March 16

Special Relativity and Maxwell's Equations

In the previous lecture, we saw that a problem that is purely magnetostatic in one frame of reference exhibits electric forces when viewed in another frame of reference. We used this connection to derive the existence of magnetic forces from electrostatic and relativity.

It is not hard to understand how a current in one frame appears as a charge density in another. We know that a current is a part of a 4-vector. So if

$$J = (\rho c, \vec{j}) = (0, j\hat{3})$$

in one frame, these same quantities will appear in a frame boosted by $v\hat{3}$ as

$$(\rho c)' = \gamma(\rho c - \beta j^3) = -\frac{v/c}{\sqrt{1-v^2/c^2}}\, j$$

$$j^{3\prime} = \gamma(j^3 - \beta\rho c) = \frac{j}{\sqrt{1-v^2/c^2}}$$

In particular

$$\rho' = -\frac{j}{\sqrt{1-v^2/c^2}} \cdot \frac{v}{c^2}$$

so that the <u>mean</u> charge density in the new frame is

$$\rho'_{c/m} = -\frac{1}{\sqrt{1-v^2/c^2}} \; \frac{\underline{v}}{c^2} \cdot I$$

exactly as we saw last time.

So there is a more general reason why electrostatics implies the presence of additional forces when combined with special relativity. Electrostatics responds to ρ, which is a part of a 4-vector. Since the components of a 4-vector mix up when we make a change of frame, there is another frame in which the field is a response to $\vec{J}$, and so we need to have another field in the problem ~~that~~ couples to $\vec{J}$.

In this lecture, I would like to formalize this argument and use it to derive Maxwell's equations. In particular, I would like to show that Maxwell's equations are the simplest field equations consistent with special relativity in which the source of the field is a conserved current 4-vector J.

What do we mean when we ask that a set of equations be consistent with special relativity? Basically, we would like the laws of physics to be the same in every frame. To insure this for an equation $\quad \mathcal{O}\phi = \chi$

we would like to see

(1) If $\chi = 0$, if $\phi(x)$ is a solution of the equation, any <u>boost</u> of ϕ is also a solution

(2) If $\chi \neq 0$, if $\phi(x)$ is a solution of the equation with χ, then the <u>boost</u> of ϕ is a solution to the equation with the <u>boost</u> of χ

An equation with these properties is said to be <u>relativistically covariant</u>.

In the previous lecture, when we analyzed energy-momentum conservation, we also had the problem of insuring ~~that~~, if the conservation laws held in one frame, they would also hold for any boost of the situation. We found a simple method to insure that this would be true: Give to quantities to be boosted simple <u>transformation</u> laws under Lorentz transformations. I will now show how to apply this trick more systematically.

Let $\underline{X}$ be a 4-vector measured in a frame F, and let $\underline{X}'$ be the same 4-vector measured in F'. Then

$$\underline{X} = \Lambda \underline{X}'$$

Let's write this equation with indices $\mu, \nu = 0,1,2,3$:

$$\underline{X}^\mu = \Lambda^\mu{}_\nu (\underline{X}')^\nu$$

I have lowered the second index on Λ for a reason that I will now explain.

Let x or $x^\mu = (x^0 x^1 x^2, x^3)$ be the space and time coordinates. What is the relation between $\partial/\partial x^\mu$ and $\frac{\partial}{\partial x'^\mu}$.

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

In a notation free of indices: $\quad \frac{\partial}{\partial x} = (\Lambda^{-1})^T \frac{\partial}{\partial x'}$

This is a different transformation law from that of a 4-vector. In fact, it is just the dual relation. If J or J^μ is a 4-vector

$$\frac{\partial}{\partial x'^\mu} J'^\mu = \frac{\partial}{\partial x^\nu} \underbrace{\Lambda^\nu{}_\mu \ (\Lambda^{-1})^\mu{}_\sigma}_{= \underline{1}} J^\sigma = \frac{\partial}{\partial x^\nu} J^\nu$$

(Actually this equation gives the reason that the electric charge current J must be a 4-vector: If J is conserved in one frame, J should be conserved in any frame. This equation guarantees it.)

