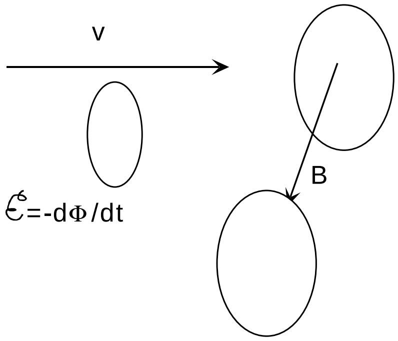
(Chapter head:)Special Relativity

1 The need for special relativity

It is perplexing aspect of electrodynamics that a charge moving at a constant velocity has an Magnetic field whereas a stationary charge has only an electric field. Thus the laws of electrodynamics would seem to be written with respect to a fixed frame of reference (formally known as the ether.) It is surprisingly difficult, however, to come up with an experiment that could determine what the ether is stationary with respect to. Let us think of a few:

Imagine you lived before Faraday and you decided to test for a hoop passing through a magnetic field in the frame of reference of the magnet:



(1)

Even the pre-Faraday physicist would not be surprised to find that an electromotive force is induced in the wire. Recall that electromotive force is given by

$$Emf = \oint \vec{f} \cdot d\vec{\ell} \tag{2}$$

and that the charges in the wire feel a force per unit charge because $\vec{f} = \vec{v} \times \vec{B}$. From the frame of reference of the hoop however, its charges are stationary so no force should be seen. Thus the emf induced in the wire should be proportional to the velocity of the hoop with respect to the ether- measure the emf and you know the ether speed, right? Wrong! Faraday found that a change magnetic field creates an electric field according to

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{3}$$

causing the emf to be the same, regardless of the velocity of the observer. The moving observer does observe both an electric and magnetic field whereas the observer at rest with respect to the magnetic field observes only a magnetic field - but no matter, both predict the same measurable quantity.

Now let us jump forward in time to the point in time where Maxwell had already fixed up E&M with the displacement current. Now Maxwell's equations are

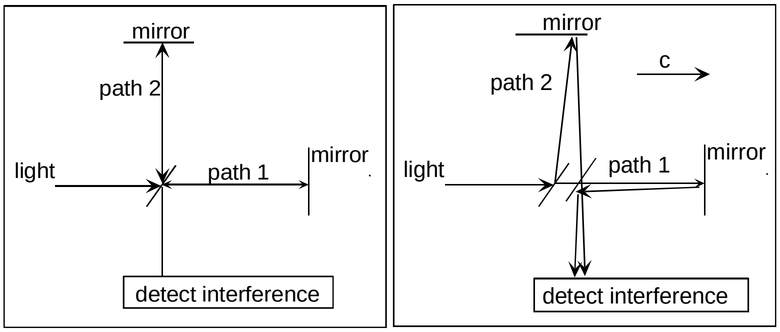
$$\vec{\nabla} \cdot \vec{E} = \rho/\varepsilon_o \tag{4}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{5}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \tag{6}$$

$$\vec{\nabla} \times \vec{B} = \mu_o \vec{J} + \mu_o \varepsilon_o \frac{\partial \vec{E}}{\partial t} \tag{7}$$

and light has been predicted to exist. With these equations, it is still not entirely easy to come up with an experiment that will show a measurable effect to reveal the ether, but Michaelson and Morely did. They figured that the earth was traveling at $30 km\, s^{-1}$ not too far from the speed of light ($300000 km$ s^{-1}). Since it seemed improbably that the ether would choose to travel with the earth, they thought the following device would notice an effect:



(8)

Light was allowed to take one of two paths of equal length, assuming the device was not moving. If the device is moving along the direction of the incident light, then a maximum difference in path length would occur. This difference is given by the difference of d_2 and d_1. d_2 can be found in terms of the speed of light c, the speed of the device with respect to the either v, and distance between the beam splitter and end mirror d.

$$d_2 = 2\sqrt{(\frac{vt}{2})^2 + d^2} \tag{9}$$

$$= ct \tag{10}$$

Solving for ct we have

$$\begin{aligned} c^2t^2 &= v^2t^2 + 4d^2 && (11) \\ ct &= d_2 = \frac{2d}{(1 - v^2/c^2)^{1/2}} && (12) \\ d_2 &= \frac{2d}{(1 - v^2/c^2)^{1/2}} && (13) \\ &\approx 2d(1 + \frac{v^2}{2c^2}) && (14) \end{aligned}$$

The time to take the path in the direction of the ether was (by Galilean logic) thought to be longer. In the frame of the ether, the both mirrors move with a velocity $\vec{v}$. The time it takes light to get from the angled mirror to the back-reflecting mirror is determined from

$$\begin{aligned} ct_+ &= d + vt_+ && (15) \\ t_+ &= \frac{d}{c - v} && (16) \end{aligned}$$

On the way back the angled mirror is approaching, so

$$\begin{aligned} ct_- &= d - vt_- && (17) \\ t_- &= \frac{d}{c + v} && (18) \end{aligned}$$

so the total distance traveled by the light going by path 1 is given by

$$\begin{aligned} d_1 &= (t_+ + t_-)c && (19) \\ &= d(\frac{c}{c+v} + \frac{c}{c-v}) && (20) \\ &= 2d\frac{1}{(1 - (v/c)^2)} && (21) \\ &\approx 2d(1 + (\frac{v}{c})^2) && (22) \end{aligned}$$

The difference between the two paths was then given by

$$\begin{aligned} \Delta d &= d_1 - d_2 && (23) \\ &\approx d\left(\frac{v}{c}\right)^2 && (24) \\ &\approx 10\,\mathrm{nm}\ (\text{for } v = v_{earth}, d = 1.0m) && (25) \end{aligned}$$

The accuracy in the original Michelson Morely experimentalist was good enough to say that if the ether hypothesis was correct, the ether must move with the earth to within one part in three of the earth's velocity.

What they found was one of the greatest physics puzzles ever: The interference pattern did not depend in the slightest on the orientation of their device.

What could be going on? Everything from Maxwell's equations being in error to the Earth dragging the ether was proposed. It took Einstein to realize that neither was true: Rather classical mechanics was in error.

Einstein proposed two postulates:

1. The laws of physics apply in all inertial reference system (26)
2. The speed of light is the same in every frame. (27)

The first postulate is reasonable and not too revolutionary. The second requires to throw away something so "obvious" that it had never been questioned: Galileo's addition rule for velocities (That is an object a observed to moving at a velocity $\vec{v}_A$ by an observer traveling at a velocity $\vec{v}_o$ will be traveling $\vec{v}_o + \vec{v}_A$ with respect to an observer at

$$\vec{v}'_A = \vec{v}_o + \vec{v}_A \tag{28}$$

If this rule is not right, then what could be. Suppose we just consider one dimension (motion along the z axis.) Then

$$v'_A = v_o + v_A \tag{29}$$

makes so much sense that we had better get this answer for the limit of speeds we typically experience. A simple thing we can write down that behaves in this way for small velocities, but has c a universal constant is

$$v'_A = \frac{v_o + v_A}{1 + v_o v_A / c^2} \tag{30}$$

Of course one can arrive at this result simply by messing around with combinations of v_o and v_A and there is nothing wrong with this exercise. Assuming this result is true simply because it is consistent with the Michelson Morely experiment is of course silly. After all, I could write down an expression such as

$$v'_A = \begin{cases} v_o + v_A \quad (v_o + v_A < c) \\ c \quad (v_o + v_A < c) \end{cases} \tag{31}$$

and it would also be consistent with the experimental results. The transformation of eq 30 is, however, the result of a Lorentz transformation which is a change in variables that leave Maxwell's equations intact. The fantastic success of Maxwell's equations together with the fact that Eq 30 rids us of the puzzling requirement for an ether to exist is what makes this form the one of choice.

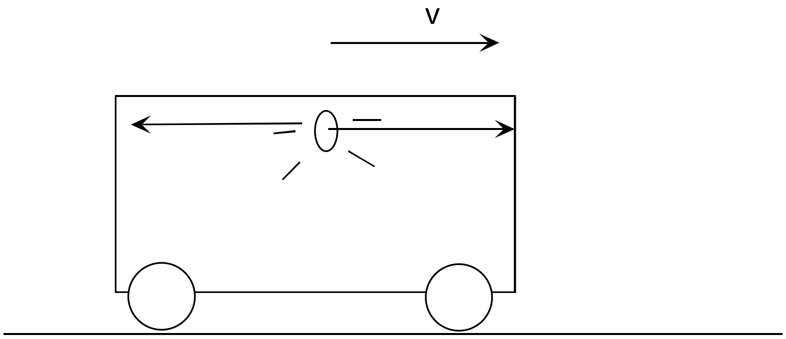
We will eventually show explicitly how Maxwell's equations are invariant under a Lorentz transformation, but not a Galilean transformation. Thus demonstration takes a fair bit of sophistication, so for now we will simply take it on faith and explore the implications of 30 and the theory of special relativity.

2 Implications of special relativity

Following Griffiths, let us discuss three incredible implications of special relativity, namely The relativity of simultaneity, time dilation, and Lorentz contraction

2.1 the relativity of simultaneity

Imagine a box car on a high-speed rail. The box car travels at a sizeable fraction of the speed of light. In the middle of the box car, a flash bulb goes off. From the point of view of someone in the box car, the light hits the sides of the car at the same instant.



(32)

From the point of view of someone observing from the side of the track, the light moving forward must travel longer in order to reach the end of the car, whereas the light traveling backward must travel less. Thus the two events (light hitting the back of the car and light hitting the front) must occur at different times! Thus simultaneous events in one frame will not be simultaneous events in another.

You might have just done a quick calculation in your head, setting $ct_{front} = \frac{L}{2} + vt_{front}$, $ct_{back} = \frac{L}{2} - vt_{back}$, $\Delta t = t_{front} - t_{back} = \frac{L}{2}(\frac{1}{c-v} - \frac{1}{c+v})$ but this is actually not correct. It is short a factor of $\sqrt{1-(v/c)^2}$, as we shall soon see.

