discussion we had since we shall not have a meeting next week. I shall also include the sketch of what we have discussed so far. Some assignments in the form of fill in the blanks are suggested but again this is not a homework assignment. You shall receive the first homework assignment shortly.

**Lorentz boost transformations:**

Assume there exists an inertial observer $S$. Now consider another inertial observer $S'$ which moves with respect to $S$ with a velocity $\vec{v} = c\beta \hat{i}$ (note that this velocity is defined in frame $S$). Now let us consider an event $A$. This event will have spacetime co-ordinates $\{ct, x, y, z\}$ in frame $S$ and $\{ct', x', y', z'\}$ in $S'$.

The linear (as we have justified from the homogeneity of vacuum for inertial observers) relation between these different co-ordinates are given as follows:

$$ct' = \gamma (ct - \beta x)$$
$$x' = \gamma (x - \beta ct)$$
$$y' = y$$
$$z' = z$$
$$\gamma \equiv \frac{1}{\sqrt{1 - \beta^2}}$$

The above is generalized to the following for a generic $\vec{v} = c\vec{\beta}$

$$\vec{r}' = \vec{r} + (\gamma - 1)(\vec{r} \cdot \hat{\beta})\hat{\beta} - \gamma \beta ct$$
$$ct' = \gamma (ct - \vec{\beta} \cdot \vec{r})$$

Derive the form above if you feel like it. Hint: It can be shown that space co-ordinates perpendicular to the relative velocity should receive no change.

We have also shown that these boost transformations leave the following quantity invariant

$$c^2 \Delta t^2 - \Delta r^2 = c^2 \Delta t'^2 - \Delta r'^2 \equiv \Delta s^2.$$ 

The quantity $\Delta s^2$ is called the invariant interval and genuinely defines three categories: $\Delta s^2 > 0$ which is called a time-like interval, $\Delta s^2 < 0$ which is called a space-like interval and $\Delta s^2 = 0$ which is called a light-like interval.

**Lengths of rods are measured differently if they move:** We did not discuss this during the lectures as I was more focused on the form of equations of motion. Now we can do it.

What is a "length measurement"? Consider a rod which moves and its length is to be measured. Different observers should agree on an algorithm for this experiment. The simplest algorithm is that the observer has to identify the locations of both ends of the rod simultaneously in its own frame. But this means that these are two events which are simultaneous but separated with some length so it is a space-like separation and thus for different observers this experiment will not be classified as a "length measurement". This follows since
we have shown that for space-like separations time ordering of events is not an invariant concept.

Now consider the observer $S$ and the rod as the observer $S'$. The rod moves with respect to $S$ with velocity $\vec{v} = c\beta\hat{i}$. Let us also assume that the observer $S'$ says that the rod -which is stationary for $S'$- has length $L_0$.

Observer $S$ identifies the two ends of the rod simultaneously. So there are two events: $A_l$ with $ct_l = 0$ and $x_l = 0$ and $A_r$ with $ct_r = 0$ and $x_r = L$ where $L$ is the measured length of the rod in $S$. The co-ordinates of these two events can be translated to the language of $S'$ and the required link will be established with the observation that these events occur at the ends of the rod: That is we have $x'_l = 0$ (this one immediate from the choice of origins so not a true condition) and $x'_r = L_0$ (this one is the true condition). As mentioned we also have $x'_r = \gamma(x_r - \beta ct_r)$ from Lorentz transformations; so we get $L_0 = \gamma L$, or better

$$L = L_0\sqrt{1 - \beta^2}$$

So a moving rod’s length is measured shorter than its length while it is at rest. Is there a paradox here? No because even though $S$ made a length measurement in its own frame $S'$ do not agree that it is a length measurement since in its frame $S$ did not identify two ends at the same time. In fact from the perspective of $S'$ we can get $ct'_l = 0$ and $ct'_r = -\gamma\beta L$ which show that the co-ordinate of the right side of a rod which moves to the right is measured earlier than its left side! $S'$ says that under these circumstances it is all too meaningful that you did measure it shorter but this was not the standard we agreed upon for what a length measurement is!

Note that if we are given that all the rod’s are standard: That is if all the rods are known to have the standard proper (in their own frame) length $L_0$ this phenomenon allows us to measure the speed of a rod by making two location measurements simultaneously. Is this possible in Newtonian physics?

Now devise an experiment that will unambiguously show that there can be no contracted measurements of lengths along the directions perpendicular to velocity of an object. This somewhat fills a logical gap since we have put $y' = y$ and $z' = z$ in the Lorentz transformations a bit arbitrarily.

