DEPARTMENT OF PHYSICS
INDIAN INSTITUTE OF TECHNOLOGY, MADRAS

PH3500 Classical Physics **16.11.2009**

4 Electromagnetism: a review

Contents: Recapitulation of Lorentz transformations; four-vectors and scalars; comparison of Lorentz and Coulomb gauges; Lorentz transformations of the electric and magnetic fields; form-invariance of Maxwell's equations under Lorentz transformations; invariants of the electromagnetic field.

We have seen that the motion of electric charges leads to magnetic effects. Maxwell's equations clearly show that electricity and magnetism are intimately related to one another. Most notably, a time-varying electric field leads to a magnetic field, and vice versa. It turns out that these equations, and hence the phenomenon of electromagnetism, have properties that are closely linked to the general principles on which special relativity is based. The unification of electricity, magnetism and optics emerges quite naturally from these considerations.

To understand the elements of these aspects, we first need to recall some basic facts about Lorentz transformations.

4.1 Special relativity

We start with a very brief summary of some relevant aspects of special relativity.

Nearly 100 years of observation and deduction support the following conclusions, valid in the absence of strong gravitational fields:
(i) The **principle** of relativity: **The laws governing physical phenomena are the same in form for mutually inertial observers**, i.e., observers whose space-time coordinate frames of reference are related to each other by rotations, constant velocity transformations ("boosts"), and shifts of the origin of space-time coordinates.
(ii) The **postulate** of relativity: **The speed of light in vacuum is a fundamental quantity that is the same for all mutually inertial observers**.

The statement in (i) is a general *principle* that is in fact already implicit in Newtonian mechanics. The (major) difference is of course the fact that there is no preferred set of frames in reality, as opposed to Newtonian relativity which assumes (erroneously) that there is an absolute fixed frame of reference ("a frame that is stationary with respect to the fixed stars").

Postulate (ii) is a new *physical* input that goes beyond Newtonian physics. It has profound consequences: e.g., the Lorentz transformation equations connecting the space-time coordinates of two mutually inertial frames, the law of

addition of velocities, the equivalence of mass and energy, and so on – in fact, all the consequences and relationships that follow in special relativity.

Before we go any further, the following clarification must be noted:
(a) The set of Lorentz transformations actually consists of the full set of arbitrary **rotations** of the spatial coordinates, *as well as* velocity transformations (**boosts**) to frames moving with any constant speed $u < c$ in an arbitrary direction. The latter alone are called "Lorentz transformations" in accounts of the subject at the high-school level, but we shall call them by their proper name, boosts.
(b) Two successive rotations of the spatial axes can be combined into a single resultant rotation. But two successive boosts in different directions cannot be combined into a single resultant boost. Rather, the resultant is a boost *together* with a rotation.

4.2 Lorentz transformation equations

The Lorentz transformation equations connecting the space-time coordinates $(\vec{r}, t) = (x, y, z, t)$ and $(\vec{r}\,', t') = (x', y', z', t')$ of two mutually inertial frames of reference S and S' are derived from the postulate of relativity. This postulate implies that we must have

$$(ct)^2 - \vec{r}\cdot\vec{r} = (ct')^2 - \vec{r}\,'\cdot\vec{r}\,' . \qquad (1)$$

Clearly this is satisfied if S and S' are related by a **rotation** of the spatial axes alone, because we then have $r = r'$, while the time coordinate remains unaffected, or $t = t'$. But a more general possibility also arises – namely, that S' is moving with respect to S with any constant velocity $\vec{u}$, where $u < c$: in other words, a **boost**.