Quantities that transform from frame to frame according to

$$A'_\mu = A_\nu \ \Lambda^\nu{}_\mu$$

are called contravariant 4-vectors, as distinguished from quantities for which

$$B^{\mu\prime} = (\Lambda^{-1})^\mu{}_\nu B^\nu$$

which are our usual 4-vectors (covariant 4-vectors).

If A is contravariant and B is covariant

$$A_\mu B^\mu = A_0 B^0 + A_1 B^1 + A_2 B^2 + A_3 B^3$$

is independent of frame. On the other hand, if B and C are covariant

$$B^0C^0 - B^1C^1 - B^2C^2 - B^3C^3$$

is independent of frame. (We proved this earlier for the case in which B and C are identical, but it is true as long as B and C are 4-vectors.)

So the relation between covariant and contravariant 4-vectors must be very simple: If we multiply the space components of a covariant vector by (-1), we get a contravariant vector, and vice versa. I would like to notate this by writing covariant vector indices as raised, writing contravariant vector indices as lowered, and define the following matrices to raise and lower indices:

$$\eta^{\mu\nu} = \begin{pmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix} \qquad \eta_{\mu\nu} = \begin{pmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix}$$

$\eta^{\mu\nu}$ is the inverse of $\eta_{\mu\nu}$, as it must be.

$$B_\mu = \eta_{\mu\nu} B^\nu = (B^0, -B^1, -B^2, -B^3)$$

is contravariant;

$$A^\mu = \eta^{\mu\nu} A_\nu = (A_0, -A_1, -A_2, -A_3)$$

is covariant. This notation implies a relation between η

and a Lorentz transformation matrix Λ:

$$(B')_\mu = B_\nu \Lambda^\nu{}_\mu = B^\lambda \eta_{\lambda\nu} \Lambda^\nu{}_\mu$$

$$= \eta_{\mu\nu} (B')^\nu = \eta_{\mu\nu} (\Lambda^\nu)_\lambda B^\lambda$$

so $$\Lambda^T \eta = \eta \Lambda^{-1}$$

Let's check this relation for $\Lambda = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

$$\begin{pmatrix} \gamma & \beta\gamma & & \\ \beta\gamma & \gamma & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & & \\ \beta\gamma & -\gamma & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

equal! ✓

$$\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} \gamma & -\beta\gamma & & \\ -\beta\gamma & \gamma & & \\ & & 1 & \\ & & & 1 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & & \\ \beta\gamma & -\gamma & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

and it should work, because this relation is just

$$\Lambda^T \eta \Lambda = \eta$$

contract with $x^\mu \; x^\nu$

$$((\Lambda x)^0)^2 - ((\Lambda x)^1)^2 - ((\Lambda x)^2)^2 - ((\Lambda x)^3)^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$$

which is the relation by which we defined Lorentz transformations in the first place.

Contracting raised and lowered indices gives a Lorentz-invariant quantity. Some examples of Lorentz invariants are:

$$x^\mu \eta_{\mu\nu} x^\nu \quad (\text{often written} = x^2)$$

$$= (x^0)^2 - (\vec{x})^2 = s^2, \text{ the interval}$$

$$p^\mu \eta_{\mu\nu} p^\nu = \left(\frac{E}{c}\right)^2 - (\vec{p})^2 \quad \text{often written} = p^2$$

$$= \left(\frac{mc}{\sqrt{1-v^2/c^2}}\right)^2 - \left(\frac{m\vec{v}}{\sqrt{1-v^2/c^2}}\right)^2$$

$$= (mc)^2 \frac{1-v^2/c^2}{1-v^2/c^2} = (mc)^2$$

So for a particle, the quantity

$$p^2 = p^\mu p_\mu$$

is independent of frame and defines the <u>mass</u> of the particle

$$p^2 = m^2 c^2$$

Another apparently invariant quantity is

$$\frac{\partial}{\partial x^\mu} \eta^{\mu\nu} \frac{\partial}{\partial x^\nu}$$

$$\frac{\partial}{\partial x^\mu} \equiv \partial_\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, \vec{\nabla}\right)$$

$$\partial^\mu = \eta^{\mu\nu}\frac{\partial}{\partial x^\nu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\frac{\partial}{\partial x^1}, -\frac{\partial}{\partial x^2}, -\frac{\partial}{\partial x^3}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\vec{\nabla}\right)$$

then

$$\partial_\mu \partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 = \Box$$

the wave operator. So it is reasonable to think that the equation

$$\partial_\mu \partial^\mu \phi = 0$$

is Lorentz-invariant.