**Tic-tac duration of a moving clock is measured differently if they move:** Consider a moving clock in a frame $S'$. The tic is an event in $S'$ with co-ordinates $ct'_\text{tic} = 0$ and $x'_\text{tic} = 0$ and the tac is an event in $S'$ with co-ordinates $ct'_\text{tac} = T_o$ and $x'_\text{tac} = 0$ so $\Delta t' = T_o$ is the duration between the tic and the tac. One can using Lorentz transformations that in the frame $S$ this duration is measured as

$$T = \frac{T_o}{\sqrt{1 - \beta^2}}$$

**About Four-Vectors and index notation**

We already know that our spacetime continuum requires 4 co-ordinates for an event. In exact analogy to the euclidean three-vectors a four-vector can be expanded in terms of basis vectors. So a four-vector is

$$X \equiv \sum_{\mu=0}^{3} X^\mu e_\mu$$

2
where the boldface is used four vectors (in analogy for three vectors) which are elements of the four dimensional vector space where we define the events and index zero refers to the temporal co-ordinate of the four vector. For instance if we are talking of an event the zero index refers to the time of the event.

The difference lies in the so-called inner product of vectors. Since we know that our beloved invariant interval has a relative minus sign we can not make this spacetime Euclidean. Let us define

$$X \circ X = (X^0)^2 - (X^1)^2 - (X^2)^2 - (X^3)^2$$

If $X$ is to denote an event $X^0 = ct$, $X^1 = x$, $X^2 = y$ and $X^3 = z$ as expected.

But since this vector space is also linear the object above can also be written as

$$X \circ X = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} X^\mu X^\nu \epsilon_\mu \circ \epsilon_\nu$$

To have these match we need to define

$$\epsilon_\mu \circ \epsilon_\nu \equiv \eta_{\mu \nu}$$

with $\eta_{00} = 1$ and $\eta_{11} = \eta_{22} = \eta_{33} = -1$ whereas for all the other possibilities $\eta_{\mu \nu}$ vanishes. This is in analogy to the usual $i$, $j$ and $k$ objects and · vector product operation in Euclidean space. So we call call the object $\eta_{\mu \nu}$ the metric.

A good notational discipline which is generally used in physics is the following: Whenever an index is denoted by a greek letter it refers to all the spacetime components and whenever an index is denoted by a latin letter it refers to space components only. For instance $\eta_{ij} = -\delta_{ij}$.

Surely one can pick a different set of base vectors and describe the same four-vector with components that are in accord with these vectors. This can not change the vector itself. Let us use an example from ordinary three vectors: Assume that a rod with one end painted red sits in space. This object is itself whatever basis one chooses to describe it. For instance say $\vec{A} = a \hat{i}$ in one dimensions. Another choice of basis vector is say $-\hat{i}$ and then its component must be $-a$ since $\vec{A} = (-a)(-\hat{i}) = a\hat{i}$.

Of course the choice of basis is completely arbitrary except that all these choices must be in linear relation with each other and that these relations must be invertible. So if from one set of unit vectors one linearly finds another set via a matrix than that matrix must have non-zero determinant.

What does the principle of relativity imposes on this structure. Well in one frame $e_0$ refers to a time-like unit vector, that is it defines the direction of time. For another inertial observer another set of unit vectors are needed. Let us call them $e'_\mu$. This set must contain one and only one time-like direction as well as three space-like directions and whenever such directions are found one can normalize the norms to the standard. This is simply saying that we can endow all inertial observers with standard clocks and rods. So we have the following conditions immediately from the principle of relativity concerning inertial observers
where $\Lambda^{-1}$ refers to the invertible matrix that makes us perform the change of language between these to inertial frames: That is to linearly transform the unit vectors. The definition via the inverse is just for future convenience. We immediately get from the above

$$
\sum_{\mu=0}^{3} \bar{\mu} = 0 \quad (\Lambda^{-1})_{\mu}^{\bar{\mu}} \eta_{\bar{\mu} \bar{\mu}} = \eta_{\lambda \nu}
$$

Now it is a good time to introduce a convention to get rid of those summation signs. Note that the sums turn out to contain one upper and one lower index. So we assume any index which is repeated once in the upper and once in the lower is to be summed from 0 to 3. If any another meaning is to be placed upon and expression in conflict with this it has to be mentioned. For instance $(\Lambda^{-1})_{\mu}^{\mu}$ no sum means we consider the collection of 4 objects instead of a single one. Thus the last expression above reads

$$
(\Lambda^{-1})_{\mu}^{\mu} (\Lambda^{-1})_{\nu}^{\nu} \eta_{\bar{\mu} \bar{\nu}} = \eta_{\lambda \nu}
$$

