MSci 4621 ELECTROMAGNETISM

LECTURE NOTES VII

In this section of the notes, we will turn to the discussion of the Lorentz covariant formulation of elecctrodynamics.

7.1 Special Relativity

The fact that the speed c of propagation of the electromagnetic waves predicted by Maxwell's equations is a universal constant, independent of the motion of the source or of the detector of the waves was the basis on which Einstein built the special theory of relativity (1905). The basic postulate is that the fundamental laws of physics have the same form no matter in which inertial frame they are expressed. What is *special* about the theory is that it restricts its formulation to *inertial frames of reference.* Since inertial frames are in uniform rectilinear motion with respect to one another, the laws of nature must preserve their form under the transformations of coordinates appropriate to passage from one inertial frame (K) to another (K'), where K' is moving uniformly in a straight line with respect to K.

Note that the fundamental laws of nature also preserve their form under that change of coordinates which results from a *rotation* of the frame of reference, which is most simply appreciated by expressing them using the notation of *vectors.* For then so long as both sides of an equation are together scalar, or vector, or whatever as the case may be, the effect of a rotation is the same on both sides, and what was true in one frame remains true in the rotated frame. The equation is *covariant* under rotations.

It is already clear that Maxwell's equations are indeed covariant under rotations, since we regard $\mathbf{E}$ and $\mathbf{B}$ as vectors. We will see that they are also covariant under the *boosts*, i.e., the transformations which transform one inertial (K) frame to another (K'). As usual, we suppose that the frames K and K' coincide at $t = t' = 0$, and consider a flash of light emanating from their common origin at the instant they coincide. Then the (spherical) wave front described in K by

$$x^2 + y^2 + z^2 = c^2t^2$$

will be described in K' by

$$x'^2 + y'^2 + z'^2 = c^2t'^2.$$

Homogeneity and isotropy of space and homogeneity of time require a linear relationship between (x', y', z', t') and (x, y, z, t). If we have the "standard orientation" of the axes, so that the frames are parallel, with their relative motion along the x- direction, then consistency of

$$x^2 + y^2 + z^2 = c^2t^2 \Leftrightarrow x'^2 + y'^2 + z'^2 = c^2t'^2$$

with linearity of the transformation gives

$$\begin{aligned} ct' &= \gamma(ct - \beta x) \\ x' &= \gamma(x - \beta ct) \\ y' &= y \\ z' &= z \end{aligned}$$

where $\gamma = \frac{1}{\sqrt{1-\beta^2}}$, $\beta = |\boldsymbol{\beta}|$, $\boldsymbol{\beta} = \frac{\mathbf{v}}{c}$, and $\mathbf{v}$ is the relative velocity of K' with respect to K. The inverse relations are

$$\begin{aligned} ct &= \gamma(ct' + \beta x') \\ x &= \gamma(x' + \beta ct') \\ y &= y' \\ z &= z'. \end{aligned}$$

For parallel axes, but when $\mathbf{v}$ is not necessarily along the x-direction, one has

$$
\begin{aligned}
ct' &= \gamma(ct - \boldsymbol{\beta}\cdot\mathbf{x}) \\
\mathbf{x}' &= \gamma\mathbf{x}_{||} + \mathbf{x}_{\perp} - \gamma\boldsymbol{\beta}ct
\end{aligned}
$$

where $\mathbf{x}_{||} = \frac{\mathbf{x}\cdot\boldsymbol{\beta}}{\beta^2}\boldsymbol{\beta}$ is the component of $\mathbf{x}$ parallel to $\mathbf{v}$ and $\mathbf{x}_{\perp} = \mathbf{x} - \mathbf{x}_{||}$ is the component perpendicular to $\mathbf{v}$.

7.2 Four-Vectors

We now introduce the notation

$$
\begin{aligned}
x^0 &= ct \\
x^1 &= x \\
x^2 &= y \\
x^3 &= z
\end{aligned}
$$

and define the *rapidity* variable ζ by $\gamma = \cosh\zeta$ (so that $\beta = \tanh\zeta$), and then find that the equations for the Lorentz transformation of the coordinates (with standard orientation of the axes) can be written as

$$
\begin{aligned}
x'^0 &= \cosh\zeta\, x^0 - \sinh\zeta\, x^1 \\
x'^1 &= -\sinh\zeta\, x^0 + \cosh\zeta\, x^1 \\
x'^2 &= x^2 \\
x'^3 &= x^3
\end{aligned}
$$

which is formally very similar to the transformation law for rotations (e.g., for a rotation through an angle θ about the x^3- or $z-$axis);

$$
\begin{aligned}
x'^0 &= x^0 \\
x'^1 &= \cos\theta\, x^1 + \sin\theta\, x^2 \\
x'^2 &= -\sin\theta\, x^1 + \cos\theta\, x^2 \\
x'^3 &= x^3.
\end{aligned}
$$

