Physics 200-04
Einstein

Poincare and Lorentz: Throughout the last few years of the 19th century, two people were the chief advocates for a radical rethinking of physics in the face of the problems that the aether presented to physics. In both cases their emphasis was on trying to figure out how and why matter behaved in such a strange way in order to give the outcomes for the experiments. They were searching for a constitutive theory of matter which would make sense of the various experiments. Lorentz realized that instead of Fresnel's strange partial aether drag, one could get the same results by assuming that in the equations of motion for the waves, one assumed that in the new frame not only did $x \rightarrow x - vt$ but also that the time dependence be changed by substituting $t \rightarrow t - vx/c^2$. both examined Maxwell's equations and realized that there were a set of transformations which would leave the Maxwell equations the same (a set of transformations which we call the Lorentz transformations). But both were confused as to how to make these transformations compatible with matter. Lorentz for example tried to make electrons out of the electromagnetic field (so that it would automatically obey the same transformations as the EM field did) but always got hung up with the fact that in order to hold the electron together, non-electromagnetic forces were required.

Einstein It was Einstein who suddenly pulled it all together and showed how, but looking at the problem, and in particular at the nature of space and time in a new way. He based his approach, not on any analysis of how non-electromagnetic matter behaved, but instead postulated two principles. Both principles were based in part on experiment and in part on intuition. The first was that any attempt to measure the velocity of light by any means would always give the value c. Light never traveled at any different speed. The second principle he borrowed from Galileo– namely that there was not experiment that anyone could do which could determine the velocity of a system internally (ie without any reference to or interaction with some external system). This second was probably the more radical, since, although Galileo had paid it lip service, he had denied it in his theory of tides. Newton had explicitly denied it in his theory of absolute space, even though his equations were Galilean transformation invariant. The Aether theorists, and the wave nature of light seemed to contradict this principle. Nevertheless, Einstein based his whole approach on it.

I will not follow Einstein in his derivation, but rather take the road we began on three lectures ago.

The statement that light always travels with velocity c can be phrased by writing down the statement that the equation of motion for a particle of light, or for a wave packet of light made to look like a particle, is

$$\text{distance traveled} = ct \tag{1}$$

Or, if we assume that the light starts out from the point $x = y = z = 0$ at $t = 0$, then this can be written as

$$\sqrt{x^2 + y^2 + z^2} = ct \tag{2}$$

or, to make this more symmetrical,

$$x^2 + y^2 + z^2 - c^2t^2 = 0 \tag{3}$$

Now, let us imagine that this same light ray, or set of light rays is observed by someone traveling past us with some velocity. They will have a time defined t' and will have spatial coordinates defined x', y', z'. Because there should be no way in which this observer could tell that he is traveling just by looking at how the light moves, light must obey exactly the same equation in this new system of coordinates as in the old. Ie, the equation of motion for the light must again be

$${x'}^2 + {y'}^2 + {z'}^2 - c^2{t'}^2 = 0 \tag{4}$$

Now, there most be some relationship between the new coordinates and the old. Since in each case I assumed that the light started out at the origin at time zero, we will need that $x = y = z = t = 0$ corresponds to $x' = y' = z' = t' = 0$. Furthermore, it is hard to see how the relationship between the unprimed and primed coordinates could be non-linear, as that would surely introduce terms which are more than just quadratic in the new coordinates.

I will assume that the relation is linear. furthermore, I will assume for simplicity that we are working in two dimensions rather than four– ie, with coordinates x, t and x', t'. We will come back to the four dimensions later.

Thus we have

$$x = \alpha x' + \beta ct' \tag{5}$$

$$ct = \gamma x' + \delta ct' \tag{6}$$

Substituting this into

$$x^2 - (ct)^2 \tag{7}$$

we get

$$(\alpha^2 - \gamma^2){x'}^2 + 2(\alpha\beta - \gamma\delta)x'ct' - (\delta^2 - \beta^2)(ct')^2 \tag{8}$$

In order that the equation

$${x'}^2 - c^2{t'}^2 = 0 \tag{9}$$

follows from

$$x^2 - c^2t^2 = 0 \tag{10}$$

we need that

$$\alpha\beta - \gamma\delta = 0 \tag{11}$$

$$\alpha^2 - \gamma^2 = \delta^2 - \beta^2 \tag{12}$$

Let us begin by assuming that $\alpha^2 - \gamma^2 = 1$ and we will look at the possibility that this is not unity later on.