2.2 Time dilation

Now let us consider light that bounces off the floor of the train and returns to the bulb. In the inertial frame of the train, this event takes a time

$$\Delta t' = 2h/c \tag{33}$$

whereas in the frame of one watching the train pass by, the path length l is longer than $2h$. Specifically

$$\begin{aligned} l &= 2\sqrt{h^2 + (\frac{v\Delta t}{2})^2} && (34) \\ \Delta t &= l/c && (35) \\ &= \sqrt{\left(\frac{2h}{c}\right)^2 + (\frac{v\Delta t}{c})^2} && (36) \\ &= \sqrt{\Delta t'^2 + (\frac{v\Delta t}{c})^2} && (37) \end{aligned}$$

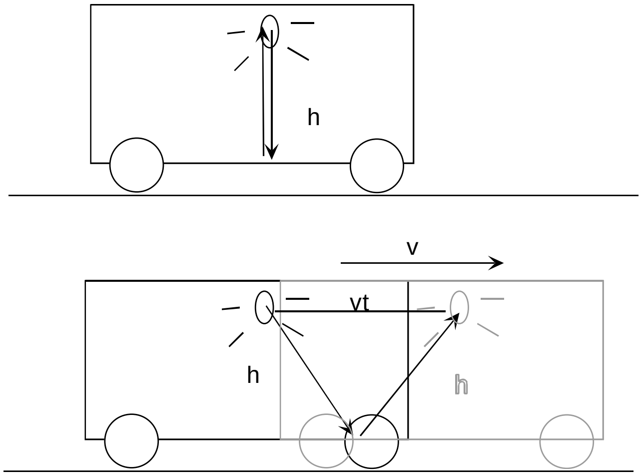


Figure 1:

or

$$\Delta t' = \Delta t\sqrt{1 - (\frac{v}{c})^2} \tag{38}$$

$$= \Delta t/\gamma \tag{39}$$

Suppose $v = 3/5\ c$ so that $\sqrt{1 - (\frac{v}{c})^2} = \frac{4}{5}$. Then if it takes 4 ns for the light to travel up and down the box car in the frame of someone traveling with the box car, a person stationary with respect to the tracks will observe a $5ns$ delay for the light to travel up and down. But if this event is slowed down by a factor $1/\sqrt{1 - (\frac{v}{c})^2}$, so must all other events. Thus if we observe a clock on the train as it passes, the clock will appear to be running slow by an amount $1/\sqrt{1 - (\frac{v}{c})^2}$. This dilation factor is always greater than 1 and appears so often that it is given its own Greek letter:

$$\gamma = \frac{1}{\sqrt{1 - (\frac{v}{c})^2}} \tag{40}$$

Problem: How fast must a muon created a height $h = 15km$ above the earths atmosphere travel if it is to reach the earth before in its lifetime ($\tau = 2 \times 10^{-6}s$.)

Answer: $\tau c = 600m$ so it would appear that the muon could never get to the earth. However we must remember that if we observe a muon as it travels, its decay time is a marker of how fast physics is going on in its rest frame. If a time Δt passes in our frame of reference, only $\Delta t/\gamma$ has passes in the frame of reference of the particle. Thus the muon can travel to the earths surface. We have

$$\begin{aligned}
\Delta t &= h/v && (41)\\
&= \tau/\sqrt{1-(\frac{v}{c})^2} && (42)
\end{aligned}$$

Solving for v we have

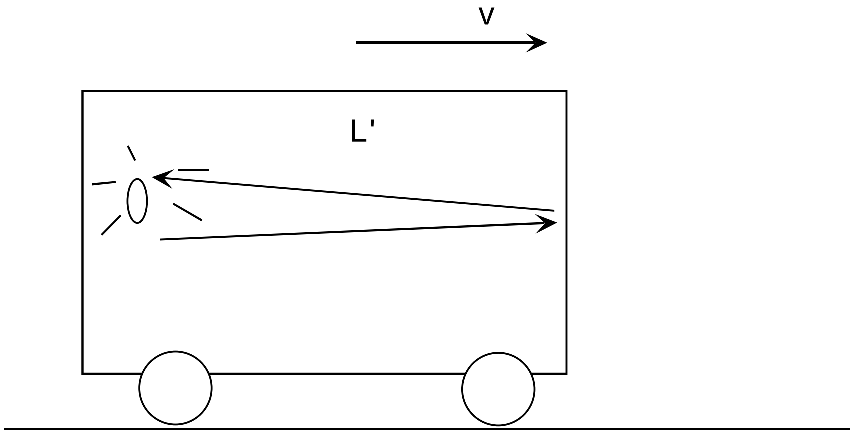
$$\begin{aligned}
h\sqrt{1-(\frac{v}{c})^2} &= v\tau && (43)\\
1-\left(\frac{v}{c}\right)^2 &= \left(\frac{v\tau}{h}\right)^2 && (44)\\
1 &= \frac{v^2}{c^2}(1+\left(\frac{\tau c}{h}\right)^2) && (45)\\
v &= \frac{c}{\sqrt{1+(\frac{\tau c}{h})^2}} && (46)\\
&= c\left[1+\left(\frac{600m}{15000m}\right)^2\right]^{-1/2} && (47)\\
&= 0.992c && (48)
\end{aligned}$$

Notice that no matter what the combination of τ or h, it is always possible to get that particle to the earths surface before it is likely to decay (v never has to be greater than c.)

Question to resolve: Who is right?

A man on a train has a pocket watch whereas a women on the ground wears a wrist watch. The man on the train passes the women and observes her wrist watch running slow by a factor γ. The women on the ground watches the man pass by and observes his pocket watch to be slow by a factor γ. How can both the man and women be correct?

2.3 Lorentz contraction



$$(49)$$

Imagine a flash lamp producing a burst of light at one end of a box car and one measures the time it takes the light to hit the opposite and return. From the frame of reference of one traveling with the car,

$$\Delta t' = 2L'/c \tag{50}$$

Here I have added a prime to L' because the length of the box car in the frame of reference of one that travels is not necessarily the same as one that does not. As we shall see more formally in the next lecture, dimensions perpendicular to the motion (such as h) do not change.

From the frame of reference of one who is at rest, the light makes it to the opposite wall in a time t_+

$$\begin{aligned} ct_+ &= L + vt_+ \qquad (51) \\ t_+ &= \frac{L}{c-v} \qquad (52) \end{aligned}$$

and back in a time t_-

$$\begin{aligned} ct_- &= L - vt_- \qquad (53) \\ t_- &= \frac{L}{c+v} \qquad (54) \end{aligned}$$

so the round trip time must be

$$\begin{aligned}
\Delta t &= (t_+ + t_-) && (55)\\
&= L\left(\frac{1}{c-v} + \frac{1}{c+v}\right) && (56)\\
&= L\left(\frac{2c}{c^2 - v^2}\right) && (57)\\
&= \frac{2L}{c}\frac{1}{1-(v/c)^2} && (58)\\
&= \frac{2L}{c}\gamma^2 && (59)
\end{aligned}$$

But we know already that the round trip time must be the time that occurred in the box car multiplied by a factor $1/\gamma$:

$$\begin{aligned}
\Delta t' &= \Delta t/\gamma && (60)\\
&= (2L'/c) && (61)\\
&= (2L/c)\gamma && (62)
\end{aligned}$$

This can only happen if $L = L'/\gamma$. Thus objects that are moving are observed to be a factor of γ shorter then they are in there rest frame.

Question: A man on a train holding a wooden meter stick observes a women at the side of the track holding a metal meter stick. The man observes the women's metal meter stick to be shorter by a factor of γ whereas the women observes than man's wooden stick also to be short. How can they both be right?

2.4 The Lorentz Transformation

Up to this point in our discussion we have simply investigated the consequences of the Einstein's postulate that the speed of light is the same in all reference frames, making an additional assumption that the z-dimension of an object is not affected by the speed of the observer in the x-y plane. With these statements we found that simultaneity is not conserved in different inertial frames, that moving clocks run slow by a factor γ (time dilation,) and moving objects appear shorter by a factor γ (Lorentz contraction.) These crazy features are consequences of Lorentz transformations, the topic of today's lecture.

3 Transformations that don't change Maxwell's equations.

3.1 Charge, parity, and time-reversal symmetry.

The starting point of this conversation is Maxwell's equations and the Lorentz force law:

$$\vec{\nabla}\cdot\vec{E} = \rho/\varepsilon_o \tag{63}$$

$$\vec{\nabla}\times\vec{E} = -\frac{\partial\vec{B}}{\partial t} \tag{64}$$

$$\vec{\nabla}\cdot\vec{B} = 0 \tag{65}$$

$$\vec{\nabla}\times\vec{B} = \mu_o\vec{J} + \mu_o\varepsilon_o\frac{\partial\vec{E}}{\partial t} \tag{66}$$

$$\vec{F} = q(\vec{E} + \vec{v}\times\vec{B}) \tag{67}$$

Suppose you have made contact with the Quebots living in an alternate Universe. You have also been told that this Universe might consist entirely of particles with opposite signs from ours (their electrons are positively charged, protons negatively, etc). Suppose the only way you have of making contact with the Quebots is by sharing your fundamental laws of Physics. The question is, could you tell by sharing your laws of electrodynamics that a parallel Universe consists of oppositely charged particles? To find out, we ask how our equations change when q is replaced by $-q$. We have

$$\vec{\nabla}\cdot\vec{E} = \frac{-\rho}{\varepsilon_o} \tag{68}$$

$$\vec{\nabla}\times\vec{E} = -\frac{\partial\vec{B}}{\partial t} \tag{69}$$

$$\vec{\nabla}\cdot\vec{B} = 0 \tag{70}$$

$$\vec{\nabla}\times\vec{B} = -\mu_o\vec{J} + \mu_o\varepsilon_o\frac{\partial\vec{E}}{\partial t} \tag{71}$$

$$\vec{F} = -q(\vec{E} + \vec{v}\times\vec{B}) \tag{72}$$

At first glance it seems that Maxwell's equations and the Lorentz force law are now very different. But we must be careful. We can compare our laws of physics with each other, but not our fields. What if we simply let $\vec{E} = -\vec{E}$, and $\vec{B} = -\vec{B}$. This would mean what we call $\vec{E}$ the Quebots call $-\vec{E}$, but we have no way of knowing for sure. You can check the above five laws of electrodynamics one at a time, but in the end, you would end up with exactly the same set of equations.

What if we are told that in the other Universe, everything is a mirror image of our world. That is $\vec{r} \to -\vec{r}$ when you enter the Quebot world. Maxwell's equations would read

$$
\begin{aligned}
-\vec{\nabla}\cdot\vec{E} &= \rho/\varepsilon_o && (73)\\
-\vec{\nabla}\times\vec{E} &= \frac{\partial\vec{B}}{\partial t} && (74)\\
-\vec{\nabla}\cdot\vec{B} &= 0 && (75)\\
-\vec{\nabla}\times\vec{B} &= -\mu_o\vec{J}+\mu_o\varepsilon_o\frac{\partial\vec{E}}{\partial t} && (76)\\
\vec{F} &= q(\vec{E}-\vec{v}\times\vec{B}) && (77)
\end{aligned}
$$

(Here the charge per unit volume ρ does not change sign, but $\vec{r}$, $\vec{v}$ and $\vec{J}$ do. If all of this is getting confusing, see the attached table from Jackson's famous E&M book.) It is easy to see that, by letting $\vec{E}\to-\vec{E}$, and $\vec{F}\to-\vec{F}$, we return to the same set of equations. You may be saying "Doesn't changing $\vec{F}$ to $-\vec{F}$ change Newton's laws?", but $\vec{F}=\frac{d\vec{p}}{dt}$ and for either the relativistic or classical definition of momentum, we expect $\vec{r}$ going to $-\vec{r}$ to change the sign of $\vec{p}$.

The same game can be played with time: One could be told the Quebots live in a world where everything happens in reverse. As strange as this sounds, Maxwells equations could not be used to tell the forward-moving Universe from the backward moving one. Before you conclude that this game is silly, consider the fact that there are fundamental Physics theories (The theory of the electroweak interaction that led to Lee & Wang's Nobel) that can tell a left-handed universe from right a right-handed one.