So it tells us about a set of linear transformation that do not change the metric of the spacetime. Using it we can show the following

$$
X_{\mu}^{\nu} (\Lambda^{-1})_{\mu}^{\bar{\mu}} = X_{\mu}^{\bar{\mu}}
$$

Thus defining a new object

$$
X_{\bar{\mu}}^{\mu} \equiv \Lambda^{\mu}_{\bar{\mu}} X_{\mu}
$$

We realize that we must say

$$
(\Lambda^{-1})_{\mu}^{\bar{\mu}} = \Lambda^{\bar{\mu}}_{\mu}
$$

This notation makes it possible to use the following symbol

$$
(\Lambda^{-1})_{\mu}^{\alpha} \Lambda^{\mu}_{\beta} = \delta_{\beta}^{\alpha}
$$

where $\delta_{\beta}^{\alpha}$ is the Kroenecker symbol: It is 1 whenever $\alpha = \beta$ and 0 otherwise.

Using the Kroenecker symbol one can define the object $\eta^{\alpha \mu}$ via the following definition

$$
\eta^{\alpha \mu} \eta_{\alpha \nu} \equiv \delta^{\alpha}_{\mu}
$$

Thus using $\eta$ new objects of the form

$$
X_{\mu} = X^{\nu} \eta_{\mu \nu}
$$

and
\[ X^\mu = X^\nu \eta^{\mu\nu} \]

With these new objects the Lorentz norm of a four-vector can be written as
\[ X^\mu X_\mu \]
which of course becomes the invariant interval if we talk about the difference of two four-vectors describing two events.

For the standard Lorentz transformation mentioned at the beginning of this supplemental material we shall have
\[
\begin{align*}
\Lambda^0_0 &= \Lambda^1_1 = \gamma \\
\Lambda^0_1 &= \Lambda^1_0 = -\beta \gamma \\
\Lambda^2_2 &= \Lambda^3_3 = 1
\end{align*}
\]

The above completely makes sense of the following
\[ X'^\mu = \Lambda^\mu_\nu X^\nu \]
as a formal expression of the Lorentz boost transformations.

The road to the equations of motion

We have seen that we have to describe an event via a four-vector in a unified structure called spacetime. This is in contrast to the Newtonian picture where an event is described with a three vector and a number called time which have no relation between them.

So to describe the motion of a point particle we need to use the position four-vector. Furthermore the physical motion of a particle should be a succession of time-like separated events with an invariant ordering: The particle is a particle at time \( t_2 \) because it was a particle at an earlier time \( t_1 \). This requirement makes it possible to use the infinitesimal invariant interval as an invariant quantity which will label all these successive events which form the trajectory. So we define
\[ V^\mu \equiv c \frac{dX^\mu}{ds} \]
where in a particular frame \( S \)
\[ ds = c dt \sqrt{1 - u^2/c^2} \equiv \frac{cdt}{\Gamma(u)} \]
with \( \vec{a} \equiv d\vec{x}/dt \). The quantity \( ds/c \) is also called the proper time of the particle because with respect to an inertial observer instantaneously in congruent motion with the particle it coincides with the time measured with a clock moving with the particle (since the particle does not move in its own frame). We shall talk about the clock hypothesis if we have time.

Anyway... The elements of this four-vector are as follows
\[
\begin{align*}
V^0 &= c \Gamma(u) \\
V^i &= \Gamma(u) u^i
\end{align*}
\]
The space components can also be combined in a three-vector notation $\vec{V} = \Gamma(u)\vec{u}$.

$V^\mu$ is a four-vector in the sense that it transforms as $X^\mu$ since $ds$ is the same for all inertial observers. Thus we have the elements with respect to another inertial frame

$$V^{\prime \mu} = \Lambda^{\mu \nu} V^\nu$$

But by definition -as obvious from the principle of relativity- in this new frame the elements must be given as

$$V^{\prime 0} = c\Gamma(u^\prime)$$
$$V^{\prime i} = \Gamma(u^\prime) u^{\prime i}$$

which provides a relation between $\vec{\beta}$ which is the relative speed of these two inertial frames and $\vec{u}$ and $\vec{u}^\prime$.

In view of its four-vector properties the object $V^\mu V_\mu$ must be the same for all inertial observers. A direct calculation will yield

$$V^\mu V_\mu = c^2$$

So this quantity is not only invariant between inertial frames but it is also a constant. And this is true no matter what the actual trajectory of the particle is. One can say that we have found the first kinematical law of motion in special relativity!

Now why we should take this object seriously? One can show that it is the only object which contains the ordinary velocity for speed much less than the speed of light: One can show that $V^0 \rightarrow c$ and $\vec{V} \rightarrow \vec{u}$ under $u << c$.

