Lecture 12 March 4, 2010

Galileo: if $\mathcal{O}$ uses $\vec{x}, t$, and $\mathcal{O}'$, moving at constant velocity $\vec{v}$ w.r.t. $\mathcal{O}$, uses $\vec{x}\,', t'$, where the coordinates of any event are related by

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

then the laws of physics are the same for both observers. Physics is about forces and their effect on motion. Ancient Greeks and Medieval philosophers would not agree with Galileo. They thought velocities required force. Newton said no, Forces give accelerations, and with

$$\vec{u}(t) := \frac{d\vec{x}}{dt}, \quad \vec{a}(t) := \frac{d\vec{u}}{dt}, \quad \text{and} \quad \vec{u}\,'(t) := \frac{d\vec{x}\,'}{dt} = \vec{u}(t) - \vec{v},$$

$$\text{so} \qquad \vec{a}\,'(t) := \frac{d\vec{v}\,'}{dt} = \vec{a}(t).$$

The observers agree on the acceleration and the forces! So Newton's laws are consistent with Galilean relativity.

Are Maxwell's?

The wave equation is not invariant:

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

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

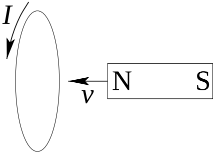
$$\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} = \nabla'^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t'^2} + \frac{2}{c^2}\frac{\partial}{\partial t}\vec{v}\cdot\vec{\nabla}\,' - \frac{1}{c^2}(\vec{v}\cdot\vec{\nabla}\,')^2.$$

Wave in sound not invariant, because the air is at rest in one frame, not others. Does Maxwell hold only in rest frame of ether? Michelson and Morley (reluctantly) say no.

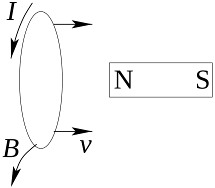
But E&M obeys relativity

Einstein says E&M obeys relativity. Consider a loop of wire and a bar magnet.

If loop at rest and magnet approaching, B and flux increasing with time, EMF generates current in loop.



If magnet at rest and loop moving towards it, charges in wire are moving right in a magnetic field with a component away from the loop center, so the Lorentz force causes current as shown.



Same current in either case — two different causes, but same effect. Einstein says: must be one theory good in either reference frame.

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.

No to Fitzgerald Contraction to explain M-M.

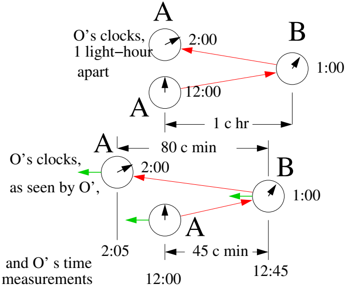
But these have strange consequences:

- Relativity of Simultaneity
- Length contraction
- Time dilation

Relativity of Simultaneity

A and B are at rest wrt $\mathcal{O}$, and are one light-hour apart. A sends a light signal to B at noon, B reflects it as his clock strikes 1, and it returns to A at 2PM. (of course A's clock and B's clock are synchronized).

But $\mathcal{O}'$ is watching all this, and has his own synchronized clocks. As B is travelling left as the light beam travels right, it takes less time to reach B, which it does at 12:45 on $\mathcal{O}'$'s clock.



As A is travelling away from the return beam, it takes 80 minutes to reach him, at 2:05. Not only does $\mathcal{O}'$ claim A's clock is running slow, but also that his clock striking 1 (at $1{:}02\frac{1}{2}$) and B's clock striking 1 (at 12:45) are not simultaneous, although $\mathcal{O}$ claims they are.

They disagree on the time an event takes place, $t' \neq t$.

The Lorentz Transformation

What gives $\mathcal{O}'$'s coordinates in terms of $\mathcal{O}$'s?
Homogeneity implies a linear relation, as displacements can't depend on the origin. As all observers agree light travels at the same speed c (in vacuum), we should measure distances in light-seconds, or time in meters/c, so $x^0 := ct$ has the same units as $\vec{x}$.
Let $x^i = (\vec{x})_i$, and use Greek index x^α, $\alpha = 0, 1, 2, 3$.
Then x'^α is linear in x^β, $x'^\alpha = \sum_\beta A^\alpha{}_\beta x^\beta$, where the 4×4 matrix A depends only on $\vec{v}$.
Two events along the history of a ray of light satisfy $|\Delta\vec{x}| = c\Delta t$, or

$$(\Delta s)^2 := c^2(\Delta t)^2 - (\Delta\vec{x})^2 = 0.$$

This quadratic form plays a fundamental role, and is called the invariant length, though it is not always positive.

More elegantly:[1]

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

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

Minkowski metric: $\eta_{00} = 1$, $\eta_{ij} = -\delta_{ij}$ and $\eta_{0i} = \eta_{i0} = 0$ for i and j from 1 to 3.

