Physics 504, Lecture 15
March 24, 2011

Copyright©2009 by Joel A. Shapiro

1 Relativity

Galileo said that the laws of physics are the same in different reference frames that move, relative to each other, with constant velocity. That is, if $\mathcal{O}$ expresses physics in coordinates $\vec{x}, t$ and $\mathcal{O}'$ expresses physics in coordinates $\vec{x}\,', t'$, where the coordinates for a physical event are related by

$$\vec{x}\,' = \vec{x} - \vec{v}t, \qquad t' = t, \tag{1}$$

for constant $\vec{v}$, then the equations in terms of $\vec{x}, t$ have the same form as those in terms of $\vec{x}\,', t'$, though the two observers do not agree on the values of the coordinates themselves. Of course, all aspects of the physical situation need to be so reexpressed. But Newton told us that all forces are what one body exerts on the other, and as long as the forces only depend on the displacements of one from the other, and not on their actual coordinates, the forces will be unchanged, and though the velocities of an object differ ($\vec{v}\,'_i = \vec{v}_i - \vec{v}$), the accelerations agree, so Newton's laws are invariant under this Galilean transformation.

The wave equation is not invariant, for under (1)

$$\left.\frac{\partial}{\partial t}\right|_{\vec{x}} = \sum_j \left.\frac{\partial x'_j}{\partial t}\right|_{\vec{x}} \left.\frac{\partial}{\partial x'_j}\right|_{x'_k, t'} + \left.\frac{\partial}{\partial t}\right|_{\vec{x}\,'}$$

or

$$\left.\frac{\partial}{\partial t}\right|_{\vec{x}} = -\vec{v}\cdot\vec{\nabla}\,' + \left.\frac{\partial}{\partial t}\right|_{\vec{x}\,'},$$

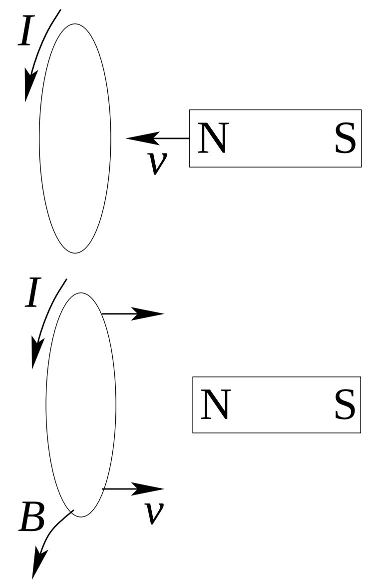
while $\left.\vec{\nabla}\right|_t = \left.\vec{\nabla}\right|_{t'}$. Thus

$$\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} = \nabla'^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t'^2} + \frac{2}{c^2}\frac{\partial}{\partial t}\vec{v}\cdot\vec{\nabla}\,' - \frac{1}{c^2}(\vec{v}\cdot\vec{\nabla}\,')^2.$$

Of course if this was the wave equation for sound through motionless air, Galileo would not be concerned, because we neglected to account for the fact

that the air is moving according to $\mathcal{O}'$, and we don't expect the wave to satisfy the simple wave equation. The same might be true for electromagnetic waves if they represented disturbances in a medium with a definite rest-frame, but various 19^{th} century experiments showed no sign and considerable difficulties with the assumption that the ether has a rest frame.

In his first paper on relativity, Einstein began by pointing out that the conventional explanation of why a current is induced in a stationary loop of wire when a magnet approaches it, and the conventional explanation of why a current is induced in the loop if it is moved towards a stationary magnet, are totally different, even though the effect is the same, depending only on the relative velocities of the loop and magnet. In the first case, the magnetic flux through the coil is changing, and this induces an electric field by Faraday's law, which produces a current. In the second case, the charges in the wire have a velocity towards the right, with a magnetic field diverging to the left, so the Lorentz force is in the axial direction, and creates the same current.



Galileo would have thought these two situations are really the same, and Maxwell predicts the effects are the same, so, Einstein said, the explanations should be the same as well. And we shouldn't invoke Fitzgerald contractions to explain why we can't detect our motion through the preferred reference frame, and just conclude the truth of the postulates of relativity:

- All the laws of physics are the same, whether described by either of two systems in uniform relative motion. Included here is the notion that space is homogeneous and isotropic.