When S' is obtained from S by a rotation of the spatial axes, for example a rotation about the z-axis (i.e., in the xy-plane) through an angle α, the respective space-time coordinates are related by

$$x' = x\cos\alpha + y\sin\alpha\,,\; y' = -x\sin\alpha + y\cos\alpha\,,\; z' = z\,,\; t' = t\,. \qquad (2)$$

On the other hand, when S' is obtained from S by a boost, for example along the z-direction by a velocity $\vec{u} = u\,\hat{e}_z$, the respective space-time coordinates are related by

$$x' = x\,,\; y' = y\,,\; z' = \gamma\,(z - ut)\,,\; t' = \gamma\,(t - \beta z/c)\,, \qquad (3)$$

where we have used the convenient notation

$$\beta = u/c \;\text{ and }\; \gamma = 1/\sqrt{1-\beta^2} = 1/\sqrt{1-u^2/c^2}\,. \qquad (4)$$

It is easy to verify that both these sets of transformations satisfy Eq. (1). The noteworthy point is of course that, in a boost, both space *and* time coordinates are affected, and get "mixed up" with each other. This is what leads to all the interesting consequences of special relativity such as length contraction, time dilatation, and so on.

4.3 Scalars and vectors: what they are, and why we need them

The elementary or high-school definition of a vector ("a quantity with both magnitude and direction") is rather unsatisfactory, to put it mildly.

We know that a more accurate definition is: **"A vector is a set of three quantities that transforms, under rotations of the coordinate axes, exactly as the set of coordinates itself transforms."** In other words: we know precisely how the coordinates $(x, y, z) = \vec{r}$ transform to $(x', y', z') = \vec{r}\,'$ under a rotation of the coordinate frame. $(A_x, A_y, A_z) = \vec{A}$ is a vector if, under a rotation of the coordinates axes, the new set $(A_x{}', A_y{}', A_z{}') = \vec{A}\,'$ is related to $\vec{A}$ in precisely the same way in which $\vec{r}\,'$ is related to $\vec{r}$.

A **scalar** is then defined as **a quantity that does not change under a rotation of the coordinate axes**. The most familiar example is of course $\vec{r}\cdot\vec{r} = r^2$, the square of the distance of a point from the origin. More generally, the dot product of two vectors, say $\vec{A}\cdot\vec{B}$, is a scalar – hence the name, "scalar product".

What is the need to write physical laws in terms of scalars, vectors and their generalizations (such as tensors)? The answer is that since these laws are unchanged or **form-invariant** under rotations of the coordinated axes, they must be *expressed in terms of quantities whose transformation properties are well-defined and prescribed.* This way the formulas carry their own dictionary, that enables us to go from one frame to another. When we say that $\vec{F} = m\vec{a}$, we **don't need to prescribe any particular coordinate system**. We are **guaranteed**, once it is specified that $\vec{F}$ is a vector, m is a scalar, and $\vec{a}$ is a vector, that the same law in another frame of reference would be just $\vec{F}\,' = m\vec{a}\,'$, and not anything else.

It is the **invariance** *of the laws under a given set of transformations that becomes manifest if they are expressed in terms of* **covariant** *quantities (i.e., quantities whose transformation properties are known).*

4.4 Four-vectors

But we have now generalized the possible transformations of the space-time coordinates to the full set of Lorentz transformations, rather than restricting ourselves to rotations of the spatial axes. We have asserted that the laws of physical phenomena are form-invariant under the full set of Lorentz transformations, rather than just the set of rotations. Therefore they must be expressed in terms of quantities with given transformation properties under *Lorentz transformations* (i.e., rotations as well as boosts). These quantities are the four-dimensional counterparts of the usual scalars and vectors just mentioned. To distinguish them from the latter, we could call them "Lorentz scalars", "Lorentz vectors", etc. The more usual terminology is **four-vector** instead of "Lorentz vector". And when no confusion is likely to arise, we just say "scalar" instead of "Lorentz scalar".

Thus we define: **a four-vector is a set of four quantities that changes under a Lorentz transformation in exactly the same way that the set of four space-time coordinates changes**. Now, the different components of a vector should of course have the same physical dimensions. For the space-time coordinates, the obvious way of taking care of this minor problem is by using the combination ct rather than t. Then ct, x, y and z all have the dimensions of length. They are combined into the four-vector

- space-time coordinates, $\underline{x} \equiv (ct, x, y, z) = (ct\,,\, \vec{r})$

To distinguish between an ordinary vector (a "three-vector") and a four-vector, we shall use an overhead arrow for the former and an underbar for the latter.