What's up with superscripts? x^α is called a *contra-variant vector*. Never sum over two contravariant indices! But *always* sum over a repeated covariant (subscript) and contravariant index. In fact, Einstein suggests you leave out the summation sign, so for example

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

According to Einstein, the suggestion that you can leave out the summation sign was his greatest contribution to human knowledge.

[1] Jackson uses $g_{\alpha\beta}$ where I use $\eta_{\alpha\beta}$.

Determining $A^\mu{}_\nu$

We now turn to finding the requirements for $A^\mu{}_\nu$ so that it represents a correct relationship between $\mathcal{O}$'s coordinates and $\mathcal{O}'$'s. This is best done by considering a number of clever *gedanken* experiments,

- Clock made of parallel mirrors: $\Longrightarrow$ time dilation
- Light pulse from center of railroad car hits ends: $\Longrightarrow$ non-simultaneity
- Rods perpendicular to relative motion crossing: $\Longrightarrow$ no transverse contraction
- Michelson-Morley: $\Longrightarrow$ Fitzgerald length contraction.

But you have seen these often in undergrad courses.
If not, you must read Smith.
We will present an abstract, mathematical treatment.

$(\Delta s')^2$ must be zero whenever $(\Delta s)^2$ is, as both agree on light travel speed. As $(\Delta s')^2 = \eta_{\alpha\beta} A^\alpha{}_\mu A^\beta{}_\nu \Delta x^\mu \Delta x^\nu$, the matrix $M_{\mu\nu} := \eta_{\alpha\beta} A^\alpha{}_\mu A^\beta{}_\nu$ must satisfy $M_{\mu\nu} x^\mu x^\nu = 0$ for any lightlike $x^\mu = (x^0, \pm\vec{x})$ with $x^0 = |\vec{x}|$. This tells us first that $M_{00}|\vec{x}|^2 \pm x^0 \sum M_{0i} x_i + \sum M_{ij} x_i x_j = 0$ for any vector $\vec{x}$.

Difference $\Longrightarrow M_{0i} = M_{i0} = 0$, and then $\sum M_{ij} x_i x_j$ depends only on the length and not the direction of $\vec{x}$, so M_{ij} is a multiple $\lambda(\vec{v})$ of the identity. The sum tells us $M_{00} = \lambda(\vec{v})$ as well, so $\eta_{\alpha\beta} A^\alpha{}_\mu A^\beta{}_\nu = \lambda(v) \eta_{\mu\nu}$, and $(\Delta s')^2 = \lambda(v)(\Delta s)^2$ for arbitrary Δx^μ.

Isotropy: λ depends only on $|\vec{v}|$, and Lorentz transforming $\mathcal{O} \to \mathcal{O}' \to \mathcal{O}$ tells us $(\lambda(v))^2 = 1$, so $\lambda(v) = 1$ (not -1 by continuity in v), and

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

So the condition for relating coordinates, for A being a Lorentz transformation, is

$$\eta_{\alpha\beta} A^{\alpha}{}_{\mu} A^{\beta}{}_{\nu} = \eta_{\mu\nu}. \tag{1}$$

As an example, suppose $\vec{v}$ is in the x direction.
$\mathcal{O}'$'s origin: $x'^{\mu}_{\mathcal{O}'} = (ct', 0, 0, 0)$ corresponds[2] to $x^{\mu}_{\mathcal{O}'} = (ct, vt, 0, 0)$. From $\eta_{\mu\nu} x'^{\mu} x'^{\nu} = \eta_{\mu\nu} x^{\mu} x^{\nu}$ we have $t' = t\sqrt{1 - v^2/c^2}$.

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

Now look 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.

[2] We assume $(0,0,0,0) \to (0,0,0,0)$. See lecture notes for why.

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 $i = 2, 3$) we see that $A^0{}_i = A^1{}_i = 0$.

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

$$A^{\mu}{}_{\nu} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (2)$$

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

β and γ are convenient parameters to use for relativistic transformations, but 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. For a rotation $\mathbf{O}$ is an orthogonal matrix,

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

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

Vectors

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

$$B \cdot C = \eta_{\mu\nu} B^\mu C^\nu, \quad \text{and } \textit{not} \quad \sum_\mu B^\mu C^\mu.$$

We can define, for every contravariant vector V^μ, a *covariant vector* $V_\mu := \eta_{\mu\nu} V^\nu$, $V_0 = V^0, V_i = -V^i$, with the same physical content[3].

[3] 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.

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 $\eta_{\mu\rho}$.
Note it is important to keep the order of indices unchanged.