Let's check the Lorentz-invariance of this equation explicitly. If $\phi(x)$ is a field configuration on space-time, the boost of this field configuration is

$$\phi'(x) = \phi(\Lambda^\mu{}_\nu x^\nu)$$

Watch the way things transform: if $\phi(x)$ has a maximum at $x^\mu = a^\mu$, the boost should have a maximum at

$$x'^\mu = (\Lambda^{-1})^\mu{}_\nu a^\nu$$

and, since $\Lambda \cdot (\Lambda^{-1} a) = a$, the function $\phi(\Lambda x)$ has this property. Now check that the boosted function satisfies the equation

$$\partial_\mu \partial^\mu \phi' = \frac{\partial}{\partial x^\mu} \eta^{\mu\nu} \frac{\partial}{\partial x^\nu} \phi(\Lambda x)$$

Let $y = \Lambda x$, then $\frac{\partial}{\partial x^\mu} \phi(y) = \Lambda^\nu{}_\mu \frac{\partial}{\partial y^\nu} \phi(y)$

so $$\partial_\nu \partial^\nu \phi(\Lambda x) = \eta^{\mu\nu} \Lambda^\lambda{}_\mu \frac{\partial}{\partial y^\lambda} \Lambda^\sigma{}_\nu \frac{\partial}{\partial y^\sigma} \phi(y)$$

and, since $$\Lambda \eta \Lambda^{-1} = \eta$$

$$= \eta^{\lambda\sigma} \frac{\partial}{\partial y^\lambda} \frac{\partial}{\partial y^\sigma} \phi(y)$$

so indeed
$$\partial^2 \phi = 0 \Rightarrow \partial^2 \phi' = 0$$

Similarly, if ϕ is a solution to

$$\partial^2 \phi = \chi$$

and χ is a scalar field whose boost is $\chi' = \chi(\Lambda x)$

then $\phi'(x) = \phi(\Lambda x)$ solves

$$\partial^2 \phi' = \chi'.$$

Now, how do we build an equation with a vector source.

A 4-vector field is boosted by the relation:

$$J'^\mu(x) = (\Lambda^{-1})^\mu{}_\nu J^\nu(\Lambda x)$$

This is, if $J^\nu(x)$ has a maximum at $x = a$ where the function points in the direction of V^ν, then J'^μ has a maximum at $x' = \Lambda^{-1} x = a$ where J' points in the boosted direction.

A possible covariant equation is

$$\partial^2 A^\mu = a J^\mu$$

where a is a constant. This equation also have the property of Lorentz covariance: if $A^\mu(x)$ solves $\partial^2 A^\mu = 0$, $A'^\mu(x) = (\Lambda^{-1})^\mu{}_\nu A^\nu(\Lambda x)$ also solves this equation, and if

$$\partial^2 A^\mu = J^\mu$$

then A' solves

$$\partial^2 A'^\mu = J'^\mu$$

This equation, however, is second-order, and we might wish to find a first-order equation which generalizes

$$\vec{\nabla}\cdot\vec{E} = \frac{1}{\epsilon_0}\rho$$

This equation can be written $\qquad \mu = 0,1,2,3 \qquad i,j = 1,2,3$

$$\partial_i (E^i) = \frac{1}{c\epsilon_0} J^0$$

The right-hand side is the 0 component of a 4-vector. The left-hand side must then also be written as the 0 component of a 4-vector to insure that the two sides of the equation have the same transformation law. This can be done by identifying the E^i as components of a 4x4 matrix with two covariant 4-vector indices: $\quad F^{\mu\nu}$

such that

$$E^i = F^{i0}$$

If $F^{\mu\nu}$ is symmetric it has $\frac{4 \times 5}{2} = 10$ components. But then

$$\partial_\mu F^{\mu\nu}\Big|_{\nu=0} = \frac{1}{c}\frac{\partial}{\partial t} F^{00} + \frac{\partial}{\partial x^i} F^{i0}$$

which does not lead back to Gauss' law. An alternative is to make $F^{\mu\nu}$ antisymmetric. Then it has