The wonderful fact (which shows that the principle of relativity is indeed borne out in nature) is that familiar physical quantities do fall into natural groupings that permit us to identify them as Lorentz scalars, four-vectors, Lorentz tensors, etc. We give (without proof) some examples of four-vectors that are relevant to our present purposes:

- energy-momentum or four-momentum of a particle, $\underline{p} = (E/c\,,\, \vec{p})$
- four-current density, $\underline{J} = (c\,\rho\,,\, \vec{J})$
- four-vector potential, $\underline{A} = (\Phi/c\,,\, \vec{A})$

We have adopted the convention of writing the "time-like" component of a four-vector first, and then the three "space-like" components. Note that the energy of a particle is *not a scalar in relativistic physics* – rather, it is the "time-like" component of a four-vector!

If the time ct and the space coordinates $\vec{r}$ form the four-vector $\underline{x}$, what about the derivatives with respect to these? It turns out that these, too, form the components of a four-vector – the four-dimensional counterpart of the gradient operator:

- four-dimensional gradient, $\underline{\partial} = (\partial/\partial(ct)\,,\, -\vec{\nabla})$

It is very important to note that the spatial components of the four-dimensional gradient are given by **minus** $\vec{\nabla}$, while the spatial components of $\underline{x}$ itself are given by $+\vec{r}$. The reason for this is technical, and we do not go into it here. But the form of the combination $(ct)^2 - \vec{r}\cdot\vec{r} = (ct)^2 - x^2 - y^2 - z^2$ that is unchanged under Lorentz transformations (see Eq. (1)) already shows that time and space are intrinsically somewhat different – note the relative minus sign in the combination.

4.5 Lorentz scalars

Given two four-vectors, it is natural to define a dot product between them to form a Lorentz scalar, i.e., a quantity that is unchanged under Lorentz transformations (rotations as well as boosts). Keeping in mind the minus sign in the expression $(ct)^2 - \vec{r}\cdot\vec{r}$, the appropriate *definition* of the scalar product of the four-vector $\underline{x}$ with itself is

$$\underline{x}\cdot\underline{x} = (ct,\,\vec{r})\cdot(ct,\,\vec{r}) = (ct)^2 - \vec{r}\cdot\vec{r}. \tag{5}$$

Unlike the square of a three-vector such as $\vec{r}$, the quantity in Eq. (5) can be positive, zero, or negative. What is guaranteed is that its numerical value will remain unchanged under Lorentz transformations. Any four-vector $\underline{u}$ is said to be

$$\left.\begin{array}{l}\text{time}-\text{like}\\ \text{light}-\text{like}\\ \text{space}-\text{like}\end{array}\right\}\ \text{depending on whether } \underline{u}\cdot\underline{u}\ \text{is}\ \left\{\begin{array}{l}>0\\ =0\\ <0\,.\end{array}\right.$$

The scalar product of two different four-vectors is defined similarly, *with a minus sign in between the product of the respective time components and the respective space components*. For example, taking two of the four-vectors defined above, we have

$$\underline{J}\cdot\underline{A} \equiv \rho\,\Phi - \vec{J}\cdot\vec{A}.$$

As a check, on the consistency of this definition of the scalar product, note that we automatically get

$$\underline{\partial}\cdot\underline{x} = \frac{\partial(ct)}{\partial(ct)} + \vec{\nabla}\cdot\vec{r} = 4,$$

the number of space-time dimensions, as expected.

The equation of continuity now takes on a very suggestive form: we find

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{J} \equiv \underline{\partial}\cdot\underline{J} = 0\,. \tag{6}$$

In other words, the (four-dimensional) divergence of the four-vector current density vanishes. But $\underline{\partial}\cdot\underline{J}$ is a Lorentz scalar, so that its value remains unchanged under Lorentz transformations! This leads to a deep conclusion: **electric charge is a scalar under Lorentz transformations**.