3.2 Transformation of Maxwell's equations under a Galilean transformation

Now let us see if we could use are fundamental theories to see if the Quebots where moving in a reference frame that is not stationary with respect to ours. In a classical world, this is accomplished by

$$
\begin{aligned}
t' &= t && (78)\\
\vec{r}' &= \vec{r}-\vec{v}t && (79)
\end{aligned}
$$

The chain rule leads to the expressions

$$
\begin{aligned}
\vec{\nabla} &= \vec{\nabla}' && (80)\\
\frac{\partial}{\partial t} &= \frac{\partial}{\partial t'}-\vec{v}\cdot\vec{\nabla}' && (81)
\end{aligned}
$$

and, with a wee bit of work, Maxwell's equations reduce to

$$\begin{aligned}
\vec{\nabla}' \cdot \vec{E} &= \rho'/\varepsilon_o && (82)\\
\vec{\nabla}' \cdot \vec{B} &= 0 && (83)\\
\vec{\nabla}' \times (\vec{E} + \vec{v} \times \vec{B}) &= -\frac{\partial \vec{B}}{\partial t} && (84)\\
\vec{\nabla}' \times (\vec{B} - \frac{1}{c^2}\vec{v} \times \vec{E}) &= \mu_o \vec{J'} + \frac{1}{c^2}\frac{\partial \vec{E}}{\partial t} && (85)
\end{aligned}$$

(Here I have made the transformation $\vec{J'} = \vec{J} + \rho\vec{v}, \rho' = \rho$ because the transformation will create a current where there was none before.) Try as you like, there is no redefinition of $\vec{E}$ and $\vec{B}$ that can lead to something that looks like Maxwell's equations: (You could fix up one of the last two equations, but that would spoil the other first three.) This may not seem like a problem. If the Quebot's are moving with respect to us, then they won't write down Maxwell's equations, but instead they will write down an equation with four fundamental constants, namely c and the components of $\vec{v}$. But this is highly problematic. First of all, as Michelson & Morley demonstrated, we can move into another frame of reference and we do not observe a change in Maxwell's equations. Secondly, it would seem like a wild coincidence that we would be deemed to live in a solar system that moved exactly at $\vec{v}$. The fix is that Newtonian mechanics is wrong and an inertial change of coordinates does not lead to $t' = t, \vec{r'} = \vec{r} - \vec{v}t$, but something else.

3.3 Linear Transformations that do not change of Maxwell's equations: THE LORENTZ TRANFORMATION

In the last section we learned that Maxwell's equations are altered by a Galilean transformation. Here we seek to find transformation matrix Λ does not change Maxwell's equations. Namely what 4x4 matrix Λ leads to a transformation

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \Lambda \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{86}$$

Such that, as strange as the new form of Maxwells equations is, there is some redefinition of E&M fields that gives us back Maxwell's equations.

At this point you might ask why a constant matrix. For example, why can't $t' = x'^2$? The reason is as follows. Suppose we have any vector field $\vec{D}$ that comes with a scalar field D_o to create the conservation law

$$\vec{\nabla} \cdot \vec{D} + \frac{\partial}{c\partial t} D_o = 0 \tag{87}$$

Note the speed of light does not have to have any physical significance at this point. We are just introducing a speed scale so that D_o and $\vec{D}$ have the same units. Two important examples are conservation of charge and energy. For conservation of charge, we have

$$\vec{\nabla} \cdot \vec{J} + \frac{\partial \rho}{\partial t} = 0$$

Letting $\vec{D} = \vec{J}$ and $D_o = \rho c$, we have the stated form of the conservation law. For conservation of energy density we had

$$\vec{\nabla} \cdot \vec{p} + \frac{1}{c^2}\frac{\partial u}{\partial t} = 0$$

which can be written in the form of 87 provided we let

$$\vec{D} = \vec{p}, \quad D_o = u/c$$

Note that the $\vec{D}$ and D_0 together form a dimension four vector which may be written

$$D^u = \begin{matrix} D_o \\ D_1 \\ D_2 \\ D_3 \end{matrix}$$

where we use the subscript 1,2,3 for xyz components of $\vec{D}$. In this way

$$\begin{aligned} J^u &= \begin{matrix} \rho c \\ J_x \\ J_y \\ J_z \end{matrix} \\ p^u &= \begin{matrix} u/c \\ p_x \\ p_y \\ p_z \end{matrix} \end{aligned}$$

Now let us assume

$$\begin{aligned} \begin{pmatrix} x'_o \\ x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} &= \Lambda \begin{pmatrix} x_o \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= x'^u \end{aligned}$$

Then the conservation law is given by

$$\begin{aligned}
\vec{\nabla}\cdot\vec{D}+\frac{1}{c}\frac{\partial D_o}{\partial t} &= \sum_{v}\frac{\partial D^v}{\partial x^v} \\
&= \sum_{u,v}\frac{\partial x'^u}{\partial x^v}\frac{\partial D^v}{\partial x'^u} \\
&= \sum_{u,v}(\Lambda^{-1})^{\mu\nu}\frac{\partial D^v}{\partial x'^u} \\
&= \sum_{u,v}\frac{\partial((\Lambda^{-1})^{\mu\nu}D^v)}{\partial x'^u} \\
&= \sum_{u,v}\frac{\partial(D'^{\mu})}{\partial x'^u}
\end{aligned}$$

Here we have decided that the four vector transforms in the exact way that the coordinates do. This is not so unusual in physics.

We would expect that in any frame of reverence, conserved quantities should remain conserved.

To continue, we write down Maxwell's equations in terms of the potentials V and $\vec{A}$. Provided the Lorentz Gauge conditions

$$\vec{\nabla}\cdot\vec{A}+\frac{1}{c^2}\frac{\partial V}{\partial t}=0 \tag{88}$$

Then Maxwell's equations are simply

$$\Box^2\begin{pmatrix} V/c \\ A_x \\ A_y \\ A_z \end{pmatrix} = -\mu_o\begin{pmatrix} \rho c \\ J_x \\ J_y \\ J_z \end{pmatrix} \tag{89}$$

But the Lorentz gauge condition looks just like a conservation law. Thus, provided Λ is a constant and invertible matrix, if we transform A^μ as a four-vector (i.e., just like we tranformed the coordinates) we have

$$\begin{pmatrix} V'/c \\ A'_x \\ A'_y \\ A'_z \end{pmatrix} = \Lambda\begin{pmatrix} V/c \\ A_x \\ A_y \\ A_z \end{pmatrix} \tag{90}$$

and

$$\vec{\nabla}'\cdot\vec{A'}+\frac{1}{c^2}\frac{\partial V'}{\partial t'}=0. \tag{91}$$

Similarly, we will still have charge conservation

$$\vec{\nabla}'\cdot\vec{J'}+\frac{\partial\rho'}{\partial t'}=0 \tag{92}$$

provided we let

$$\begin{pmatrix} \rho' c \\ J'_x \\ J'_y \\ J'_z \end{pmatrix} = \Lambda \begin{pmatrix} \rho c \\ J_x \\ J_y \\ J_z \end{pmatrix} \tag{93}$$

This leads to the following idea. Why don't we let the new fields be defined as

$$\vec{E}' = -\vec{\nabla}' V' - \frac{\partial \vec{A}'}{\partial t'} \tag{94}$$

$$\vec{B}' = \vec{\nabla}' \times \vec{A}' \tag{95}$$

Then all of Maxwell's equations will continue to hold true (because the gauge condition has not changed) provided

$$\Box'^2 \begin{pmatrix} V'/c \\ A'_x \\ A'_y \\ A'_z \end{pmatrix} = -\mu_o \begin{pmatrix} \rho' c \\ J'_x \\ J'_y \\ J'_z \end{pmatrix} \tag{96}$$

We start with

$$\Box^2 \begin{pmatrix} V/c \\ A_x \\ A_y \\ A_z \end{pmatrix} = -\mu_o \begin{pmatrix} \rho c \\ J_x \\ J_y \\ J_z \end{pmatrix} \tag{97}$$

Transforming both sides:

$$\Lambda \Box^2 \begin{pmatrix} V/c \\ A_x \\ A_y \\ A_z \end{pmatrix} = -\Lambda \mu_o \begin{pmatrix} \rho c \\ J_x \\ J_y \\ J_z \end{pmatrix} \tag{98}$$

And writing the four potential in terms of the inverse of the matrix Λ :

$$\Lambda \Box^2 \Lambda^{-1} \begin{pmatrix} V'/c \\ A'_x \\ A'_y \\ A'_z \end{pmatrix} = -\mu_o \begin{pmatrix} \rho c' \\ J'_x \\ J'_y \\ J'_z \end{pmatrix} \tag{99}$$

So we have found our necessary condition. It is

$$\Lambda \Box^2 \Lambda^{-1} = \Box'^2 \tag{100}$$

But things are even easier than that because Λ is a constant matrix. The condition is simply

$$\Box^2 = \Box'^2 \tag{101}$$

What are the constraints on Λ so the d'Alembertian $\Box^2$ is invariant? To see this we note that the transformation equation

$$\left(\begin{array}{c} ct' \\ x' \\ y' \\ z' \end{array}\right) = \Lambda \left(\begin{array}{c} ct \\ x \\ y \\ z \end{array}\right) \tag{102}$$

can be re-written as

$$x'^{\nu} = \sum_{\mu=0}^{3} \Lambda^{\nu}_{\mu} x^{\mu} \tag{103}$$

Here $x^o = ct$, $x^1 = x, x^2 = y$, and $x^3 = z$. the matrix Λ has been given the indices Λ^{ν}_{μ} rather than the usual $\Lambda_{\nu\mu}$. This is part of the Einstein Notation. Eventually we will assume that the summation in $\Lambda^{\nu}_{\mu} x^{\mu}$ is assumed because the lower index μ is the same as the upper index μ. Their is more to the Einstein Notation which will be covered shortly. For now, let us continue with the derivation on hand. From the transformation equation, we recognize

$$\frac{\partial x'^{\nu}}{\partial x^{\mu}} = \Lambda^{\nu}_{\mu} \tag{104}$$

So the chain rule for derivatives leads to

$$\frac{\partial}{\partial x^{\mu}} = \sum_{\alpha=0}^{3} \Lambda^{\alpha}_{\mu} \frac{\partial}{\partial x'^{\alpha}} \tag{105}$$

And the second-derivative becomes

$$\frac{\partial^2}{\partial x^{\mu\,2}} = \sum_{\alpha,\beta=0}^{3} \Lambda^{\alpha}_{\mu} \Lambda^{\beta}_{\mu} \frac{\partial^2}{\partial x'^{\alpha} \partial x'^{\beta}} \tag{106}$$

Finally, the d'Alembertian $\Box^2 = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ becomes

$$\Box^2 = \sum_{\mu} \sum_{\alpha,\beta=0}^{3} g^{\mu\mu} \Lambda^{\alpha}_{\mu} \Lambda^{\beta}_{\mu} \frac{\partial^2}{\partial x'^{\alpha} \partial x'^{\beta}} \tag{107}$$

Here $g^{\mu\mu'}$ are the components of a diagonal matrix called the Minkowski metric. It is defined so that $g^{\mu m'} = -1$ for $\mu = \mu' = 0$,$g^{\mu\mu'} = 1$ for $\mu = \mu' = 0$, and $g^{\mu\mu'} = 0$ for $\mu \neq \mu'$. Explicitly,

$$g = \left(\begin{array}{cccc} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \tag{108}$$