$\frac{4 \times 3}{2} = 6$ independent components. Three of them are $E^i = F^{i0} = -F^{0i}$. The other three independent components are

$$F^{ij} = -F^{ji} \quad \text{for} \quad i \neq j$$

Thus, we have 6 fields obeying a set of equations

$$\partial_\mu F^{\mu\nu} = \frac{1}{c\epsilon_0} J^\nu$$

setting $\nu = 0$, $\partial_i F^{i0} = \frac{1}{c\epsilon_0} \rho c$

or $\vec{\nabla} \cdot \vec{E} = \frac{1}{\epsilon_0} \rho$

What if we set ~~[scribbled]~~ $\nu = 3$?

$$\partial_0 F^{03} + \partial_1 F^{13} + \partial_2 F^{23} = \frac{1}{c\epsilon_0} J^3 .$$

or, using $\quad \mu_0 = \frac{1}{\epsilon_0 c^2} \quad \partial_0 = \frac{1}{c}\frac{\partial}{\partial t}$

$$-\frac{1}{c}\frac{\partial}{\partial t} E^3 + \partial_1 F^{13} + \partial_2 F^{23} = c\mu_0 J^3$$

Let's define $\vec{B}$ by

$$F^{13} = c B^2 \qquad F^{23} = -c B^1$$

in general $\quad F^{ij} = -c\, \epsilon^{ijk} B^k$

then $\quad \partial_1 F^{13} + \partial_2 F^{23} = c\,(\partial_1 B^2 - \partial_2 B^1) = c(\vec{\nabla}\times\vec{B})^3$

and so the terms in

$$\partial_\mu F^{\mu\nu} = \frac{1}{c\epsilon_0} J^\nu$$

with $\nu = i$ are:

$$(\vec{\nabla}\times\vec{B})^i = \frac{1}{c^2}\frac{\partial E^i}{\partial t} + \mu_0 J^i$$

$\vec{B}$ satisfies $\vec{\nabla}\cdot\vec{B} = 0$. The relativistic generalization of this formula is written using the totally antisymmetric symbol with four indices

$$\epsilon^{\mu\nu\lambda\sigma} = \begin{cases} 1 & \mu\nu\lambda\sigma = 0123 \text{, or any even permutation} \\ -1 & \mu\nu\lambda\sigma = \text{an } \underline{\text{odd}} \text{ permutation of } 0123 \\ 0 & \text{any } \underline{\text{two}} \text{ of } \mu\nu\lambda\sigma \text{ are equal} \end{cases}$$

$\epsilon^{\mu\nu\lambda\sigma}$ is Lorentz invariant, since

$$\epsilon^{\mu\nu\lambda\sigma} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Lambda^\lambda{}_\gamma \Lambda^\sigma{}_\delta \underbrace{(\varepsilon')^{\alpha\beta\gamma\delta}}_{\text{also totally antisymmetric}}$$

$$= (\det \Lambda)(\varepsilon')^{\mu\nu\lambda\sigma}$$

but for a boost e.g. $\Lambda = \begin{pmatrix} \gamma & \beta\gamma & & \\ \beta\gamma & \gamma & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$

or a rotation e.g. $\Lambda = \begin{pmatrix} 1 & & & \\ & \cos\theta & -\sin\theta & \\ & \sin\theta & \cos\theta & \\ & & & 1 \end{pmatrix}$