We call any set of four quantities V^μ, $\mu = 0, 1, 2, 3$ which transform in this fashion under boosts and rotations a *four-vector*, thus

$$
\begin{aligned}
V'^0 &= \cosh\zeta\, V^0 - \sinh\zeta\, V^1 \\
V'^1 &= -\sinh\zeta\, V^0 + \cosh\zeta\, V^1 \\
V'^2 &= V^2 \\
V'^3 &= V^3.
\end{aligned}
$$

Note that if V^μ and U^μ are the components of any two four-vectors, the combination

$$
V^0U^0 - V^1U^1 - V^2U^2 - V^3U^3
$$

is an *invariant*, that is to say it is unchanged when one replaces V by V' and U by U'. We can call this combination the *(Lorentz) scalar product* between the vectors, and write it as

$$
V\cdot U = V^0U^0 - \mathbf{V}\cdot\mathbf{U}.
$$

In particular

$$
dx^0dx^0 - dx^1dx^1 - dx^2dx^2 - dx^3dx^3 = c^2dt^2 - |d\mathbf{x}|^2
$$

is invariant, and we write it as ds^2 or as $c^2d\tau^2$.

Two 'nearby' *events* in space-time, separated by dx^μ are said to be *space-like* separated iff (= if and only if) $ds^2 < 0$, or *time-like* separated iff $ds^2 > 0$, or *null-* or *lightlike-* separated iff $ds^2 = 0$; and these notions are independent of the reference frame.

7.3 Time Dilation and the Lorentz-FitzGerald Contraction

If a particle moves with a 3-velocity $\mathbf{u}(t)$ with respect to a frame K, then in the time interval dt (as determined in K) it changes its position by $d\mathbf{x} = \mathbf{u}(t)\,dt$; so the space-time interval traversed is given by

$$\begin{aligned} ds^2 = c^2dt^2 - |d\mathbf{x}|^2 &= c^2dt^2 - \mathbf{u}^2dt^2 \\ &= c^2dt^2(1 - \frac{u^2}{c^2}) \\ &= c^2(1-\beta_u^2)dt^2, \end{aligned}$$

and this is invariant. Consider the instantaneous *rest-frame*, i.e., the frame in which the particle is instantanously at rest. In this frame, the time interval corresponding to dt is called $d\tau$, and evidently the space interval is zero, since the particle is at rest. Thus we have

$$d\tau = \sqrt{1-\beta_u^2}\,dt,$$

or

$$d\tau = \frac{dt}{\gamma_u}$$

which gives the relation between the time interval $d\tau$ (the so-called *proper* time) as measured by a clock moving with the particle, and the time interval dt as measured in the frame K in which the instantaneous speed of the particle is $u(t)$. Since $\gamma \geq 1$, we have

$$t_2 - t_1 = \int_{\tau_1}^{\tau_2} \frac{dt}{d\tau}\,d\tau = \int_{\tau_1}^{\tau_2} \gamma\,d\tau \geq \tau_2 - \tau_1,$$

so that *moving clocks run slow.* This is the phenomenon of *time dilation.*

Likewise, if we consider a rod of length L as measured in the frame K in which it is at rest, its end point may be taken to be at $x = 0,\ x = L$. Events which occur at the end-points thus have coordinates

$$\begin{aligned} x^0 &= ct & \qquad x^0 &= cT \\ x^1 &= 0 & x^1 &= L \\ x^2 &= 0 & x^2 &= 0 \\ x^3 &= 0 & x^3 &= 0. \end{aligned}$$

These same events will, in frame K' have coordinates

$$\begin{aligned} x'^0 &= \gamma ct & \qquad x'^0 &= \gamma(cT - \beta L) \\ x'^1 &= -\beta\gamma ct & x'^1 &= \gamma(L - \beta cT) \\ x'^2 &= 0 & x'^2 &= 0 \\ x'^3 &= 0 & x'^3 &= 0. \end{aligned}$$