Because g is diagonal, there is no harm in writing the sum over a second index μ'. Doing this we get

$$\Box^2 = \sum_{\mu,\mu'=0}^{3} \sum_{\alpha,\beta=0}^{3} g^{\mu\mu'} \Lambda_{\mu}^{\alpha} \Lambda_{\mu'}^{\beta} \frac{\partial^2}{\partial x'^{\alpha} \partial x'^{\beta}} \tag{109}$$

$$= \sum_{\alpha,\beta=0}^{3} (\Lambda^T g \Lambda)^{\alpha\beta} \frac{\partial^2}{\partial x'^{\alpha} \partial x'^{\beta}} \tag{110}$$

Here $(\Lambda^T g\Lambda)^{\alpha\beta}$ are the α, β components of the matrix $\Lambda^T g\Lambda$. We see that for the d'Alembertian not to change, we require

$$\Lambda^T g \Lambda = g \tag{111}$$

What does this imply? To see this, let us consider the scalar product

$$(ct, x, y, z)\, g \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = (-ct, x, y, z) \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{112}$$

$$= -c^2t^2 + x^2 + y^2 + z^2 \tag{113}$$

But apparently, if the d'Alembertian is not changed, this quantity can also be written

$$(ct, x, y, z)\, \Lambda^T g \Lambda \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \left[\Lambda \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \right]^T g \left[\Lambda \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \right] \tag{114}$$

$$= (ct', x', y', z') g \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} \tag{115}$$

$$= -c^2t'^2 + x'^2 + y'^2 + z'^2 \tag{116}$$

Apparently the transformations that leave the d'Alembertian unchanged are those that conserve the quantity

$$-c^2t^2 + x^2 + y^2 + z^2 \tag{117}$$

These transformations are known as Lorentz transformations. This is very similar to the result that the transformations that leave the Laplacian $\bigtriangledown^2$ unchanged are those that do not change $x^2 + y^2 + z^2$. In the following sections we go over the Lorentz transformations. But first, a bit of notation.

Einstein Notation Einstein notation gives a convenient way of dealing with the four-dimensional space we live in. The rules are as follows.

(1) A contravariant four vector is a column vector that transforms under the Lorentz transformation 86. Some examples are

$$x^\mu = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{118}$$

$$A^\mu = \begin{pmatrix} V/c \\ A_x \\ A_y \\ A_z \end{pmatrix} \tag{119}$$

$$J^\mu = \begin{pmatrix} \rho c \\ J_x \\ J_y \\ J_z \end{pmatrix} \tag{120}$$

(2) For every contravariant vector x^u, a covariant vector x_u can be formed using the metric:

$$x_\mu = \left[g \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \right]^T \tag{121}$$

$$= (-ct, x, y, z) \tag{122}$$

$$A_u = (-V/c, A_x, A_y, A_z) \tag{123}$$

$$J_u = (-\rho c, J_x, J_y, J_z) \tag{124}$$

(3) When a four vector or tensor is written as a product, the summation over like contravariant and covariant indices are assumed. For example

$$x^\mu x_\mu = -c^2 t^2 + x^2 + y^2 + z^2 \tag{125}$$

(4)The Lorentz transformation of a contravariant or covariant vector is written using the notation Λ^ν_μ. For example, the Lorentz transform of the spatial coordinates can be written.

$$x'^\nu = \Lambda^\nu_\mu x^\mu \tag{126}$$

Note the summation over μ is assumed.

(5) Indices of a Tensor can be flipped using the metric $g_{\mu\mu'} = g^{\mu\mu'}$. For example

$$\Lambda^\mu_\nu = g_{\nu\nu'} g^{\mu\mu'} \Lambda^{\nu'}_{\mu'} \tag{127}$$

3.3.1 The Lorentz Transformations

Translations A translation is defined by adding a constant to each coordinate. Physically it corresponds to resetting the origin and time $t = 0$ of the coordinate system.

$$\begin{pmatrix} x'^o \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} x^o \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} + \begin{pmatrix} d^o \\ d^1 \\ d^2 \\ d^3 \end{pmatrix} \tag{128}$$

Here d^u are constants and we have used the convention

$$x^o = ct,\ x^1 = x,\ x^2 = y,\ x^3 = z \tag{129}$$

We can write Eq 128 in more compact notation

$$x'^u = x^u + d^u \tag{130}$$

It is easy to see that the addition of a constant does not change the form of any of the partial derivatives. Thus the d'Alembertian is not altered.

Rotations Because the d;Alembertian may be written

$$\Box^2 = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2 \tag{131}$$

it is easy to see that any change of coordinate system that does not change ∇^2 will not change $\Box^2$. Thus rotations (and inversions) of the $x - y - z$ coordinate system will not change the d;Alembertian. The most general rotation consists of a rotation through the three Euler angles θ, ϕ and χ. These rotations consist of a rotation through an angle ϕ about the z axis, followed by a rotation of an angle θ about the newly created y axis, followed by a rotation χ about the new z axis:

$$\begin{pmatrix} x'^o \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c\chi & s\chi & 0 \\ 0 & -s\chi & c\chi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{132}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c\theta & 0 & -s\theta \\ 0 & 0 & 1 & 0 \\ 0 & s\theta & 0 & c\theta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c\phi & s\phi & 0 \\ 0 & -s\phi & c\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^o \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{133}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c\phi c\theta c\chi - s\phi s\chi & -c\phi c\theta s\chi - s\phi c\chi & c\phi s\theta \\ 0 & s\phi c\theta c\chi + c\phi s\chi & -s\phi c\theta s\chi + c\phi s\chi & s\phi s\theta \\ 0 & -s\theta c\chi & s\theta s\chi & c\theta \end{pmatrix} \begin{pmatrix} x^o \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{134}$$

This rotation will take any right-handed coordinate system to all other right handed coordinate systems that share the same origin. Notice that the rows of each of the rotation matrices are orthonormal, as required for the Laplacian not to change. It is hardly outside our intuition that the laws of physics should remain the same in two frames related by a rotation of the spatial coordinates.

Boosts What is comforting about rotations is that $x^o = x'^o$ for all time. this implies $t = t'$. For translations, it is possible that $t = t' + t_o$, but this is not so disturbing to our intuition. If someone resets their watch so that zero-time does not agree with another's, we do not expect the laws of physics to change. The boosts we are about to see mix up the space at time coordinates. This is highly non-intuitive stuff. Consider the transformation

$$\begin{pmatrix} x'^o \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^o \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{135}$$

where

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{136}$$

This transformation corresponds to the new coordinates of an observer traveling at a velocity $v = \beta c$ along the x axis. Let us consider two points A and B in the unprimed coordinate system. The invariant interval is

$$\Delta x_u \Delta x^u = -c^2(t_A - t_B)^2 + (x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2 \tag{137}$$

$$\Delta x_u = x_{Au} - x_{Bu} \tag{138}$$

$$\Delta x^u = x^u_A - x^u_B \tag{139}$$

Here we have used a special notation: Namely that

$$x_o = -x^o = -ct \tag{140}$$

$$x_1 = x^1 = x \tag{141}$$

$$x_2 = x^2 = y \tag{142}$$

$$x_3 = x^3 = z \tag{143}$$

The vectors with the superscript are call contravariant vectors. The vectors with the subscript are called covariant vectors. When a contravarient and covariant vector are written together with the same index, a summation is assumed:

$$\Delta x_u \Delta x^u = \sum_{u=0}^{3} \Delta x_u \Delta x^u \tag{144}$$

Now let us look at the invariant interval for the boosted coordinate system:

$$\Delta x'_u \Delta x'^u = -c^2(t'_A - t'_B)^2 + (x'_A - x'_B)^2 + (y'_A - y'_B)^2 + (z'_A - z'_B)^2 \tag{145}$$

$$= -(\gamma c t_A - \gamma\beta x_A - \gamma c t_B + \gamma\beta x_B)^2 + \tag{146}$$

$$(\gamma x_A - \beta\gamma c t_A - \gamma x_B + \beta\gamma c t_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2 \tag{147}$$

$$= -\gamma^2 \left[c(t_A - t_B) - \beta(x_A - x_B)\right]^2 + \tag{148}$$

$$\gamma^2 \left[(x_A - x_B) - \beta c(t_A - t_B)\right]^2 + (y_A - y_B)^2 + (z_A - z_B)^2 \tag{149}$$

$$= -\gamma^2 c^2 [(t_A - t_B)^2 - \beta^2 (t_A - t_B)^2] + \tag{150}$$

$$\gamma^2 \left[(x_A - x_B)^2 - \beta^2 (x_A - x_B)^2\right] + \text{(cancelling cross terms)} + \tag{151}$$

$$(y_A - y_B)^2 + (z_A - z_B)^2 \tag{152}$$

$$= -\gamma^2 c^2 (1 - \beta^2)(t_A - t_B)^2 + \tag{153}$$

$$\gamma^2 (1 - \beta^2)(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2 \tag{154}$$

$$= -c^2 (t_A - t_B)^2 + (x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2 \tag{155}$$

$$= \Delta x_u \Delta x^u \tag{156}$$

Here we see that the boost does not change the invariant length. Not only does the boost and rotations and translation leave the invariant length invariant, but also the four-dimensional scalar product

$$a_u b^u = a'_u b'^u \tag{157}$$

Of course we can consider boosts in any direction, not just the x direction. I leave investigation of boosts in arbitrary directions as an exercise for the reader.

Lorentz transformations for small β: One important test of special relativity is that it is consistent with classical mechanics. The postulate of Einstein led to Lorentz transformations connected intertial reverence frames rather than Galilean transformations. Let us make sure this happens first we write the boost transformation with c shown explicitly

$$\begin{pmatrix} ct' \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1-(v/c)^2}} & \frac{-v/c}{\sqrt{1-(v/c)^2}} & 0 & 0 \\ \frac{-v/c}{\sqrt{1-(v/c)^2}} & \frac{1}{\sqrt{1-(v/c)^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{158}$$

The new time becomes

$$ct' = \frac{ct}{\sqrt{1-(v/c)^2}} - \frac{v/c}{\sqrt{1-(v/c)^2}} x \tag{159}$$

$$\approx ct - \frac{vx}{c} \tag{160}$$

$$t' \approx t - \frac{vx}{c^2} \approx t \tag{161}$$

The new x coordinate becomes

$$x' = \frac{-vt}{\sqrt{1-(v/c)^2}} + \frac{x}{\sqrt{1-(v/c)^2}} \tag{162}$$

$$\approx x - vt \tag{163}$$

It is easy to see why classical mechanics works so well. At low velocities, the speed of light drops out and we have the relationship we would expect for changing to a reference frame that moves with a velocity v, namely $x' = x - vt$. The relationship $t' = t$ is so "obvious" in classical mechanics that it was not even written down.