$\det \Lambda = 1$, so $\varepsilon^{\mu\nu\lambda\sigma} = \varepsilon'^{\mu\nu\lambda\sigma}$. Now write

$$\varepsilon^{\mu\nu\lambda\sigma} \partial_\nu F_{\lambda\sigma} = 0$$

for $\mu = 0$ $\quad \partial_1 F^{23} + \partial_2 F^{31} + \partial_3 F^{21} = 0$

which is $\quad \partial_1 B^1 + \partial_2 B^2 + \partial_3 B^3 = \vec{\nabla} \cdot \vec{B} = 0$

for $\mu = 1$ $\quad \epsilon^{1023} = -1$

$$-\partial_0 F_{23} - \partial_2 F_{30} - \partial_3 F_{02} = 0$$

If we raise the indices on F

$$F_{23} = +F^{23} = -cB^1$$

$$F_{30} = -F^{30} = -E^3$$

$$F_{02} = -F^{02} = +E^2$$

so this equation takes the form

$$-\frac{1}{c}\frac{\partial}{\partial t}(-cB^1) + \partial_2 E^3 - \partial_3 E^2 = 0$$

or $$\frac{\partial}{\partial t} B^1 = -(\vec{\nabla}\times\vec{E})^1$$

so, the covariant equation

$$\partial_\mu F^{\mu\nu} = \frac{1}{c\epsilon_0} J^\nu$$

summarizes

$$\vec{\nabla}\cdot\vec{E} = \frac{1}{\epsilon_0}\rho \qquad \vec{\nabla}\times\vec{B} = \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t} + \mu_0\vec{J}$$

the covariant equation

$$\epsilon^{\mu\nu\lambda\sigma}\partial_\nu F_{\lambda\sigma} = 0$$

summarizes

$$\vec{\nabla}\cdot\vec{B} = 0 \qquad \frac{\partial\vec{B}}{\partial t} = -\vec{\nabla}\times\vec{E}$$

Let's do one more step. The equation

$$\epsilon^{\mu\nu\lambda\sigma}\partial_\nu F_{\lambda\sigma} = 0$$

is the integrability condition that allow us (in a simply connected region) to write

$$F_{\lambda\sigma} = c(\partial_\lambda A_\sigma - \partial_\sigma A_\lambda)$$

the 12 component of this equation is

$$F^{12} = -cB^3 = c(\partial^1 A^2 - \partial^2 A^1)$$

$$= c\left(-\frac{\partial}{\partial x^1} A^2 + \frac{\partial}{\partial x^2} A^1\right)$$

or

$$B^3 = (\vec{\nabla} \times \vec{A})^3$$

more generally, this equation generalizes the relation

$$\vec{\nabla} \cdot \vec{B} = 0 \quad \Rightarrow \quad \vec{B} = \vec{\nabla} \times \vec{A}$$

to 4-vectors. If we write

$$F^{\lambda\sigma} = c(\partial^\lambda A^\sigma - \partial^\sigma A^\lambda)$$

then the Maxwell's equations from $\epsilon^{\mu\nu\lambda\sigma} \partial_\nu F_{\lambda\sigma} = 0$ are automatically satisfied. The other Maxwell equations become

$$c \; \partial_\mu (\partial^\mu A^\nu - \partial^\nu A^\mu) = \frac{1}{\epsilon_0 c} J^\nu$$

or

$$\partial_\mu \partial^\mu A^\nu - \partial^\nu(\partial_\mu A^\mu) = \mu_0 J^\nu$$

This is very close to the equation at the top of p. 10. Actually, it is a better equation, because if we take ∂_ν of this equation, we find

$$\partial_\mu \partial^\mu \partial_\nu A^\nu - \partial_\nu \partial^\nu \partial_\mu A^\mu = 0 = \mu_0 \partial_\nu J^\mu$$

so the new equation for A^μ requires that its source be a conserved current. On the other hand, we can choose A^μ such that $\partial_\mu A^\mu = 0$ — this is a choice of <u>gauge</u> — and then

$$\Box A^\mu = J^\mu$$

so

Maxwell's equations are the natural relativistically covariant equations whose source is a conserved current.

We'll learn more about the foundations of Maxwell's equations, and more about their applications, next term.