- The speed of light in vacuum is a specific finite c, independent of the motion of its source.

These sensible-seeming postulates require us to give up long-held intuitive notions of how two such coordinate systems are related. In particular, looking at the times that a light ray travels from A to B and back shows that simultaneity is relative, for an observer at rest with respect to A and B say the two trips take equal times, and so the time at B is halfway between the start and end at A. But for an observer traveling in the direction from A to B, he sees B retreating from the incident ray, so A to B takes longer, while A is approaching the returning ray, so B to A takes less time. So the two observers must disagree on the times certain events take place.

There must be a relation that gives $\mathcal{O}'$'s coordinates $\vec{x}\,', t'$ in terms of $\vec{x}, t$. By homogeneity the relation must be linear, as it can't depend on the origin of the coordinates, so we can write the relation as multiplication by a matrix. With the fundamental constancy of the speed of light it makes sense to measure times as distance/c. Another notational change is that the three coordinates of space will be written with superscript indices rather than subscripts, so we write $x^0 = ct$, $\vec{x} = (x^1, x^2, x^3)$. We also use greek letters for indices which run from 0 to 3 rather than roman letters. Thus we can write the transformation of coordinates in Einsteinian relativity

$$x'^{\,\alpha} = \sum_\beta A^\alpha{}_\beta x^\beta,$$

where the 4×4 matrix A depends only on the relative velocity $\vec{v}$ of $\mathcal{O}'$ relative to $\mathcal{O}$. The path of a ray of light must satisfy $c^2(\Delta t)^2 - (\Delta\vec{x})^2 = 0$ and $c^2(\Delta t')^2 - (\Delta\vec{x}\,')^2 = 0$. This quadratic form plays a fundamental role, and is called the invariant length $(\Delta s)^2$, though it is not always positive. In order to write it elegantly,

$$(\Delta s)^2 = \sum_{\alpha\beta} \eta_{\alpha\beta}(\Delta x^\alpha)(\Delta x^\beta),$$

where we define the Minkowski metric tensor[1] $\eta_{00} = 1$, $\eta_{ij} = -\delta_{ij}$ for i and j from 1 to 3, and $\eta_{0i} = \eta_{i0} = 0$. Notice that the indices on η here are subscripts, and that the index pairs summed over are one superscript with one subscript. With a suggestion that he called his greatest contribution to human knowledge, Einstein said we could leave out the summation sign whenever we see a repeated index, once up and once down. So

$$(\Delta s)^2 = \eta_{\alpha\beta}(\Delta x^\alpha)(\Delta x^\beta).$$

[1]Jackson uses g rather than η, but g is used for the more general metric tensor of general relativity, and it is now more common to use η for the *special* relativity metric. Many authors use opposite signs, $\eta_{00} = -1$, $\eta_{ij} = \delta_{ij}$.

Indices upstairs are called contravariant, and those downstairs covariant. A 4-vector such as x^μ with a contravariant index is called contravariant. All contractions (that is, summation over repeated indices) are of one covariant and one contravariant index.

Now the requirement that observers agree that a light beam travels at c says $(\Delta s')^2 = 0$ whenever $(\Delta s)^2$ is. We will use this to determine the restrictions on $A^\alpha{}_\beta$ for a Lorentz transformation, that is, for the transformation between two inertial observers. Our approach here is formal and mathematical, though a much more physical (even better) approach comes from considering *gedanken* experiments, such as a clock made of parallel mirrors, lightning bolts hitting trains, meter rods perpendicular and parallel to the motion, and the Michelson-Morley result. You have presumably seen such discussions several times in undergraduate courses[2] So, from $(\Delta s')^2 = \eta_{\alpha\beta}A^\alpha{}_\mu A^\beta{}_\nu \Delta x^\mu \Delta x^\nu$, we see that the matrix $M_{\mu\nu} := \eta_{\alpha\beta}A^\alpha{}_\mu A^\beta{}_\nu$ is a real symmetric matrix, which vanishes when sandwiched between lightlike vectors $x^0 = |\vec{x}|$. Applying this to $(x^0, \vec{x})$ and $(x^0, -\vec{x})$ tells us first that $M_{00}|\vec{x}|^2 \pm |\vec{x}| \sum M_{0i}x_i + \sum M_{ij}x_i x_j = 0$ for any vector $\vec{x}$. The difference tells us $M_{0i} = M_{i0} = 0$, and then that the spacial part of M gives the same $\vec{x} \cdot \mathbf{M} \cdot \vec{x}$ independent of the direction of $\vec{x}$ tells us that part of M is proportional to the unit matrix, so altogether