Addition of velocities Suppose an object moves with a velocity v along the $x - axis$ as viewed in a stationary reference frame. What is the speed $v'$of the object as recorded by an observer moving at a velocity v_o in the $-x$ direction? To answer this question, we note the trajectory of the particle, observed from the stationary frame, has the space time coordinates. Now let us record the coordinate a time t latter. Its new coordinate will be

$$x^u = \begin{pmatrix} ct \\ vt \\ 0 \\ 0 \end{pmatrix} \tag{164}$$

Now a boost of velocity v_o will take us to the coordinates of the new point, as viewed by the moving observer:

$$\begin{pmatrix} x'^o \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma_o & \gamma_o\beta_o & 0 & 0 \\ \gamma_o\beta_o & \gamma_o & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ vt \\ 0 \\ 0 \end{pmatrix} \tag{165}$$

$$= \begin{pmatrix} \gamma_o ct + \gamma_o\beta_o vt \\ \gamma_o\beta_o ct + \gamma_o vt \\ 0 \\ 0 \end{pmatrix} \tag{166}$$

$$= \begin{pmatrix} (\gamma_o + \gamma_o\beta_o\beta)ct \\ (\gamma_o\beta_o + \gamma_o\beta)ct \\ 0 \\ 0 \end{pmatrix} \tag{167}$$

The moving observer will record a speed

$$
\begin{aligned}
v' &= \frac{x^1}{x^o} c && (168)\\
&= \frac{(\gamma_o \beta_o + \gamma_o \beta)}{(\gamma_o + \gamma_o \beta_o \beta)} c && (169)\\
&= \frac{v_o + v}{1 + \beta_o \beta} && (170)\\
&= \frac{v_o + v}{1 + (v_o v)/c^2} && (171)
\end{aligned}
$$

Notice that if $v = c$, $v' = c$. In fact, this is precisely the form of v' we obtained by fooling around with possible velocity addition rules. Here we have come across the addition rule by considering those transformations that leave the d'Alembertian (and hence Maxwell's equations) unaltered.

4 The physical meaning of the invariant interval & causality

The invariant interval between two points

$\Delta x_u \Delta x^u = -c^2 (t_A - t_B)^2 + (x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2$

has been shown to be independent of reference frame. Just as the distance between two points, the invariant quantity in classical mechanics, has an intuitive physical meaning, one might expect $\Delta x_u \Delta x^u$ to have physical meaning. Unlike the distance between two points, $\Delta x_u \Delta x^u$ can be negative. Let us investigate the meaning of the sign of the invariant interval:

TIME-LIKE INTERVALS: $\Delta x_u \Delta x^u < 0$: Two events that happen at the same exact time can not have a negative invariant interval. On the other hand, two events that occur at the same point in space but different times can be time-like. If two events are time-like, there exists an inertial system in which they occur at the same point. Suppose event A takes place at $\{0, 0, 0, 0\}$ and B at $\{t_B, x_B, 0, 0\}$ with $x_B, t_B > 0$. The events are separated by a time-like interval if $-c^2 t_B^2 + x_B^2 < 0$, or $c > x_B / t_B$. Now imagine a reference frame (say a train) that travels from A to B at a velocity $v = x_B / t$. Since $v < c$, this is possible. Both events will obviously occur at the spatial origin of the new coordinate system, but at different times. Explicitly we have $x'^u_A = 0$ and

$$\begin{pmatrix} x'^o_B \\ x'^1_B \\ x'^2_B \\ x'^3_B \end{pmatrix} = \begin{pmatrix} 1/\sqrt{1-(v/c)^2} & -(v/c)/\sqrt{1-(v/c)^2} & 0 & 0 \\ -(v/c)/\sqrt{1-(v/c)^2} & 1/\sqrt{1-(v/c)^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t_B\, c \\ v t_B \\ 0 \\ 0 \end{pmatrix} \tag{172}$$

$$= \begin{pmatrix} (t_B\, c)\sqrt{1-(v/c)^2} \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{173}$$

Thus the time interval viewed from the moving frame will be $(t_B\, c)\sqrt{1-(v/c)^2}$, but both events will happen at the same point.

SPACE-LIKE INTERVALS: $\Delta x_u \Delta x^u > 0$: Two events that happen at the same point can not have a positive invariant interval. On the other hand, two events that happen at the same time but at different points are space like. If two events are space-like, there exists an inertial system in which they occur at the same time, but different points. Suppose event A takes place at $\{0,0,0,0\}$ and B at $\{t_B, x_B, 0, 0\}$ with $x_B, t_B > 0$. The events are separated by a space-like interval if $-c^2 t_B^2 + x_B^2 > 0$, or $c < x_B/t_B$. A boosts relates the $x - ct$ coordinates by

$$\begin{pmatrix} x'^o_B \\ x'^1_B \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} ct_B \\ x_B \end{pmatrix} \tag{174}$$

If A and B are to happen at the same time, but a different points, $x'^o_B = x'^o_A = 0$

$$\begin{pmatrix} 0 \\ x'^o_B \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} ct_B \\ x_B \end{pmatrix} \tag{175}$$

This will occur when

$$\gamma c t_B - \gamma x_B \beta = 0 \tag{176}$$

$$1 - \frac{x_B \beta}{c t_B} = 0 \tag{177}$$

$$\beta = \frac{c t_B}{x_B} \tag{178}$$

Thus a reference frame moving with a speed $c^2/(x_B/t_B)$ will observe both events at the same time, but separated by a distance given by the Lorentz transformation

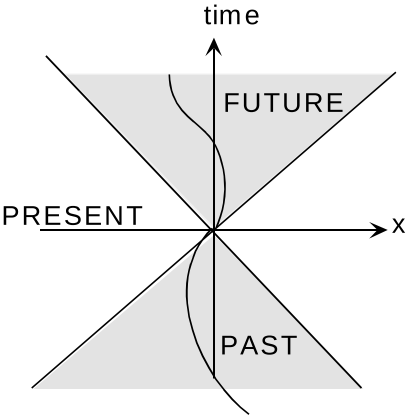
$$\begin{pmatrix} 0 \\ x'^o_B \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} ct_B \\ x_B \end{pmatrix} \tag{179}$$

$$= \begin{pmatrix} 0 \\ x_B\sqrt{-(\frac{ct_B}{x_B})^2 + 1} \end{pmatrix} \tag{180}$$

Because nothing can travel faster than the speed of light, two events occurring at the same time but different locations can not influence each other. Thus events separated by space-like intervals are completely isolated.

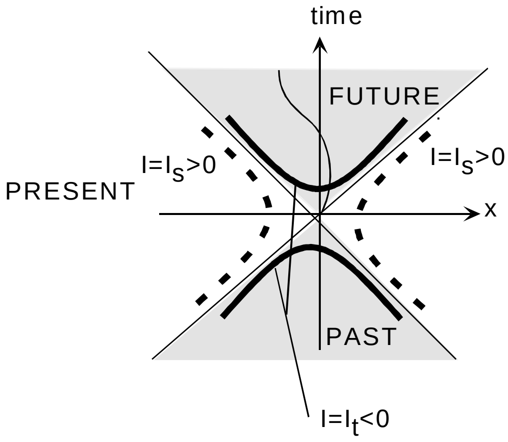
LIGHT-LIKE INTERVALS:: $\Delta x_u \Delta x^u = 0$ When two events are connected by a signal traveling at the speed of light, the events are called light-like and $\Delta x_u \Delta x^u = 0$.

Minkowski diagrams Minkowski diagrams are simply plots of the position of a particle as a function of time, but with time given the y axis. Imagine the particle is me and I cross the origin of this plot at $t = 0$. The only possible points that I could have come from (without moving faster than the speed of light) are those within the 45^o lines that define motion at the speed of light. Thus my trajectory must come from the cone behind me. In addition, I can only move into the cone in front of me (unless I can exceed the universal speed limit c.)



(181)

Now let us consider all points connected by a Lorentz transformation. If we consider all x-y points connected by a rotation, we get a circle. But if we consider all t-x points connected by a boost, we get hyperbolas of constant invariant intervals $x^2 - c^2t^2$



(182)

For the case of space-like intervals, we get one smooth surface connecting all points with constant $x^u x_u$. (It looks like two in the above diagram, but if you imagine rotating this picture so as to include a y dimension, then you would make a single surface. Thus, by an appropriate Lorentz transformation, an event occurring at $t > 0$ could be boosted to an event occurring at $t < 0$, however this event must be separated from you by a space-like interval and therefore will be impossible to observe. On the other hand, events in your past or future (connected to your present by a time-like interval) that share the same $x^u x_u$ create two surfaces, one occurring at $t > 0$, and the other at $t < 0$. No amount of Lorentz transforming will take a point from the upper surface and bring it to the lower surface, even though $x^u x_u$ is constant. Thus the time ordering is unique. This is very important because it gives us the notation of causality: If in one frame we observe A caused B, in all other frames that can witness the event, A happens before event B. (Of course, we could perform a simple change of variables $t = -t'$. This change is not considered to be a Lorentz transformation, even though it does conserve the d'Alembertian. Lorentz transformations are exclusively boosts, rotations, and translations.)

5 Relativistic Mechanics

5.1 Proper time and proper velocity

Let us suppose we observe a person traveling from point $A = \{0, 0, 0, 0\}$ to point $B = \{ct_B, ut_B, 0, 0\}$ where v is the constant speed we observe the traveler moving at. What are we to conclude about how the person has aged during his trip? We will assume his internal clock is running slow and that he has aged

$$\tau = t_B \sqrt{1 - (u/c)^2} \tag{183}$$

and that the distance the person travels per time aged is given by

$$\begin{aligned} \eta &= \frac{ut_B}{t_B\sqrt{1-(u/c)^2}} && (184) \\ &= \frac{u}{\sqrt{1-(u/c)^2}} && (185) \end{aligned}$$

Here τ is called the persons proper time and η the proper velocity.

Note that no matter who observes an object, he or she will make the same conclusion about the amount the object has aged between to events. Thus the proper time is invariant (frame independent.) To see this, let us boost the points A and B to the frame of reference of our traveler:

$$\begin{aligned} \begin{pmatrix} x'^o_B \\ x'^1_B \end{pmatrix} &= \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} ct_B \\ \beta ct_B \end{pmatrix} && (186) \\ &= \begin{pmatrix} \gamma ct_B(1-\beta^2) \\ 0 \end{pmatrix} && (187) \\ &= \begin{pmatrix} ct_B\sqrt{1-(u/c)^2} \\ 0 \end{pmatrix} && (188) \end{aligned}$$

Thus the wrist watch of the traveler will change by an amount

$$\tau = t_B\sqrt{1-(u/c)^2}. \qquad (189a)$$

The fact that τ is invariant implies that the proper velocity of an object will transform exactly like the distance between two points between to objects: Thus the same Lorentz transformation that lets us know how to measure space-time coordinates in a new frame of reference tells us how to measure an objects proper velocity in a new frame of reference. The proper velocity of an object is formally defined as the change in position of an object per unit proper time, or per unit of aging:

$$\eta^u = \frac{dx^u}{d\tau} \qquad (190)$$

Here the zeroth component is defined as the change in time per proper time and has no classical analog

$$\eta^o = \frac{dx^o}{d\tau} = \frac{c}{\sqrt{1-u^2/c^2}} \qquad (191)$$

The transformation of proper velocity is, as we indicated, simple:

$$\begin{pmatrix} \eta'^o \\ \eta'^1 \\ \eta'^2 \\ \eta'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \eta^o \\ \eta^1 \\ \eta^2 \\ \eta^3 \end{pmatrix} \qquad (192)$$