The length of the rod as determined in K' is the distance between *simultaneous* positions of its end-points; i.e., one must set $x'^0 = \gamma ct = \gamma(cT - \beta L)$, and then the difference between the x'^1 coordinates of the end-points is

$$\begin{aligned} L' &= \gamma(L - \beta cT) - (-\beta\gamma ct) \\ &= \gamma L + \beta\gamma(ct - cT) \\ &= \gamma L + \beta\gamma(-\beta L) \\ &= \frac{1}{\gamma}L. \end{aligned}$$

And since $\gamma > 1$, this shows that $L' < L$, which is the *Lorentz-FitzGerald contraction.*

7.4 The Four-velocity

Since dx^μ is a four-vector, and $d\tau = \frac{1}{c}\sqrt{(dx^0)^2 - (d\mathbf{x})^2}$ is a scalar, it follows that

$$\begin{aligned}
\frac{dx^\mu}{d\tau} &= \frac{c\,dx^\mu}{\sqrt{(dx^0)^2 - (d\mathbf{x})^2}} \\
&= \frac{dx^\mu}{dt}\frac{1}{\sqrt{1 - \left(\frac{d\mathbf{x}}{dt}\right)^2 \frac{1}{c^2}}} \\
&= \frac{dx^\mu}{dt}\gamma_u
\end{aligned}$$

is a four-vector, where $\gamma_u = \frac{1}{\sqrt{1-\beta_u^2}}$, $\beta_u = \frac{u}{c}$, $u = |\mathbf{u}|$ and $\mathbf{u} = \frac{d\mathbf{x}}{dt}$. We write U^μ for this four-vector, giving

$$\begin{aligned}
U^0 &= \gamma_u\, c \\
\mathbf{U} &= \gamma_u\, \mathbf{u}.
\end{aligned}$$

For a particle with a velocity $\mathbf{u}$, (or better to say, 3-velocity $\mathbf{u}$), this four-vector $U = \gamma_u(c, \mathbf{u})$ is called its *four-velocity.* Note that $U^2 \equiv U \cdot U = \gamma_u^2(c^2 - \mathbf{u}^2) = c^2$.

7.5 Energy and Momentum

The non-relativistic definitions $\mathbf{p} = m\mathbf{u}$ of momentum and KE= $\frac{1}{2}mu^2$ of kinetic energy are replaced by

$$\mathbf{p} = m_0\mathbf{U} = m_0\gamma_u\mathbf{u} = m_u\mathbf{u}$$

together with

$$p^0 = m_0U^0 = m_0\gamma_u c = m_u c,$$

so that $p^0 c = m_u c^2$ which may be recognised as the total relativistic energy E for a particle of *rest-mass* m_0 with speed u. Thus $(E/c, \mathbf{p})$ is a four-vector $= m_0(U^0, \mathbf{U})$, with

$$p^2 = (E/c)^2 - \mathbf{p}^2 = m_0^2[(U^0)^2 - \mathbf{U}^2] = m_0^2U^2 = m_0^2c^2,$$

i.e.,

$$E^2 = \mathbf{p}^2c^2 + m_0^2c^4$$

so that

$$\begin{aligned}
E &= \sqrt{m_0^2c^4 + \mathbf{p}^2c^2} \\
&= m_0c^2 + \frac{1}{2}\frac{\mathbf{p}^2}{m_0} + \cdots
\end{aligned}$$

which shows that the total relativistic energy E has an expansion which leads off with the rest-mass contribution m_0c^2 and has as its next term the non-relativistic kinetic energy $\mathbf{p}^2/2m_0 = \frac{1}{2}m_0u^2$.

7.6 Covariant and Contravariant Vectors

The Lorentz transformation rule can be expressed as a matrix equation

$$\begin{pmatrix} V'^0 \\ V'^1 \\ V'^2 \\ V'^3 \end{pmatrix} = \begin{pmatrix} \cosh\zeta & -\sinh\zeta & 0 & 0 \\ -\sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} V^0 \\ V^1 \\ V^2 \\ V^3 \end{pmatrix}$$

or

$$V'^\mu = \sum_\nu \Lambda^\mu{}_\nu V^\nu$$

where the elements of the transformation matrix are

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

A very useful convention, known as the *Einstein summation convention*, is to omit the summation sign in the previous equation and to write simply

$$V'^\mu = \Lambda^\mu{}_\nu V^\nu,$$

it being understood that whenever an index is repeated, it should be summed over: and whenever an index is repeated it will *always* be once 'up', and once 'down'. This form of the transformation rule is valid for any Lorentz transformation, be it a boost or a rotation or a combination of such.