$$\eta_{\alpha\beta}A^\alpha{}_\mu A^\beta{}_\nu = \lambda(v)\eta_{\mu\nu},$$

and $(\Delta s')^2 = \lambda(v)(\Delta s)^2$ for arbitrary δx. But by isotropy $\lambda(\vec{v})$ is independent of the direction of $\vec{v}$, and the fact that $\mathcal{O}'$ and $\mathcal{O}$ can be interchanged by $\vec{v} \to -\vec{v}$ gives $(\lambda(v))^2 = 1$. We see that $\lambda(v) = 1$. (Not -1 because that would conflict with continuity in $\vec{v}$ and the nonrelativistic limit).

Thus the condition that the transformation matrix A be a Lorentz transformation is

$$\eta_{\alpha\beta}A^\alpha{}_\mu A^\beta{}_\nu = \eta_{\mu\nu} \tag{2}$$

If $\vec{v}$ is in the x^1 direction, the origin of $\mathcal{O}'$ is $x'^\mu = (t', 0, 0, 0)$ corresponds[3]

[2]If not, you **must** read about them, in Smith or other reference listed on the otherrefs web page.

[3]Invariance under translation, $x^\mu \to x^\mu + c^\mu$ for constant c^μ is unchanged by Einstein, and trivially handled. So for convenience we will deal with Lorentz transformations where the origins of the coordinate systems are the same point at $t = t' = 0$. The broader class of invariant transformations, throwing in translations as well, is called Poincaré transformations.

to $x^\mu = (ct, vt, 0, 0)$. From $\eta_{\mu\nu}x'^\mu x'^\nu = \eta_{\mu\nu}x^\mu x^\nu$ we have $t' = t\sqrt{1 - v^2/c^2}$.

$$\text{Notation:} \qquad \beta = v/c, \quad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}} = \frac{1}{\sqrt{1-\beta^2}}$$

Looking at $\mathcal{O}$'s origin, with $x^\mu_{\mathcal{O}} = (ct, 0, 0, 0)$ and $x'^\mu_{\mathcal{O}} = (ct', -vt', 0, 0)$, where this t and t' are not the same as above (indeed, $t' = \gamma t$ here, while $t' = t/\gamma$ for $\mathcal{O}$'s origin). As $x'^\mu_{\mathcal{O}} = A^\mu{}_\nu x^\mu_{\mathcal{O}} = A^\mu{}_0 t$, we see that the first column, $A^0{}_0 = \gamma$, $A^1{}_0 = -\beta\gamma$, $A^2{}_0 = 0$, $A^3{}_0 = 0$, is determined.

So we have the first ($\mu = 0$) column of $A^\mu{}_\nu$. Looking again at

$$x'^\mu_{\mathcal{O}'} = \begin{pmatrix} ct' \\ 0 \\ 0 \\ 0 \end{pmatrix} = \mathbf{A}\begin{pmatrix} c\gamma t' \\ \beta\gamma t' \\ 0 \\ 0 \end{pmatrix} = ct'\begin{pmatrix} \gamma^2 + A^0{}_1\beta\gamma \\ -\beta\gamma^2 + A^1{}_1\beta\gamma \\ A^2{}_1\beta\gamma \\ A^3{}_1\beta\gamma \end{pmatrix}$$

we find $A^1{}_1 = \gamma$, $A^2{}_1 = A^3{}_1 = 0$, and $A^0{}_1 = (1-\gamma^2)/\beta\gamma = \gamma(\gamma^{-2} - 1)/\beta = -\beta\gamma$. We now have the first two columns, and from

$$\begin{aligned} 0 = \eta_{0i} = \eta_{\mu\nu}A^\mu{}_0 A^\nu{}_i &= \gamma A^0{}_i + \beta\gamma A^1{}_i \\ 0 = \eta_{1i} = \eta_{\mu\nu}A^\mu{}_1 A^\nu{}_i &= -\beta\gamma A^0{}_i - \gamma A^1{}_i \end{aligned}$$