5.2 Relativistic Energy and Momentum

In classical mechanics, $\vec{p} = m\vec{v}$. In a classical world, there is no difference between proper velocity and ordinary velocity, so there is no reason to suspect that $m\vec{v}$ would have more or less significance than $m\vec{\eta}$. As it turns out, the relativistic momentum $m\vec{\eta}$ is of much greater physical significance. We define the four momentum

$$p^u = m\eta^u \tag{193}$$

For a particle traveling at a constant velocity, this becomes the four vector

$$p^u = \begin{pmatrix} \frac{mc}{\sqrt{1-u^2/c^2}} \\ \frac{m\vec{u}}{\sqrt{1-u^2/c^2}} \end{pmatrix} \tag{194}$$

Let us rearrange this a bit: First note that the squared magnitude of the proper momentum $\vec{p}$ is given by

$$\vec{p}^2 = \frac{m^2c^2(u^2/c^2)}{1-u^2/c^2} \tag{195}$$

Solving for $1-u^2/c^2$ we have

$$\begin{aligned} \vec{p}^2(1-u^2/c^2) &= -m^2c^2(1-(u^2/c^2)) + m^2c^2 && (196) \\ (1-u^2/c^2)(\vec{p}^2 + m^2c^2) &= m^2c^2 && (197) \\ 1-u^2/c^2 &= \frac{m^2c^2}{\vec{p}^2 + m^2c^2} && (198) \end{aligned}$$

Solving for the zero'th component of the four-vector, we have

$$\frac{mc}{\sqrt{1-u^2/c^2}} = \sqrt{\vec{p}^2 + m^2c^2} \tag{199}$$

Thus we have a four-momentum given by

$$p^u = \begin{pmatrix} \sqrt{\vec{p}^2 + m^2c^2} \\ \vec{p} \end{pmatrix} \tag{200}$$

Where

$$\vec{p} = \frac{m\vec{u}}{\sqrt{1-u^2/c^2}} \tag{201}$$

What is the first term of this four vector we have created? Einstein told us it was the energy of the particle divided by the speed of light. That is

$$\begin{aligned} E &= p^o c && (202) \\ &= \sqrt{\vec{p}^2c^2 + m^2c^4} && (203) \end{aligned}$$

If an object is at rest with respect to an observe, this reduced to

$$E_{rest} = mc^2 \tag{204}$$

If we assume that $\vec{p}$ is small (so that mc^2 is much greater that pc) then we have

$$\begin{aligned} E &\approx mc^2 + \frac{\vec{p}^2}{2m} && (205) \\ &= mc^2 + E_{classical} && (206) \end{aligned}$$

The the brilliance of this deduction is not in the equations (mostly notational manipulation) but in the experimental fact that, in a closed system, p^u is conserved.

If p^u transforms like four vector, we must have $p_u p^u$ be an invariant quantity. In this case, $p^u p_u = -m^2c^2$ is invariant. Notice that the fact that this is invariant does not imply that the mass of an object can not change to release energy.

6 Relativistic Electrodynamics

6.1 What should be a four vector?

The first four vector we encountered was the space-time coordinate

$$x^u = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{207}$$

and we also constructed a new four-vector by defining the invariant proper time τ where

$$\frac{d\tau}{dt} = \sqrt{1-(u/c)^2} \tag{208}$$

gives the amount an object ages per unit of time in our reference frame. (If we choose a time origin where $\tau = t = 0$ and integrate

$$\tau = \int_o^t \sqrt{1-(u(t')/c)^2} dt' \tag{209}$$

we can find an explicit expression for τ in terms of t. If u is constant this reduces to Eq 189a.) Because τ is the same in all reference frames, a second four vector is the proper velocity η^u of a particle.

$$\eta^u = \frac{dx^u}{d\tau} \tag{210}$$

If we give a particle a mass m, then we can define the proper momentum

$$p^u = m\eta^u. \tag{211}$$

All this is very nice, very formal, and does not contain much physics, until we make the bold statement that p^u is a conserved quantity: in any isolated process, the proper momentum is conserved. This statement has been observed to be true over and over again, both in the relativistic and nonrelativistic limits. Thus we say to ourselves p^u is a useful four vector.

In the next section we introduce new four vectors of observable quantities. Each time may ask ourselves if our new quantity has any physical significance. The proof is always in experiment. Each time we choose a new four vector of an observable, we make very definite predictions about the way objects should behave as their speed approaches the speed of light. If these predictions are observed, we have chosen wisely, if not we must conclude that, although we might have discovered a nice trick, we have not discovered new physics.

.

6.2 The transformation of fields

We have seen that the four-vector potential A^μ transforms just like the coordinates do. That is

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

implies that

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

So, we can easily determine how the field transforms by considering the chain rule for derivatives applied to

$$\begin{aligned} \vec{E}' &= -\vec{\nabla}' V - \frac{\partial \vec{A}'}{\partial t} \\ \vec{B}' &= \vec{\nabla}' \times \vec{A}'. \end{aligned}$$

The algebra gets a bit tricky, and can be avoided by using a clever trick. Notice that all the fields are proportional to first derivatives. We ask ourselves if there is any combination of first derivatives of A^μ that transforms simply. The answer is yes, the following one does:

$$R^{\alpha\beta} = g^{\alpha\mu}\frac{\partial A^\beta}{\partial x^\mu} = -\delta_{\alpha 0}\frac{\partial A^\beta}{\partial ct} + \delta_{\alpha 1}\frac{\partial A^\beta}{\partial x} + \delta_{\alpha 2}\frac{\partial A^\beta}{\partial y} + \delta_{\alpha 3}\frac{\partial A^\beta}{\partial z}$$

Let us write this in terms of the primed coordinates

$$R^{\alpha\beta} = g^{\alpha\mu}(\Lambda^{-1})^\beta_{\beta'}\Lambda^{\mu'}_\mu \frac{\partial A'^{\beta'}}{\partial x'^{\mu'}}$$

But we know that

$$\Lambda^T g \Lambda = g$$

or, in Einstein notation

$$\begin{aligned} \Lambda^\alpha_{\alpha'} g^{\alpha\mu}\Lambda^{\mu'}_\mu &= g^{\alpha'\mu'} \\ g^{\alpha\mu}\Lambda^{\mu'}_\mu &= (\Lambda^{-1})^\alpha_{\alpha'} g^{\alpha'\mu'} \end{aligned}$$

Thus we can make a substitution in our expression for $R^{\alpha\beta}$:

$$\begin{aligned} R^{\alpha\beta} &= (\Lambda^{-1})^{\beta}_{\beta'}(\Lambda^{-1})^{\alpha}_{\alpha'}g^{\alpha'\mu'}\frac{\partial A'^{\beta'}}{\partial x'^{\mu'}} \\ &= (\Lambda^{-1})^{\beta}_{\beta'}(\Lambda^{-1})^{\alpha}_{\alpha'}R'^{\alpha'\beta'} \end{aligned}$$

Thus, to transform the tensor $R^{\alpha\beta}$, one simply applies the Lorentz tranform to each index

$$R'^{\alpha'\beta'} = \Lambda^{\beta'}_{\beta}\Lambda^{\alpha'}_{\alpha}R^{\alpha\beta}$$

We now know how that the transformation of some combinations of derivatives is very easy. But will this help us find the fields? Another words, can we write the fields we want in terms of the $R^{\alpha\beta\prime}s$. The answer is yes, provided we take the *antisymmetric combination*

$$\begin{aligned} F^{\alpha\beta} &= R^{\alpha\beta} - R^{\beta\alpha} \\ &= g^{\alpha\mu}\frac{\partial A^{\beta}}{\partial x^{\mu}} - g^{\beta\mu}\frac{\partial A^{\alpha}}{\partial x^{\mu}} \end{aligned}$$

Note that the diagnoal elements are explicitly zero. (Because $F^{\alpha\beta} = -F^{\beta\alpha}$.) Of the twelve off-diagonal elements of the matrix, we only need six to define the matrix complelely, because given $F^{\alpha\beta}$, we know $F^{\beta\alpha} = -F^{\alpha\beta}$. This is the same number of degrees of freedom in $\vec{E}$ and $\vec{B}$, so perhaps we are on track. Indeed, I leave it to you to show

$$F^{\mu\nu} = \begin{matrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{matrix} \tag{212}$$

This is it! This is the rule for transforming fields. For example,consider a boost in the x direction. Then

$$\Lambda^{s}_{u} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{213}$$

It is easier to see what will happen to the $\vec{E}$ and $\vec{B}$ fields if we write the transformation in matrix form:

$$F' = \Lambda F \Lambda^{T} \tag{214}$$

$$= \Lambda F \Lambda \tag{215}$$

$$= \begin{pmatrix} 0 & E_x/c & \gamma(E_y/c - \beta B_z) & \gamma(E_z/c + \beta B_y) \\ & 0 & \gamma(B_z - \beta E_y/c) & -\gamma(B_y + \beta E_z/c) \\ & & 0 & B_x \\ & & & 0 \end{pmatrix} \tag{216}$$

Here I have left out half of the antisymmetric matrix. Notice when β goes to zero (and γ goes to 1) the field tensor is unaltered. What does this imply? That a boost in the x direction does not change ones perception of either E_x or B_x. However, the y and z components of both the magnetic and electric field change:

$$\begin{aligned} E'_y/c &= \gamma(E_y/c - \beta B_z) && (217) \\ E'_z/c &= \gamma(E_z/c + \beta B_y) && (218) \\ B'_y &= \gamma(B_y + \beta E_z/c) && (219) \\ B'_z &= \gamma(B_z - \beta E_y/c) && (220) \end{aligned}$$

This result can be written in the extremely useful vector form (see Jackson, Classical Electrodynamics) as

$$\begin{aligned} \vec{E}' &= \gamma(\vec{E} + \vec{\beta} \times \vec{B}) - \frac{\gamma^2}{\gamma+1}(\vec{\beta} \cdot \vec{E})\vec{\beta} \\ \vec{B}' &= \gamma(\vec{B} - \vec{\beta} \times \vec{E}) - \frac{\gamma^2}{\gamma+1}(\vec{\beta} \cdot \vec{B})\vec{\beta} \end{aligned}$$

Taking $\vec{\beta} \cdot \vec{E}'$ we get

$$\begin{aligned} \vec{\beta} \cdot \vec{E}' &= (\gamma - \frac{\gamma^2\beta^2}{\gamma+1})(\vec{\beta} \cdot \vec{E}) \\ &= (\gamma - \frac{\gamma^2(1-\frac{1}{\gamma^2})}{\gamma+1})(\vec{\beta} \cdot \vec{E}) \\ &= (\gamma - \frac{(\gamma+1)(\gamma-1)}{\gamma+1})(\vec{\beta} \cdot \vec{E}) = \vec{\beta} \cdot \vec{E} \end{aligned}$$

Thus we see that the components of $\vec{E}$ and $\vec{B}$ parallel to $\vec{\beta}$ are the same in both reference frames.