Any four-component object which transforms as

$$V^\mu \to V'^\mu = \Lambda^\mu{}_\nu V^\nu$$

is a four-vector, and such a four-vector is called a *contravariant* four-vector, and is always written with the index 'up'. This is to distinguish it from another kind of four-vector, which is written with the index 'down'. An example of this kind is given by the gradient of a scalar f. So if f is a scalar function, the set of its partial derivatives with respect to the coordinates

$$\partial_\alpha f \equiv \frac{\partial f}{\partial x^\alpha}$$

transforms as some sort of a vector – but not as a contravariant vector. This is clear from consideration of $df = \partial_\alpha f \; dx^\alpha$, which is of course a scalar, and is thus some sort of a scalar product between the vector whose components are dx^α and the gradient with components $\partial_\alpha f$. The transformation law can easily be derived from

$$\frac{\partial f}{\partial x'^\alpha} = \frac{\partial f}{\partial x^\beta}\frac{\partial x^\beta}{\partial x'^\alpha},$$

which is just the chain rule for differentiation. So the new transformation law is

$$U_\alpha \to U'_\alpha = U_\beta \frac{\partial x^\beta}{\partial x'^\alpha},$$

and a vector with this transformation law is called a *covariant* vector. It is also clear that we may write this as

$$U'_\alpha = U_\beta (\Lambda^{-1})^\beta{}_\alpha.$$

7.7 Tensors

Either kind of vector is an example of a more general object called a *tensor*. A tensor is something which has a *linear* transformation rule, in this case (for a Lorentz tensor) under the Lorentz group of transformations from one frame to another. The simplest kind of tensor is one for which the transformation says simply 'no change', thus

$$S \to S' = S,$$

and this is characteristic of a *scalar*, which may be called a tensor of rank zero. The *rank* of a tensor is the number of indices it carries. So a vector is a tensor of rank 1. (And we need to specify whether those indices are contravariant or covariant.) A *contravariant tensor of rank 2* is then a two-index quantity with both indices 'up', say $M^{\alpha\beta}$. Since each index ranges over the four possibilities (0,1,2,3), there are $4 \times 4 = 16$ components, and in general $M^{\alpha\beta} \neq M^{\beta\alpha}$. The transformation law for such a tensor is

$$M^{\alpha\beta} \to M'^{\alpha\beta} = \Lambda^\alpha{}_\gamma \, \Lambda^\beta{}_\delta \, M^{\gamma\delta}.$$

If $M^{\alpha\beta} = M^{\beta\alpha}$, the tensor is said to be *symmetric* and this symmetry is preserved in going from one frame to another. Similarly if $M^{\alpha\beta} = -M^{\beta\alpha}$, the tensor is said to be *antisymmetric*, or *skew symmetric*, and again this property is independent of the reference frame. A symmetric second rank tensor has only 10 independent components ($= 4 + \frac{4 \cdot 3}{2}$), an antisymmetric second rank tensor has six independent components. A *covariant* second rank tensor will be a two-index quantity like $F_{\mu\nu}$ with transformation law

$$F_{\mu\nu} \to F'_{\mu\nu} = F_{\rho\sigma} (\Lambda^{-1})^{\rho}{}_{\mu} \, (\Lambda^{-1})^{\sigma}{}_{\nu}.$$

Also to be encountered are *mixed* tensors of the second rank, like $D^{\alpha}{}_{\beta}$ with one contravariant and one covariant index, and the corresponding transformation rule

$$D^{\alpha}{}_{\beta} \to D'^{\alpha}{}_{\beta} = \Lambda^{\alpha}{}_{\gamma} D^{\gamma}{}_{\delta} (\Lambda^{-1})^{\delta}{}_{\beta}.$$

Of special interest is the tensor given by

$$\delta^{\alpha}_{\beta} = \begin{cases} 1, & \text{when } \alpha = \beta; \\ 0, & \text{otherwise.} \end{cases}$$

If these are the values of its components in the frame K, then in the frame K' they will be

$$\begin{aligned} \delta'^{\alpha}_{\beta} &= \Lambda^{\alpha}{}_{\gamma} \delta^{\gamma}_{\delta} (\Lambda^{-1})^{\delta}{}_{\beta} \\ &= \Lambda^{\alpha}{}_{\gamma} (\Lambda^{-1})^{\gamma}{}_{\beta} \\ &= \delta^{\alpha}_{\beta}, \end{aligned}$$

since thought of as a matrix, δ^{α}_{β} is just the unit matrix. This means that δ^{α}_{β} is an *invariant* tensor.