Similarly, recall that the Lorentz gauge condition is $\mu_0\,\epsilon_0\,\partial\Phi/\partial t + \vec{\nabla}\cdot\vec{A} = 0$. Since we know that $\mu_0\,\epsilon_0 = 1/c^2$, we recognise that the gauge condition is simply

$$\frac{\partial}{\partial(ct)}\frac{\Phi}{c} + \vec{\nabla}\cdot\vec{A} = \underline{\partial}\cdot\underline{A} = 0\,. \tag{7}$$

It is obvious that this is the four-dimensional analogue of the Coulomb gauge condition $\vec{\nabla}\cdot\vec{A} = 0$. The latter is unchanged under a rotation of the spatial coordinate axes, whereas **the Lorentz gauge condition is unchanged under the full set of Lorentz transformations**, because $\underline{\partial}\cdot\underline{A}$ is a Lorentz scalar. This is one of the main advantages of this gauge.

4.6 Use of the Lorentz gauge: Lorentz invariance of Maxwell's equations

We are now in a position to establish that the basic equations of electromagnetism, Maxwell's equations, are in fact unchanged under the full set of Lorentz transformations. The easiest way to do this is to work in the Lorentz gauge, because this gauge condition itself is unchanged under Lorentz transformations. Recall that in this gauge, both the scalar and vector potentials satisfy the wave equation: we have $\Box\,\Phi = \rho/\epsilon_0$ and $\Box\,\vec{A} = \mu_0\,\vec{J}$, where $\Box$ denotes the d'Alembertian or wave operator. But now, since we can write

$$\Box \equiv \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 = \underline{\partial}\cdot\underline{\partial}\,, \tag{8}$$

we recognise that **the d'Alembertian or wave operator is a Lorentz scalar** operator, just as the usual Laplacian $\nabla^2 = \vec{\nabla}\cdot\vec{\nabla}$ is a scalar operator under ordinary rotations of the spatial coordinates! Moreover, we have already asserted that $(\Phi/c\,,\,\vec{A})$ constitutes a four-vector $\underline{A}$, and that $(c\,\rho\,,\,\vec{J})$ also constitutes a four-vector, $\underline{J}$. Therefore **Maxwell's equations can be expressed in the completely Lorentz invariant (or relativistically invariant) form**

$$\begin{aligned}\underline{\partial}\cdot\underline{A} &= 0 \quad \text{(Lorentz gauge condition)}\\ \Box\,\underline{A} &= \mu_0\,\underline{J} \quad \text{(field equations)}\,.\end{aligned} \tag{9}$$

To repeat: if a Lorentz transformation is made to another frame of reference, the corresponding quantities in the new frame (denoted by primes to distinguish them from their counterparts in the original frame) are *guaranteed* to satisfy

$$\underline{\partial}'\cdot\underline{A}' = 0 \quad \text{and} \quad \Box\,'\,\underline{A}' = \mu_0\,\underline{J}'\,, \tag{10}$$

where of course $\Box = \Box\,'$ because it is a Lorentz scalar operator.

4.7 How do the physical fields $\vec{E}$ and $\vec{B}$ transform?

We now ask: while the scalar and vector potentials combine to give the four-vector potential, how do the physical electric and magnetic fields $\vec{E}$ and $\vec{B}$ themselves transform under Lorentz transformations? Together, they are made up of 6 components. Hence they cannot represent either a Lorentz scalar or a four-vector. The answer is most interesting. Recall that in three dimensions, the number of components of a tensor of rank 2 is $3^2 = 9$. Further, if the tensor is antisymmetric, the number of independent components is just 3, this being the value of $d(d-1)/2$ for $d = 3$. Similarly, in 4 dimensions, the number of independent components of an antisymmetric tensor of rank 2 is $(4\times 3)/2 = 6$. It turns out that **the components of $\vec{E}$ and $\vec{B}$ together constitute the components of an antisymmetric Lorentz tensor of rank 2, called the EM field tensor**.