7 The field of a charge moving at a constant velocity

7.1 fields as determined by a lorentz transformation

Suppose we have a charge stationary at the origin. It's electric and magnetic field are given by

$$\begin{aligned} \vec{E} &= \frac{q}{4\pi\varepsilon_o}\frac{\vec{r}}{r^3} && (221) \\ \vec{B} &= \vec{0} && (222) \end{aligned}$$

What about the field moving with a speed $v = \beta c$ in the x direction? If we introduce the displacement $\vec{R}$ of the particle from a position (x, y, z)

$$\vec{R} = (x, y, z) - (vt, 0, 0) \qquad (223)$$

and the speed v is much less than the speed of light, we expect

$$\vec{E}' \approx \frac{1}{4\pi\varepsilon_o}\frac{\vec{R}}{R^3} \tag{224}$$

To get this right in the general case, we note that if a coordinate frame is moving with the particle, then people in the moving frame will observe

$$\begin{aligned} \vec{E}' &= \frac{q}{4\pi\varepsilon_o}\frac{\vec{r}'}{r'^3} && (225) \\ \vec{B}' &= \vec{0} && (226) \end{aligned}$$

To determine what we will observe, we note our frame is moving at a velocity $-\beta$ with respect to the frame moving with the particle. The field we observe is therefore given by

$$\begin{aligned} E_x &= E'_x && (227) \\ E_y &= \gamma(E'_y + \beta c B'_z) = \gamma E'_y && (228) \\ E_z &= \gamma(E'_z - \beta c B'_y) = \gamma E'_z && (229) \\ B_x &= B'_x = 0 && (230) \\ B_y &= \gamma(B'_y - \beta E'_z/c) = -\gamma\beta E'_z/c && (231) \\ B_z &= \gamma(B'_z + \beta E'_y/c) = \gamma\beta E'_y/c && (232) \end{aligned}$$

Of course we would like to evaluate $\vec{E}$ and $\vec{B}$ in the coordinates of our frame whereas $\vec{E}'$ is defined with respect to frame in which the charge is stationary. To do this, we note the boost is given by

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} \tag{233}$$

The inverse boost is found by changing the sign of β:

$$\begin{aligned} \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} &= \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} && (234) \\ &= \begin{pmatrix} \gamma(ct - \beta x) \\ \gamma(x - \beta ct) \\ y \\ z \end{pmatrix} && (235) \end{aligned}$$

From this transformation, we find that

$$\frac{\vec{r}'}{r'^3} = \frac{(\gamma(x - vt), y, z)}{(\gamma^2(x - vt)^2 + y^2 + z^2)^{3/2}} \tag{236}$$

Here γ is the usual factor of $1/\sqrt{1-(v/c)^2}$. The electric and magnetic fields are then

$$E_x = \frac{q\gamma(x-vt)}{4\pi\varepsilon_o(\gamma^2(x-vt)^2+y^2+z^2)^{3/2}} \tag{237}$$

$$E_y = \frac{q\gamma y}{4\pi\varepsilon_o(\gamma^2(x-vt)^2+y^2+z^2)^{3/2}} \tag{238}$$

$$E_z = \frac{q\gamma z}{4\pi\varepsilon_o(\gamma^2(x-vt)^2+y^2+z^2)^{3/2}} \tag{239}$$

$$B_x = 0 \tag{240}$$

$$B_y = \frac{-vq\gamma z}{4\pi\varepsilon_o c^2(\gamma^2(x-vt)^2+y^2+z^2)^{3/2}} \tag{241}$$

$$B_z = \frac{vq\gamma y}{4\pi\varepsilon_o c^2(\gamma^2(x-vt)^2+y^2+z^2)^{3/2}} \tag{242}$$

This can be written in an very pretty way with the following manipulation:

$$\frac{1}{(\gamma^2(x-vt)^2+y^2+z^2)^{3/2}} = \frac{1}{\gamma^3((x-vt)^2+(1-\beta^2)(y^2+z^2))^{3/2}} \tag{243}$$

$$= \frac{1}{\gamma^3((x-vt)^2+(1-\beta^2)(y^2+z^2))^{3/2}} \tag{244}$$

$$= \frac{1}{\gamma^3(R^2-\beta^2(y^2+z^2))^{3/2}} \tag{245}$$

Now note that $\vec{\beta}\cdot\vec{R} = x\beta$ so that $\beta^2R^2 - (\vec{\beta}\cdot\vec{R})^2 = \beta^2(y^2+z^2)$. Thus we have

$$\beta^2(y^2+z^2) = \beta^2R^2 - (\vec{\beta}\cdot\vec{R})^2 \tag{246}$$

$$= R^2\beta^2(1-(\widehat{\beta}\cdot\widehat{R})^2) \tag{247}$$

Finally we have

$$\frac{1}{(\gamma^2(x-vt)^2+y^2+z^2)^{3/2}} = \frac{1}{\gamma^3R^3(1-\beta^2(1-(\widehat{\beta}\cdot\widehat{R})^2)^{3/2}} \tag{248}$$

With this substitution, the fields become

$$\vec{E} = \left[\frac{q}{4\pi\varepsilon_o}\frac{\vec{R}}{R^3}\right]\frac{1}{\gamma^2(1-\beta^2(1-(\widehat{\beta}\cdot\widehat{R})^2)^{3/2}} \tag{249}$$

$$\vec{B} = \frac{\vec{\beta}\times\vec{E}}{c} \tag{250}$$

$$= \left[\frac{q}{4\pi\varepsilon_o c}\frac{\vec{\beta}\times\vec{R}}{R^3}\right]\frac{1}{\gamma^2(1-\beta^2(1-(\widehat{\beta}\cdot\widehat{R})^2)^{3/2}} \tag{251}$$

Notice that in the limit of a slow moving particle, $\vec{E}$ takes on the expected value and $\vec{B}$ goes to zero.

7.2 fields as determined by the Lienard-Wiechert potentials

If you recall, we have already investigated the field from a moving charge by considering the Lienard-Wiechert potentials:

$$V(\vec{r},t) = \frac{q}{4\pi\varepsilon_o}\left[\frac{1}{|\vec{r}-\vec{r}_i| - (\vec{r}-\vec{r}_i)\cdot\vec{v}/c}\right] \tag{252}$$

and

$$\vec{A} = \vec{v}V/c^2 \tag{253}$$

Here $\vec{r}_i$ is the position of the particle evaluated at the retarded time. For a particle traveling along a straight line, $\vec{r}_i$ is related to $\vec{R}$ by

$$\vec{r}_i = \vec{r}_o + \vec{v}\left[t - \frac{\left[\frac{\vec{R}\cdot\vec{v}}{c}\right] + \sqrt{R^2(1-\beta^2) + \left[\frac{\vec{R}\cdot\vec{v}}{c}\right]^2}}{c(1-\beta^2)}\right]. \tag{254}$$

The electric and magnetic fields can be found by differentiation of V and $\vec{A}$ after making this substitution. (Here $\vec{R} = \vec{r} - (\vec{r}_o + \vec{v}t)$ because we have introduced the possibility that the particle does not start at the origin at $t = 0$. This assumption does not change the form of equations 249-251.) The fields were found in an in-class assignment using Mathematica to compute

$$\vec{E} = -\vec{\nabla}V - \frac{\partial}{\partial t}\vec{A} \tag{255}$$

$$\vec{B} = \vec{\nabla}\times\vec{A} \tag{256}$$

Do these two results match? The answer is yes. I invite you to try to prove it. If you succeed, you will have done a calculation that Mathematica has a great deal of trouble with.

8 Relativity and Forces

So far we have been concerned with the way in which an object's motion or field changes when viewed in different reference frames. We have not yet discussed how forces and Newton's laws are altered by the theory of relativity. That is the concern of this lecture.

We have already taken the first step with momentum conservation. We have introduced the idea that the proper momentum (or just momentum)

$$\vec{p} = m\eta^u = m\frac{dx^u}{d\tau^u} \tag{257}$$

is a four vector because the proper time is an invariant. (Recall that $dx^u/d\tau$ is the change in time and position of a particle with respect to how much it ages.)

When defining a force, classically, one has

$$\vec{F} = \frac{d\vec{p}}{dt}. \tag{258}$$

One might first guess that the way to define force relativistically is by replacing t with the proper time. As we shall see in the next section, there is a good reason to consider the ordinary force is defined as $d\vec{p}/dt$ when thinking about Newton's equations.

8.1 Newton's laws

The First Law Newton's first law is an object in motion stays in motion whereas an object at rest stays at rest unless acted upon by another object. This law is unchanged by Special relativity, provided we accept that an object at rest does not move unless it is hit by another object. If this is the case, then an object at rest will have the four-vector trajectory

$$x^u = \begin{pmatrix} ct \\ x_i \\ y_i \\ z_i \end{pmatrix} \tag{259}$$

where $\vec{r_i}$ is a constant. In another frame differing from the first by a boost in the $-x$ direction,

$$x'^u = \begin{pmatrix} \gamma(ct + \beta x_i) \\ \gamma(x_i + \beta ct) \\ y_i \\ z_i \end{pmatrix} \tag{260}$$

Solving for ct' and rewriting we have

$$x'^u = \begin{pmatrix} ct' \\ \gamma(x_i + \beta(\frac{ct}{\gamma} - \beta x_i)) \\ y_i \\ z_i \end{pmatrix} = \begin{pmatrix} ct' \\ x_i + vt' \\ y_i \\ z_i \end{pmatrix} \tag{261}$$

As long as the observer in the stationary frame does not see the object move when it is not hit by a force, the observer in the moving frame will see the object move at a constant velocity.

The second law The second law, $\vec{F} = m\vec{a}$, can be though of as simply a definition of force. It becomes more of a significant result when we relate mechanics to other branches of physics. To do this, we relate force to the amount of work done on the object. This is in turn related to the change in energy of the electric, gravitational, or whatever other field is acting on the particle. In classical mechanics, this relationship is made possible by the work-energy theorem:

$$\int_A^B \vec{F}\cdot d\vec{\ell} = m\int_A^B \vec{a}\cdot d\vec{\ell} = m\int_{t_A}^{t_B} \vec{a}\cdot\vec{v}dt = \frac{1}{2}m\int_{t_A}^{t_B}\frac{dv^2}{dt}dt \tag{262}$$

$$= \frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2 \tag{263}$$

This states that work done on a particle is its change in kinetic energy. To generalize this to special relativity Let us assume that the right hand side should be its relativistic energy:

$$\int_A^B \vec{F}\cdot d\vec{\ell} = \Delta E = \Delta p^o c \tag{264}$$

Here the energy is now given by $p^o c$. Does this make sense? Let us assume that $\vec{F} = d\vec{p}/dt$

$$\int_A^B \vec{F}\cdot d\vec{\ell} = \int_A^B (\vec{F}\cdot\vec{u})dt \tag{265}$$

$$= \int_A^B (\frac{d\vec{p}}{dt}\cdot\vec{u})dt \tag{266}$$

Are new form of the work energy theorem and Newton's law $\vec{F} = d\vec{p}/dt$ will only be in agreement if the term in the integrand is equal to dE/dt. But the relativistic momentum vector is given by

$$\vec{p} = \frac{m\vec{u}}{\sqrt{1-u^2/c^2}} \tag{267}$$

The derivative of $\vec{p}$ with respect to time is given by

$$\frac{d\vec{p}}{dt} = \frac{m}{\sqrt{1-u^2/c^2}}\frac{d\vec{u}}{dt} + \frac{1}{2c^2}\frac{m\vec{u}}{(1-u^2/c^2)^{3/2}}\frac{d}{dt}\frac{u^2}{c^2} \tag{268}$$

$$= \frac{m}{\sqrt{1-u^2/c^2}}\frac{d\vec{u}}{dt} + \frac{1}{c^2}\frac{m\vec{u}}{(1-u^2/c^2)^{3/2}}(u\frac{du}{dt}) \tag{269}$$

$$= \frac{m}{(1-u^2/c^2)^{3/2}}\left[(1-u^2/c^2)\frac{d\vec{u}}{dt} + \vec{u}\left(u\frac{du}{dt}\right)\frac{1}{c^2}\right] \tag{270}$$

$$= \frac{m}{(1-u^2/c^2)^{3/2}}\frac{d\vec{u}}{dt} + \frac{m}{c^2(1-u^2/c^2)^{3/2}}\left[\vec{u}\left(u\frac{du}{dt}\right) - u^2\frac{d\vec{u}}{dt}\right] \tag{271}$$