It is useful to think of the components V^{μ} of a contravariant vector arranged as a column

$$\begin{pmatrix} V^0 \\ V^1 \\ V^2 \\ V^3 \end{pmatrix},$$

and then the transformation law may be given by matrix multiplication rules as

$$V \to V' = \Lambda V.$$

Now the Lorentz transformations keep invariant the form

$$U \cdot V = U^0 V^0 - U^1 V^1 - U^2 V^2 - U^3 V^3 = U^{\alpha} \eta_{\alpha\beta} V^{\beta},$$

where we define

$$\eta_{\alpha\beta} = \begin{cases} +1, & \text{when } \alpha = \beta = 0; \\ -1, & \text{when } \alpha = \beta \neq 0; \\ 0, & \text{when } \alpha \neq \beta. \end{cases}$$

Thus for any U, V we have

$$U'^{\alpha} \eta_{\alpha\beta} V'^{\beta} = U^{\alpha} \eta_{\alpha\beta} V^{\beta},$$

so that

$$\Lambda^{\alpha}{}_{\gamma} U^{\gamma} \eta_{\alpha\beta} \Lambda^{\beta}{}_{\delta} V^{\delta} = U^{\gamma} \eta_{\gamma\delta} V^{\delta}$$

for every U, V which means that the coefficients of each and every Lorentz transformation have to satisfy

$$\Lambda^{\alpha}{}_{\gamma} \eta_{\alpha\beta} \Lambda^{\beta}{}_{\delta} = \eta_{\gamma\delta}.$$

Written in terms of matrices, this is

$$\Lambda^{\mathrm{T}} \eta \Lambda = \eta,$$

where Λ^{T} is the transpose of Λ. If η had been the unit matrix, this would be the condition that the matrix Λ was orthogonal; as it is, the matrix is said to be *pseudo-orthogonal.* A consequence, which we shall need later, is that the determinant of Λ is one. The condition on the coefficients $\Lambda^{\alpha}{}_{\beta}$ also states that $\eta_{\alpha\beta}$ may be regarded as the components of a constant second rank symmetrical covariant tensor.

Consider now any contravariant vector V^{μ}, and define V_{μ} by $V_{\mu} = \eta_{\mu\nu}V^{\nu}$, i.e., $V_0 = V^0, V_1 = -V^1, V_2 = -V^2, V_3 = -V^3$. It is then easy to see that V_{μ} transforms as a covariant four-vector. Thus the tensor $\eta_{\mu\nu}$ can be used to 'lower' a contravariant index, thereby giving a covariant index. In an exactly similar way, we may define the constant second rank symmetrical *contravariant* tensor with components $\eta^{\mu\nu}$ by

$$\eta^{\mu\nu} = \begin{cases} +1, & \mu = \nu = 0; \\ -1, & \mu = \nu \neq 0; \\ 0, & \mu \neq \nu. \end{cases}$$

This tensor can be used to 'raise' a covariant index.

Note also that $U \cdot V = U_{\alpha}V^{\alpha}$, which suggest arranging the components of a covariant vector as a row-matrix $(U_0\ U_1\ U_2\ U_3)$, and then the scalar product between the two vectors is also the matrix product

$$U \cdot V = U^{\alpha}\eta_{\alpha\beta}V^{\beta} = U_{\beta}V^{\beta} = \begin{pmatrix} U_0 & U_1 & U_2 & U_3 \end{pmatrix} \begin{pmatrix} V^0 \\ V^1 \\ V^2 \\ V^3 \end{pmatrix}.$$

Because the tensor $\eta_{\mu\nu}$ also enters into the formula

$$ds^2 = dx^{\mu}\eta_{\mu\nu}dx^{\nu},$$

it is called the *metric tensor.* In special relativity it is constant, and space-time is *flat.* But in general relativity the metric tensor is *not* constant; one has

$$ds^2 = dx^{\mu}g_{\mu\nu}(x)dx^{\nu},$$

and the metric tensor $g_{\mu\nu}$ determines the curvature of space-time. Constructed from the metric tensor and its derivatives is a tensor $G_{\mu\nu}$ which is determined through *Einstein's equations*

$$G_{\mu\nu} = \kappa T_{\mu\nu}$$

in terms of the density of energy and momentum which appear as the components of the stress-energy-momentum tensor $T_{\mu\nu}$. The constant κ equals $8\pi G/c$, where G is the (Newtonian) gravitational constant.