We can solve this by defining $\gamma = \sinh(\theta)$ for some θ where

$$\sinh(\theta) = \frac{e^{\theta} - e^{-\theta}}{2} \tag{13}$$

Then we have

$$\alpha = \cosh(\theta) \tag{14}$$

by the hyperbolic trigonometric identities.

Similarly

$$\delta = \cosh(\theta') \tag{15}$$

$$\beta = \sinh(\theta') \tag{16}$$

and finally from $\alpha\beta - \gamma\delta = 0$

$$\cosh(\theta)\sinh(\theta') - \sinh(\theta)\cosh(\theta') = 0 \tag{17}$$

or by the hyperbolic identities

$$\sinh(\theta - \theta') = 0 \tag{18}$$

from which we get that $\theta = \theta'$.

Thus we find that

$$x = \cosh(\theta)x' + \sinh(\theta)ct' \tag{19}$$

$$ct = \cosh(\theta)ct' + \sinh(\theta)x' \tag{20}$$

Note that this immediately leads to the result that in the four dimensional case the above transformation together with

$$y = y' \tag{21}$$

$$z = z' \tag{22}$$

works– ie is a transformation which leaves the speed of light constant in the second frame.

We can invert these equations as well to give x', y', z', t' in terms of x, y, z, t.

$$x' = \cosh(\theta)x - \sinh(\theta)ct \tag{23}$$

$$t' = \cosh(\theta)t - \sinh(\theta)\frac{x}{c} \tag{24}$$

$$y' = y \tag{25}$$

$$z' = z \tag{26}$$

Note that the a particle located at $x = 0$ is in the new frame given by $x' = -\frac{\sinh(\theta)}{\cosh(\theta)}ct'$. Ie, in the $x't'$ frame, that particle is moving with a velocity of

$$-v = -ctanh(\theta) \tag{27}$$

Similarly the point $x' = y' = z' = 0$ moves with velocity $v = c\tanh(\theta)$ in the plus x direction.

We can use these to express θ in terms of v since $tanh(\theta) = \frac{v}{c}$ (in your assignment you show what $\cosh(\theta)$ and $\sinh(\theta)$ are in terms of v/c.)

$$\cosh(\theta) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{28}$$

$$\sinh(\theta) = \frac{\frac{v}{c}}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{29}$$

Matrix version Define the four dimensional vector

$$X = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \tag{30}$$

Also define the square matrix

$$G = \begin{pmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{31}$$

Then the length of the vector X can be defined as

$$length(X)^2 = X^T G X = -c^2 t^2 + x^2 + y^2 + z^2 \tag{32}$$

The Lorentz transformation L is a transformation written in matrix form as $X = LX'$ which leaves the length squared the same form. This means that we need

$$(LX')^T G L X' = X'^T G X' \tag{33}$$

But the left hand side is

$$(LX')^T GLX' = X'^T L^T GLX' = (X'^T)(L^T GL)X' \tag{34}$$

The only way this will be true for all vectors X' is if

$$L^T GL = G \tag{35}$$

This is the condition for a linear transformation designated by L to be a Lorentz transformation. G is usually called the metric. Ie, the Lorentz transformations are transformations which leave the metric invariant.

If L leaves t alone ($t = t'$) then the Lorentz transformations are exactly the rotations of x, y, z.

Note that any product of Lorentz transformations is also a Lorentz transformation. If $L = L_1 L_2$ where L_1 and L_2 are Lorentz transformations, then

$$L^T GL = (L_1 L_2)^T GL_1 L_2 = L_2{}^T L_1{}^T GL_1 L_2 = L_2^T (L_1^T GL_1) L_2 \tag{36}$$

$$= L_2^T GL_2 = G \tag{37}$$

Ie, then L is also a Lorentz transformation. The set of all Lorentz transformations form a group.

[I would like to point out how easy this is to prove using matrices. Had I had to try to prove it using the full linear transformations written out in full, it would have taken about 10 pages of huge messy scribbling. It takes two lines using matrices– the power of mathematical notation. Of course you must be able to backtrack to figure out what the notion means after you are finished.]

Note that if we define

$$G^{-1} = \begin{pmatrix} -c^{-2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{38}$$

then

$$G^{-1}G = GG^{-1} = I \tag{39}$$

We then have

$$L^T GL = G \Rightarrow (G^{-1} L^T G)L = G^{-1}G = I \tag{40}$$

so that $G^{-1} L^T G$ is the inverse of L.