But we must remember that

$$\frac{du}{dt} = \frac{\vec{u}\cdot\frac{d\vec{u}}{dt}}{u} \tag{272}$$

This leads to

$$\left[\vec{u}\left(u\frac{du}{dt}\right) - u^2\frac{d\vec{u}}{dt}\right] = \vec{u}(\vec{u}\cdot\frac{du}{dt}) - u^2\frac{d\vec{u}}{dt} \tag{273}$$

$$= \vec{u}\times(\vec{u}\times\frac{d\vec{u}}{dt}) \tag{274}$$

leading to

$$\frac{d\vec{p}}{dt} = \frac{m}{(1-u^2/c^2)^{3/2}}\left[\frac{d\vec{u}}{dt} + \frac{1}{c^2}\vec{u}\times(\vec{u}\times\frac{d\vec{u}}{dt})\right] \tag{275}$$

Finally we have

$$\frac{d\vec{p}}{dt}\cdot\vec{u} = \frac{m}{(1-u^2/c^2)^{3/2}}\frac{d\vec{u}}{dt}\cdot\vec{u} \tag{276}$$

Let us see if this makes sense. The relativistic energy is given by

$$E = p^o c = \frac{mc^2}{\sqrt{1-u^2/c^2}} \tag{277}$$

Taking its time derivative, we have

$$\begin{aligned}
\frac{dE}{dt} &= \frac{m}{2}\frac{1}{(1-u^2/c^2)^{3/2}}\frac{du^2}{dt} && (278)\\
&= m\frac{1}{(1-u^2/c^2)^{3/2}}(u_x\frac{du_x}{dt} + u_y\frac{du_y}{dt} + u_z\frac{du_z}{dt}) && (279)\\
&= \vec{u}\cdot\left(\frac{m}{(1-u^2/c^2)^{3/2}}\frac{d\vec{u}}{dt}\right) && (280)\\
&= \frac{d\vec{p}}{dt}\cdot\vec{u} \text{ as required.} && (281)
\end{aligned}$$

Thus a consistent theory of dynamics, with forces and conservation of energy can be created by defining force as $\vec{F} = d\vec{p}/dt$. One may ask if the quantity

$$K = \frac{d\vec{p}}{d\tau} \tag{282}$$

is useful. After all, this is the four vector. As it turns out, this force is called the Minkowski force, but it is not related to the energy expended per unit time and as such is perhaps less physical than the ordinary force. Indeed, the Lorentz force law

$$\vec{F} = q(\vec{E} + \vec{u}\times\vec{B}) \tag{283}$$

gives the ordinary force on a particle as it moves.

The third law: The third law does not apply in general in a relativistic world. In fact, for every action there is an equal an opposite reaction only in the case that the force is a contact force.

8.2 Motion of a particle under a constant force

Here we determine the trajectory of a particle experiencing a constant force $\vec{F}$

Classical solution Here we have

$$\begin{aligned}\vec{F} &= \frac{d\vec{p}}{dt} &(284)\\ &= \frac{dm\vec{u}}{dt} &(285)\\ &= m\vec{a} &(286)\end{aligned}$$

To find $\vec{r}$, we start by integrating over the force over the mass:

$$\begin{aligned}\frac{\vec{F}}{m} &= \frac{d\vec{u}}{dt} &(287)\\ \int \frac{\vec{F}}{m}dt &= \frac{\vec{F}}{m}t+\vec{u}_o = \vec{u} = \frac{d\vec{r}}{dt} &(288)\\ \int \left[\frac{\vec{F}}{m}t+\vec{u}_o\right]dt &= \frac{1}{2}\frac{\vec{F}}{m}t^2+\vec{u}_o t+\vec{r}_o = \vec{r} &(289)\\ \vec{r} &= \frac{1}{2}\frac{\vec{F}}{m}t^2+\vec{u}_o t+\vec{r}_o &(290)\end{aligned}$$

Relativistic solution In this case

$$\begin{aligned}\vec{F} &= \frac{d\vec{p}}{dt} &(291)\\ &= \frac{d}{dt}\frac{m\vec{u}}{\sqrt{1-u^2/c^2}} &(292)\end{aligned}$$

Integrating this equation, we have

$$\begin{aligned}\frac{\vec{F}}{m} &= \frac{d}{dt}\frac{\vec{u}}{\sqrt{1-(u_x^2+u_y^2+u_z^2)/c^2}}\\ \frac{\vec{F}}{m}t+\vec{C} &= \frac{\vec{u}}{\sqrt{1-(u_x^2+u_y^2+u_z^2)/c^2}}\end{aligned}$$

The constants of integration $\vec{C}$ do not give the initial velocity, as was the case classically. Instead, as can be seen by evaluating this equation at $t=0$, $\{C_x, C_y, C_z\} = \vec{u}_o\gamma_o = \{u_{ox}, u_{oy}, u_{oz}\}/.\sqrt{1-(u_{ox}^2+u_{oy}^2+u_{oz}^2)/c^2}$. These three equations can be used to solve for the components of $\vec{u}$ with the solution

$$\begin{aligned}\vec{u} &= \frac{(\vec{u}_o+\vec{\alpha}_o t)}{\sqrt{1/\gamma_o^2+|\vec{u}_o+\vec{\alpha}_o t|^2/c^2}}\\ &= \frac{(\vec{u}_o+\vec{\alpha}_o t)}{\sqrt{1+[2(\vec{u}_o\cdot\vec{\alpha}_o)t+\alpha_o^2 t^2]/c^2}}\end{aligned}$$

Here I have the made the substitution

$$\vec{\alpha}_o = \frac{\vec{F}}{m\gamma_o} \tag{293}$$

Notice how similar this is to the classical result. For slow speeds, $\gamma_o \approx 1$, $|\vec{u}_o + \vec{\alpha}_o t|^2/c^2 \approx 0$, and $\vec{\alpha}_o$ is very nearly the acceleration. Now let us integrate again to find the trajectory of the particle:

$$\begin{aligned}
\frac{d\vec{r}}{dt} &= \frac{(\vec{u}_o + \vec{\alpha}_o t)}{\sqrt{1 + [2(\vec{\alpha}_o \cdot \vec{u}_o)t + \alpha_o^2 t^2]/c^2}} && (294)\\
x - x_o &= \int_o^t \frac{(u_{ox} + \alpha_{ox} t')}{\sqrt{1 + [2(\vec{\alpha}_o \cdot \vec{u}_o)t + \alpha_o^2 t^2]/c^2}} dt' && (295)
\end{aligned}$$

This integral is easily solved using mathematica:

$$\begin{aligned}
= \; & \alpha_{ox}\left(\sqrt{1 + [2(\vec{\alpha}_o \cdot \vec{u}_o)t + \alpha_o^2 t^2]/c^2} - 1\right)\left[\frac{c}{\alpha_o}\right]^2 + && (296)\\
& (\mu_{ox} - \frac{\alpha_{ox}}{\alpha}(\vec{u}_o \cdot \widehat{\alpha}_o))\left[\frac{c}{\alpha_o}\right] \ln\left(\frac{\alpha_o t + \vec{u}_o \cdot \widehat{\alpha}_o + \sqrt{c^2 + 2(\vec{u}_o \cdot \vec{\alpha}_o)t + \alpha_o^2 t^2}}{c + \vec{u}_o \cdot \widehat{\alpha}_o}\right) && (297)
\end{aligned}$$

Because the result is linear in u_{ox} and α_{ox}, this result can be generalized to the full solution:

$$\begin{aligned}
\vec{r} \; = \; & \vec{\alpha}_o\left(\sqrt{1 + [2(\vec{\alpha}_o \cdot \vec{u}_o)t + \alpha_o^2 t^2]/c^2} - 1\right)\left[\frac{c}{\alpha_o}\right]^2 + \\
& [\vec{u}_o - \widehat{\alpha}_o(\vec{u}_o \cdot \widehat{\alpha}_o)]\left[\frac{c}{\alpha_o}\right] \ln\left(\frac{\alpha_o t + \vec{u}_o \cdot \widehat{\alpha}_o + \sqrt{c^2 + 2(\vec{u}_o \cdot \vec{\alpha}_o)t + \alpha_o^2 t^2}}{c + \vec{u}_o \cdot \widehat{\alpha}_o}\right) + \\
& \vec{r}_o && (298)
\end{aligned}$$

Of course this result must reduce to the classical result in the limit that the motion is not relativistic. The first term can be written

$$\begin{aligned}
\vec{\alpha}_o\left(\sqrt{1 + [2(\vec{\alpha}_o \cdot \vec{u}_o)t + \alpha_o^2 t^2]/c^2} - 1\right)\left[\frac{c}{\alpha_o}\right]^2 &\approx \vec{\alpha}_o\left(\frac{(\vec{\alpha}_o \cdot \vec{u}_o)t + \frac{1}{2}\alpha_o^2 t^2}{c^2}\right)\left[\frac{c}{\alpha_o}\right]^2 \\
&= \frac{1}{2}\vec{\alpha}_o t^2 + \widehat{\alpha}_o(\widehat{\alpha}_o \cdot \vec{u}_o)t && (299)
\end{aligned}$$

The first term looks good, but there is nothing like the second term in classical mechanics. We are saved by expansion of the second term in powers of 1/c:

$$\ln\left(\frac{\alpha_o t + \vec{u}_o \cdot \widehat{\alpha}_o + \sqrt{c^2 + 2(\vec{u}_o \cdot \vec{\alpha}_o)t + \alpha_o^2 t^2}}{c + \vec{u}_o \cdot \widehat{\alpha}_o}\right) = \frac{a_o t}{c} + O\left[\frac{1}{c^2}\right] \tag{300}$$

(I determined this result by setting $1/c = ic$, $A = \alpha_o t + \vec{u}_o \cdot \widehat{\alpha}_o$, $B = 2(\vec{u}_o \cdot \vec{\alpha}_o)t + \alpha_o^2 t^2$, and $D = \vec{u}_o \cdot \widehat{\alpha}_o$ and executed the Series function to find

$$Series[\ln[\frac{A\,ic + \sqrt{1 + B\,ic^2}}{1 + D\,ic}, \{ic, 0, 1\}] = 0 + (A - D)ic + ORDER[ic^2].) \tag{301}$$

Substituting the approximate results of Equations 299 and 300 into 298 we obtain the classical result.