(for $i = 2, 3$) gives $A^0{}_i = A^1{}_i = 0$.

The remaining elements satisfy the requirements for a rotation about the x axis, but any such rotation would violate parity, so we have determined

$$A^\mu{}_\nu = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where the top row and left column are the $\mu = 0$ and $\nu = 0$ elements respectively.

Another parameter used is the rapidity $\zeta = \tanh^{-1}\beta$, with $\gamma = \cosh\zeta$, $\beta\gamma = \sinh\zeta$, and the matrix A looks much like a hyperbolized rotation.

$$A^\mu{}_\nu = \begin{pmatrix} \cosh\zeta & -\sinh\zeta & 0 & 0 \\ -\sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

This analogy to rotations is made clearer if we compare the condition that A represent a Lorentz transformation (2) to the condition that $\mathbf{O}$ is an orthogonal matrix,

$$\delta_{ij}\mathbf{O}^i{}_k\mathbf{O}^j{}_\ell = \delta_{k\ell}$$

which is required of a rotation. We see that Lorentz transformations are a kind of rotation modified to account for the relative minus sign for $(\Delta x^0)^2$ in the invariant length.

The matrix which describes how x^μ transforms,

$$A^\alpha{}_\beta = \frac{\partial x'^\alpha}{\partial x^\beta}, \tag{3}$$

is also how any other contravariant vector transforms, so if $\mathcal{O}$ describes something with a vector B^μ, $\mathcal{O}'$ will use

$$B'^\mu = A^\mu{}_\nu B^\nu.$$

The invariant product of two vectors is therefore

$$B \cdot C = \eta_{\mu\nu}B^\mu C^\nu,$$

and *not* $\sum_\mu B^\mu C^\mu$. We can define, for every contravariant vector V^μ, a *covariant vector* $V_\mu := \eta_{\mu\nu}V^\nu$, which has the same physical content but whose spatial components have the opposite sign[4]. To make a contravariant vector from a covariant one, $V^\mu := \eta^{\mu\nu}V_\nu$, where $\eta^{\mu\nu}$ is the inverse of $\eta_{\mu\nu}$. That is $\eta^{\mu\nu}\eta_{\nu\rho} = \delta^\mu_\rho$, where we now need to write the Kronecker delta with one upper and one lower index, but it is still 1 if $\mu = \rho$ and 0 otherwise. Note that the actual matrices $\eta^{\mu\nu}$ and $\eta_{\mu\nu}$ are the same.

A covariant vector transforms by

$$V'_\mu = \eta_{\mu\nu}V'^\nu = \eta_{\mu\nu}A^\nu{}_\rho V^\rho = \eta_{\mu\nu}A^\nu{}_\rho\eta^{\rho\sigma}V_\sigma,$$

so a covariant vector transforms by

$$V'_\mu = A_\mu{}^\nu V_\nu, \quad \text{where } A_\mu{}^\nu := \eta_{\mu\rho}A^\rho{}_\sigma\eta^{\sigma\nu}.$$

This is consistent with the general rule, that any index can be raised with $\eta^{\mu\rho}$ or lowered with $\eta_{\mu\rho}$.

[4]In the curved space of general relativity, the metric tensor is not trivial the way η is, so the relation of a covariant and its contravariant tensor is more complicated, though it is still true that they represent the same physical quantity in a sense.