Time The most surprising feature of this transformation is the change in the nature of time. In particular the time in the new frame is not the same as the time in the old frame. This difference manifests itself in two ways. The first is that the notion of simultaneity changes. The t=constant surface is not the same as the t'=constant surfaces. For example, the $t = 0$ surface is given by $\cosh(\theta)t' + \sinh(\theta)x'/c = 0$, or

$$t' = -\frac{v}{c^2}x' \tag{41}$$

The "same time" means something different in one frame than it does in the other.

An even more surprising feature is that the time intervals themselves does not mean the same thing. If $x' = 0$, then $t = \cosh(\theta)t'$. Ie, the time designated by clocks in the t, x frame is not the same as that in the t', x' frame. Part of this comes about because of the change in simultaneity, and part because the clocks really tick differently in the two frames.

The most dramatic example of this is the so called twin's paradox. (there is nothing paradoxical except with our preconceptions about how time behaves.) Say that Bob leaves his twin, Alice and travels off at a speed v (with respect to Alice) for a certain time. He then travels back at that same speed v until he again meets Alice. On the way out, the time that Alice claims has elapsed until the turnaround is $\cosh(\theta)t'$ where t' is the time elapsed for Bob until the turnaround. But on the way back, the time elapsed will be the same. Thus if for Bob a time $2t'$ has elapsed, for Alice a time $2cos(\theta)t'$ will have elapsed. If v is sufficiently large (ie close to c), $\cosh(\theta)$ will be large, and time elapsed from leaving to re-meeting for Alice will be much larger than it is for Bob (by the factor $\frac{1}{\cosh(\theta)} = \sqrt{1 - \left(\frac{v}{c}\right)^2}$.)

This of course seems on the face of it absurd. How can the times differ? Time is surely a universal. In fact we have all been trained since we were very young to regard time in this way. For very young children, Piaget has shown that they will regard time as situation dependent. The amount of time passing for them for a moving block is different than for a stationary block. Our education system– pointing to clocks, telling them that there is a universal clock time which is the same for all clocks, punishing them for being late and refusing to accept that they thought less time had passed than for you– ensures that we have had it ingrained into us that time exists as some unitary, universal, all encompassing thing. On this basis it is of course lunacy to talk about time being situation dependent, that time can depend on state of your motion with respect to someone else. Unfortunately the world need not share our prejudices. While Newton together with his notions of absolute space, also gave us the notion of absolute time, it turns out to conflict with the way in which the world operates.

Lengths: How long is something? How do you compare lengths? The answer at first seems obvious. You lay a ruler beside the object and look at where along the ruler the two ends of the object lie. But things become more difficult if the object is moving. How do you measure the length of a moving object? What do you mean when you say "the length of that moving object is

1 meter"? Again it is obvious. Just lay your ruler down, and when the one end of the object is lined up against the one end of the meter stick, find out where along the meter stick the other end lies. But there is a world of difficulty in that word "when". When you say "when the end...., look find out where..." clearly depends on what you mean by "when", what you mean by "the same time" (which is what you take "when" to mean.) Clearly it is not correct to simply line up the one end, wander out for a coffee and a cinnamon bun, and then come back and see where along the stick the other end lies. You want to line up the two ends "at the same time". But, as noted above, "at the same time" means different things to different people. The notion of simultaneity depends on your state of motion. What is "the same time" to one observer, is not the same time to another.

So, let us look at the notions. Let us say we have the two ends of an object lying at rest at the points $x = 0$ and $x = L$. For the observer in the x' frame, these ends will travel along the lines

$$0 = \cosh(\theta)(x'_1 + vt'_1) \tag{42}$$

$$L = \cosh(\theta)(x'_2 + vt'_2) \tag{43}$$

If we now want to determine the length of the body $(x'_2 - x'_1)$ at the same t' time, ie at $t'_1 = t'_2$, we find that

$$x'_2 - x'_1 = \frac{L}{\cosh(\theta)} = \sqrt{1 - \frac{v^2}{c^2}}L \tag{44}$$

Ie, the length of the body in the moving frame, assuming that we want the length "at the same time" in that same frame is shorter, by exactly the Fitzgerald length contraction, $\cosh(\theta) = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$.