In view of this let us define $P^\mu \equiv mV^\mu$ where we shall insist that $m$ is the inertial mass and it is a frame independent quantity. Now let us define the following

$$M^\mu \equiv c \frac{dP^\mu}{ds}$$

We shall call $M^\mu$ the Minkowski four-force. The following is also immediate if we assume $dm/ds = 0$,

$$M^\mu P_\mu = 0$$

As we have discussed at length in class these will culminate in the following results.

1) The definition of linear momentum has to be changed to $\vec{P} = m\Gamma(u)\vec{u}$.
2) The definition of kinetic energy has to be changed to $E = mc^2\Gamma(u)$.
3) The element $P^0 = E/c$ so energy and momentum live inside a four-vector together.
4) The equations of motion are

$$\frac{dE}{dt} = \vec{F} \cdot \vec{u}$$
$$\frac{d\vec{P}}{dt} = \vec{F}$$
5) The elements of the Minkowski four-force are $\vec{M} = \Gamma(u)\vec{F}$ and $M^0 = \vec{u} \cdot \vec{F} / c$.

6) The object $\vec{F}$ is called force in the sense of Newton: The object that causes the acceleration of the particle.

**The relative nature of $\vec{F}$**

Note that the Minkowski four-force is a four-vector: It transforms under Lorentz boosts as the position four-vector $X^\mu$. These mean the following

$$M'^\mu = \Lambda^\mu_\nu M^\nu$$

and of course according to the principle of relativity we must have $M'^0 = \Gamma(u')\vec{F}' \cdot \vec{u}' / c$ and $\vec{M}' = \Gamma(u')\vec{F}'$. These will make a connection between $\vec{\beta}$ which is the relative velocity between inertial observers and $\vec{F}$ and $\vec{F}'$. So in general $\vec{F}'$ will not be equal to $\vec{F}$. And thus force itself is not a special relativistically invariant concept unlike Newtonian system.

To explore this let us recall that since $M^\mu$ is a four-vector its Lorentz norm

$$M'^\mu M'^\mu = \Gamma'^2 \left[ (\vec{u}' \cdot \vec{F}')^2 - c^2 F'^2 \right]$$

which will mean

$$\Gamma^2 \left[ (\vec{u} \cdot \vec{F})^2 - c^2 F^2 \right] = \Gamma'^2 \left[ (\vec{u}' \cdot \vec{F}')^2 - c^2 F'^2 \right]$$

where $\Gamma' = \Gamma(u')$. The equation above makes it manifest that unless we confine the study to a single space dimension -which will give $F' = F$- the force is a different vector in different frames unless it identically vanishes.

**Massless Particles**

From our special relativistic definition of the momentum and kinetic energy one can find the following relations

$$E^2 - \vec{p}^2 c^2 = m^2 c^4$$

$$\vec{u} = \frac{\vec{p} c^2}{E}$$

Note that these relations are more general in that their $m \to 0$ limit is well behaved. We immediately get $E = pc$ and thus

$$|\vec{u}| = c$$

for massless particles: They must move with the speed of light if they exist which we know they do.

**Departures from some classical mechanics dogmas**

Recall that we have found the equation of motion to be for a massive particle to be

$$\frac{d}{dt} (m \Gamma(u) \vec{u}) = \vec{F}$$

defining $\vec{a} \equiv d\vec{u}/dt$ we shall get

$$\vec{F} = m \Gamma \vec{a} + m \vec{a} \frac{d\Gamma}{dt}$$
So the acceleration is not proportional to the force. Instead what can be said is that instantaneously, the force, the acceleration and the velocity define a plane in space.

An immediate consequence of this is that the change in the kinetic energy can come from directions unrelated to the acceleration. Which makes a strong emphasis on an intuitive dogma we usually get from classical mechanics. Let us rewrite the kinetic energy as follows\(^1\)

\[
E = \sqrt{p^2 c^2 + m^2 c^4}
\]

This can not be written as a separated sum of kinetic energies referring to momenta along different directions. In sharp contrast the low velocity limit

\[
E \approx mc^2 + \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2)
\]

This and the fact that the equations of motion are separated as \(F^i = dP^i/dt\) - in either formalism- we can find ourselves talking about energy along a given direction. This kind of language is of course fundamentally wrong: The energy is not a vector. It does NOT transform like a vector under rotations. It is a scalar under rotations.

An immediate consequence of the consideration above is that it is possible that the velocity might change in directions perpendicular to the force.

---

\(^{1}\)We take the positive root . The negative root is very relevant if one considers quantum mechanics and is related to the existence of anti-particles