Of course if $\mathbf{L}$ is the transformation from $\mathcal{O}$'s coordinates to $\mathcal{O}'$'s, given by the matrix $A(\mathbf{L})$, the inverse transformation $\mathbf{L}^{-1}$ from $\mathcal{O}'$'s coordinates to $\mathcal{O}$'s is also a Lorentz transformation with $A\left(\mathbf{L}^{-1}\right) = A^{-1}(\mathbf{L})$, or

$$A^\gamma{}_\beta\left(\mathbf{L}^{-1}\right) A^\beta{}_\nu(\mathbf{L}) = \delta^\gamma_\nu.$$

But if we multiply (2) by $\eta^{\mu\gamma}$ we get

$$\eta_{\alpha\beta} A^\alpha{}_\mu\left(\mathbf{L}\right) \eta^{\mu\gamma} A^\beta{}_\nu\left(\mathbf{L}\right) = A_\beta{}^\gamma\left(\mathbf{L}\right) A^\beta{}_\nu\left(\mathbf{L}\right) = \eta_{\mu\nu}\eta^{\mu\gamma} = \delta^\gamma_\nu,$$

so as $A^\beta{}_\nu\left(\mathbf{L}\right)$ is invertible, we have

$$A_\beta{}^\gamma\left(\mathbf{L}\right) = A^\gamma{}_\beta\left(\mathbf{L}^{-1}\right).$$

It is in this sense that A is pseudo-orthogonal.

Note that as $\mathbf{L}^{-1}$ is the transformation from $\mathcal{O}'$ to $\mathcal{O}$,

$$A^\gamma{}_\beta\left(\mathbf{L}^{-1}\right) = \frac{\partial x^\gamma}{\partial x'^\beta}, \qquad \text{so } A_\beta{}^\gamma\left(\mathbf{L}\right) = \frac{\partial x^\gamma}{\partial x'^\beta}.$$

For rotations we know that there are infinitesimal generators, and an arbitrary rotation R can be written as $R = e^{i\theta_j L_j}$, where j indexes the independent infinitesimal rotations (3 in 3 dimensions, but $D(D-1)/2$ in D dimensions) and L_j is an imaginary antisymmetric matrix. This latter requirement (which makes R both unitary and real, hence orthogonal) is what tells us how many independent generators there are. For our Lorentz transformations, if an infinitesimal one is

$$A^\alpha{}_\mu = \delta^\alpha_\mu + \epsilon L^\alpha{}_\mu,$$

the requirement (2) to first order in ϵ gives

$$\eta_{\alpha\beta}\left(\delta^\alpha_\mu + \epsilon L^\alpha{}_\mu\right)\left(\delta^\beta_\nu + \epsilon L^\beta{}_\nu\right) = \eta_{\mu\nu} \Longrightarrow \epsilon\left(\eta_{\alpha\nu} L^\alpha{}_\mu + \eta_{\mu\beta} L^\beta{}_\nu\right) = 0,$$

which tells us $L_{\nu\mu} + L_{\mu\nu} = 0$ or $L_{\nu\mu}$ is antisymmetric, and there are 6 independent generators.

Three of these generators can be taken to be $-(K_i)_{0i} = (K_i)_{i0} = 1$, all other elements zero, which are the Lorentz boosts, and $(S_i)_{jk} = \epsilon_{ijk}, (S_i)_{0i} = (S_i)_{i0} = 0$, which represent the generators of spatial rotations, which do satisfy the requirement of preserving the invariant length and are therefore considered among the Lorentz transformations.

We have described how vectors transform under Lorentz transformations and how to relate co- and contra-variant forms. More generally we may have tensors, with several lorentz (4-vector) indices. Then in general the transformation properties are given by transforming each index. That is, if we had some object $M^\mu{}_\nu{}^\rho$, its value for $\mathcal{O}'$ is[5]

$$M'^\mu{}_\nu{}^\rho = A^\mu{}_\alpha A_\nu{}^\beta A^\rho{}_\sigma M^\alpha{}_\beta{}^\sigma.$$

So far we have discussed only transformation properties and not any physical quantities other than x^μ, spatial and temporal positions. Particle motion is described nonrelativistically by $\vec{x}(t)$, but now we are encouraged to think of the trajectory taken in spacetime, $x^\mu(\lambda)$, a curve parameterized by λ. The physical history is described by the path itself, not on how the parameter λ runs, and it is often convenient to choose the parameter to be the elapsed proper time τ, defined by $d\tau = \sqrt{ds^2}/c = \sqrt{(dt)^2 - |d\vec{x}|^2/c^2} = dt/\gamma(t)$. Here we have $\beta(t) = |d\vec{x}/dt|/c$ and $\gamma(t) = 1/\sqrt{1-\beta^2(t)}$, which apply to the (time-dependent) speed of the particle and not some other inertial observer.