[The following paragraph is an alternative that is almost never used. It is here only to illustrate the importance of synchronisation to the definition of moving lengths]

On the other hand you might want to use the t time to define your simultaneity for some reason. The person in the x, t frame might set off flash lamps at the two ends of the rod at the same time in his frame, and you want to know how far apart those flashes are in your frame. In this case, it is not t'_1 and t'_2 which are equal, but

$$t_1 = \cosh(\theta)t'_1 + \sinh(\theta)x'_1/c \tag{45}$$

$$t_2 = \cosh(\theta)t'_1 + \sinh(\theta)x'_2/c \tag{46}$$

which are equal. Thus, we have, by subtracting the equations for x_2 and x_1 and those for t_2 and t_1,

$$L = \cosh(\theta)\left((x'_2 - x'_1) + v(t'_2 - t'_1)\right) \tag{47}$$

$$0 = \cosh(\theta)\left((t'_2 - t'_1) + \frac{v}{c^2}(x'_2 - x'_1)\right) \tag{48}$$

or

$$x'_2 - x'_1 = L\cosh(\theta) = \frac{L}{\sqrt{1-\frac{v^2}{c^2}}} \tag{49}$$

which is longer than L.

Again, lengths also depend on the situation. You must precisely define what you mean by a length, how you define "at the same time" in defining the length, to get an answer, and the answer depends on how you do it.

General Comments

It is critical to note that there is no talk of the Aether or of the interaction with matter in this whole procedure. However, the two assumptions Einstein made– constancy of light velocity and relativity– clearly have immense implications about the way that matter behaves. When we talk about times and distances in the above, we assume that these times correspond with the times as ticked off by real, mechanical, material clocks. We assume that when one talks about lengths that this is the way that real, mechanical, material rulers behave. If they did not, then when one used those clocks and rules, one would either find that the speed of light was not constant, or that one could, by simply doing local experiments, tell the difference between a moving and a non-moving system.

Relativity of sound: To emphasis the above points, we note that there is nothing in the above mathematics which demands that c be the speed of light. The mathematics would be identical if we took c to be the velocity of sound. Demanding that the velocity of sound be the same in all frames would lead to identical equations, identical Lorentz transformations, as the above. However, in this case clocks do not behave in that way. If one were to make a real mechanical clock, and have it move through the water, its rate of ticking would not obey the above relations. If one made a ruler, its length would not obey the above relations. One would find both that the velocity of sound as measured with real mechanical clocks and rods would not be the same in all frames. One would find that one could do experiments which could tell you whether or not you were moving with respect to the water or not. **If** you could make your clocks and rods purely out of sound, with no other matter involved (not even the characteristics of the water other than the sound waves), then according to those clocks and rods sound would obey Einstein's relativity. Just as Lorentz argued, that if one could make all matter out of Electromagnetism, then the clocks and rods would almost by definition, obey special relativity. But it is not true that you can make all rods and clocks out of electromagnetism (not least because of the instability of charged matter). You need other types of matter. If Einstein's relativity is correct that **ALL** matter must behave in such a way as to ensure that relativity is correct. It is possible that there exists matter which does not. However, in that case one would find that either the speed of light would not be the same in all frames for clocks and rods made out of that kind of matter, and/or that one could find a universal rest frame and measure one's velocity with respect to that absolute space and time.

Conformal Transformation In the above derivation, I took the option that $\alpha^2 - \gamma^2 = 1$. What would happen if instead I took $\alpha^2 - \gamma^2 = K^2(v)$? In this case, we would have

$$x = K(\cosh(\theta)x' + \sinh(\theta)ct') \tag{50}$$

$$ct' = K(\cosh(\theta)ct' + \sinh(\theta)x' \tag{51}$$

$$y = Ky' \tag{52}$$

$$z = Kz' \tag{53}$$

It is the transformations of the y and z componenets which now cause the problem. This K(v) cannot be universal. We want the system to remail rotationally invariant in the new frames. Thus be can boost to velocity v along the x axis as above, now rotate around the y axis by 180^o, and again boost along the x axis. But this second boost will undo the first. Ie it should take us back to the original frame. But y and z are multiplied by K on each boost. Thus you do not get back to the original frame. Thus in some frame, one must destroy rotation invariance. But then we could tell the difference between the various frames by seeing if they were rotationally invariant or not.

Thus, K would have to depend on which frame we were in. But if it does, then K would allow us to determine the frame we were in by doing a transformtion to a new frame and measuring the lengths in the y or z directions. The one in which the object had the shortest (or longest) length in that direction would now be a special frame, different from the others. Ie, Einstein's second postulate, that there is not physical difference between the physics in any of the frames, is violated. Thus, even though such a transformation would keep the velocity of light the same in all frames, it would not maintain the indistinguishability of frames.