If $\mathbf{L}$ is the abstract transformation from $\mathcal{O}$'s description of physics 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 (1) 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 $\mathcal{O}$,

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

Infinitesimal Generators

The matrices for rotations are orthogonal, with $\mathbf{R} = e^{i\sum_j \theta_j \mathbf{L}_j}$, with $\mathbf{L}_j$ imaginary antisymmetric matrices (angular momentum operators). In 3D $j = 1, 2, 3$ but in general D dimensions there are $D(D-1)/2$ generators. That $\mathbf{L}_j$ is antisymmetric can be seen for infinitesimal θ_j, with

$$\begin{aligned}
\mathbb{1}_{km} &= (\mathbf{R})_{k\ell}\left(\mathbf{R}^T\right)_{\ell m} = \mathbf{R}_{k\ell}\mathbf{R}_{m\ell} \\
&= \left(\delta_{k\ell} + i\sum \theta_j\,(\mathbf{L}_j)_{k\ell}\right)\left(\delta_{m\ell} + i\sum \theta_j\,(\mathbf{L}_j)_{m\ell}\right) \\
&= \delta_{km} + i\sum_j \theta_j\left[(\mathbf{L}_j)_{km} + (\mathbf{L}_j)_{mk}\right],
\end{aligned}$$

so $\mathbf{L}_j$ is antisymmetric.

For our Lorentz transformations, if an infinitesimal one is

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

the requirement (1) 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.

Tensors

Tensors have several Lorentz indices, each tranforms with ${A_\mu}^\nu$ or ${A^\mu}_\nu$.

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

If M is a tensor *field*, depending on x^μ, we have

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

where $\vec{x}'$ now means the four-position $({x'}^0, {x'}^1, {x'}^2, {x'}^3)$.
Isn't it time to discuss some **physics**?

So far just general transformation properties, but no physical quantities other than x^μ, spatial and temporal positions. Nonrelativistically motion is $\vec{x}(t)$, but why not consider the path in space-time, $x^\mu(\lambda)$, λ parameterizing the path, but not significant. Convenient choice is 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 $\gamma(t) = 1/\sqrt{1-\beta^2(t)}$ apply to the (time-dependent) speed of the particle and not some other inertial observer.
Newton: Forces $\frac{d\vec{p}}{dt}$ relate to acceleration $\vec{a} = \frac{d\vec{u}}{dt}$, but time is not a relativistic invariant. Instead of $\vec{u} = \frac{d\vec{x}}{dt}$, consider the *4-velocity*

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

where $\gamma(u) = 1/\sqrt{1-\vec{u}^{\,2}/c^2}$. τ is invariant, so u^μ transforms properly as a contravariant vector.

An example: Suppose $\mathcal{O}$ has purchased a gun which shoots bullets at velocity v_b, and he mounts this pointing forward on his spaceship which is moving in the $+x$ direction with velocity v_s with respect to the Earth, upon which $\mathcal{O}'$ sits. How fast does the bullet travel, according to $\mathcal{O}'$?

Note the $\vec{v}$ for Eq. (2) is $-v_s$ in the x direction, so $A^0{}_0 = A^1{}_1 = \gamma_{v_s}$, $A^0{}_1 = A^1{}_0 = \beta_{v_s}\gamma_{v_s}$ and $u^\mu = (c\gamma_{v_b}, \gamma_{v_b} v_b, 0, 0)$, so

$$\begin{aligned} u'^0 &= \gamma_{v_s}\, c\gamma_{v_b} + \beta_{v_s}\gamma_{v_s}\, \gamma_{v_b} v_b \\ u'^1 &= c\beta_{v_s}\gamma_{v_s}\, c\gamma_{v_b} + \gamma_{v_s}\, \gamma_{v_b} v_b \end{aligned}$$

so the 3-velocity according to $\mathcal{O}'$ is

$$\vec{u}' = cu'^1/u'^0 = c\frac{v_s\gamma_{v_s}\gamma_{v_b} + \gamma_{v_s}\,\gamma_{v_b} v_b}{\gamma_{v_s}\, c\gamma_{v_b} + \beta_{v_s}\gamma_{v_s}\,\gamma_{v_b} v_b} = \frac{v_s + v_b}{1 + v_s v_b/c^2}.$$

Many particles have a mass m. Never let anyone tell you the mass of a particle depends on its velocity! Define $p^\mu = m\, u^\mu$, which is clearly a contravariant 4-vector. This is the momentum of the particle. Its spatial components, $\vec{p} = m\gamma_u \vec{u}$ should be recognized as the relativistic form of the 3-momentum. Its zeroth component $p^0 = mc\gamma_u$, which you will recognize as $1/c$ times the relativistic form of the energy, including rest energy. so $p^\mu = (E/c, \vec{p})$. Note if any inertial observer finds that the total energy and momentum of a system is conserved, all observers will agree.

As p^μ and u^μ are 4-vectors, their (Minkowsky) squares are invariant.

$$\begin{aligned} p^2 &= \eta_{\mu\nu} p^\mu p^\nu = E^2/c^2 - \vec{p}^{\,2} = (mc)^2 \gamma^2(u) + (mu)^2 \gamma^2(u) \\ &= m^2 c^2 (1 - \beta^2(u)) \gamma^2(u) = m^2 c^2. \end{aligned}$$

and $u^2 = c^2 \gamma_u^2 - \vec{u}^{\,2} \gamma_u^2 = c^2 (1 - \beta_u^2) \gamma_u^2 = c^2$.