Newtonian mechanics involves accelerations and forces, which are derivatives with respect to time of velocities $\vec{u}(t)$ and momenta. But time is not an invariant, so the transformation properties of such things might be convoluted. Instead of velocity, consider the 4-velocity

$$u^\mu := \frac{dx^\mu}{d\tau} = (c\gamma(u), \gamma(u)\vec{u})$$

where $\gamma(u) = 1/\sqrt{1-\vec{u}^2/c^2}$. As τ is an invariant, the 4-velocity transforms like a contravariant vector.

I am assuming you have studied special relativity at the freshman level, so you are aware of the relativistic expressions for momentum and energy of a particle, and of the velocity addition formula, at least for colinear motion. But I will briefly rephrase these, while Jackson can derive them for you from scratch if you have never seen them before. Still, let's do the standard space ship problem: A spaceship is traveling in the z direction at velocity v_s with respect to the Earth, and it has a gun designed to shoot bullets at velocity

[5] And if M is a field,

$$M'^\mu{}_\nu{}^\rho(\vec{x}') = A^\mu{}_\alpha A_\nu{}^\beta A^\rho{}_\sigma M^\alpha{}_\beta{}^\sigma(\vec{x}),$$

where $\vec{x}'$ now means the four-position (x'^0, x'^1, x'^2, x'^3)

v_b, with respect to itself, of course, in the forward direction. What is the velocity of the bullet with respect to the Earth?

If the spaceships coordinates are x'^{μ}, the bullet has $dz'/dt' = v_b$. The correct Lorentz transformation from spaceship to Earth must give $z = v_s t$ for $z' = 0$, for all t', which gives an $A^{\mu}{}_{\nu}$ without minus signs:

$$\left.\begin{aligned} t &= \gamma(v_s)t' + v_s\gamma(v_s)z' \\ z &= \gamma(v_s)z' + v_s\gamma(v_s)t' \end{aligned}\right\} \Longrightarrow \frac{dz}{dt} = \frac{\gamma(v_s)v_b + v_s\gamma(v_s)}{\gamma(v_s) + v_s v_b\gamma(v_s)} = \frac{v_b + v_s}{1 + v_s v_b/c^2}.$$

Let m be the mass of a particle and let no one try to tell you it depends on the velocity, or the coordinate system. It is the "rest mass" — no sane person considers any other kind. Consider the 4-vector $p^{\mu} = mu^{\mu}$, where u^{μ} is the 4-velocity defined above. Clearly p^{μ} transforms as a contravariant tensor. Its spatial components are $\vec{p} = m\gamma(u)\vec{u}$, which we recognize as the relativistic form for the momentum. The zeroth component $p^0 = mc\gamma(u)$ is, we recognize, just the relativistic energy over c. This energy includes both the "rest energy" mc^2 and the kinetic energy. Thus $p^{\mu} = (E/c, \vec{p})$. The usual argument for this form of momentum and energy starts with a collision of equal mass particles in the center of mass frame, where momentum conservation requires the particles to be traveling in opposite directions with equal speeds, both before and after collision, and energy conservation requires the speed after is the same as the speed before. Then the sum of p^{μ} is conserved in this frame, and as each transforms as a 4-vector, the sum is automatically conserved in any frame.

As the momentum is a 4-vector, its appropriate square is an invariant. In fact,

$$p_{\mu}p^{\mu} = \frac{E^2}{c^2} - \vec{p}^2 = (mc)^2\gamma^2(u) + (mu)^2\gamma^2(u) = m^2c^2(1-\beta^2(u))\gamma^2(u) = m^2c^2.$$