We shall not go into this later. What is of immediate interest to us is the way $\vec{E}$ and $\vec{B}$ change under Lorentz transformations. Under a rotation of the spatial coordinates, of course, the answer is already known to us. Since $\vec{E}$ and $\vec{B}$ are ordinary three-vectors, they transform exactly like $\vec{r}$ does under such a rotation of the axes. In particular, the components of $\vec{E}$ do not get mixed up with those of $\vec{B}$, and vice versa. Under a boost, however, we may immediately expect such a mix up to happen – after all, we know that a moving charge (i.e., a current) does generate a magnetic field, and that a static charge will look like a moving charge to an observer who is moving with respect to it.

We now write down (without proof) the transformation rules for $\vec{E}$ and $\vec{B}$ under a boost from a frame S to a frame S', which is moving with an arbitrary velocity $\vec{u}$ ($u < c$) as seen from S. As usual, let unprimed and primed quantities denote variables in S and S' respectively. Further, let the subscripts $\|$ and $\perp$ denote components respectively along the direction of the boost $\vec{u}$ and perpendicular to it. Then:

$$E_\|{}' = E_\| , \quad \vec{E}_\perp{}' = \gamma\,(\vec{E}_\perp + \vec{u}\times\vec{B}_\perp) \tag{11}$$

$$B_\|{}' = B_\| , \quad \vec{B}_\perp{}' = \gamma\left(\vec{B}_\perp - \frac{\vec{u}\times\vec{E}_\perp}{c^2}\right) . \tag{12}$$

Here $\gamma = (1 - u^2/c^2)^{-1/2}$, as already defined. Several noteworthy points follow from these relations.

- The components of the fields along the direction of the boost are unaffected by the boost.
- It is the perpendicular components $\vec{E}_\perp$ and $\vec{B}_\perp$ that get mixed up with each other as a consequence of the boost.
- For sufficiently small boosts, such that $(u/c)^2$ (note the *square*) is negligible compared to unity, we have

$$\vec{E}\,' \approx \vec{E} + \vec{u}\times\vec{B} \quad \text{and} \quad \vec{B}\,' \approx \vec{B} - \frac{\vec{u}\times\vec{E}}{c^2} . \tag{13}$$

  In hindsight, these relations suggest how the Lorentz force on a moving charge arises – or, from another point of view, how the magnetic field itself is a natural consequence of charges in motion.
- It is easily verified that the combinations $\vec{E}\cdot\vec{B}$ and $\vec{E}^2 - c^2\vec{B}^2$ are unchanged under a Lorentz transformation, i.e., they are Lorentz scalars. These two combinations are called the **invariants of the EM field**.
- In particular, it follows from the previous statement that if $\vec{E}\cdot\vec{B} = 0$ in one frame of reference, it remains so for all frames obtained from it by Lorentz transformations. Therefore **transverse electromagnetic waves remain transverse electromagnetic waves for all mutually inertial observers!** This is only to be expected, given the postulate of relativity that we started out with.

4.8 Equation of motion of a charge in an EM field

Finally, consider a point charge of (rest) mass m_0 and charge q moving in an EM field. Allowing for the possibility that the particle may move with a speed that is not negligible compared to c, the correct equation of motion is

$$\frac{d}{dt}\left(\frac{m_0\vec{v}}{\sqrt{1 - v^2/c^2}}\right) = q(\vec{E} + \vec{v}\times\vec{B}) . \tag{14}$$

It is immediately clear that the acceleration of the particle, $d\vec{v}/dt$, is *not* equal to the force on the particle divided by its rest mass, in general. In relativistic mechanics it is more convenient to work with the momentum of the particle, as we know that its energy and momentum together constitute a four-vector.

4.9 Covariant form of Maxwell's equations

The Lorentz invariance of Maxwell's equations can be made manifest by first realising that the electric and magnetic fields transform as a second-rank antisymmetric tensor under Lorentz transformations. A simple and worthwhile exercise is to verify that the **electromagnetic field strength** given by[1]

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu , \tag{15}$$

with $\mu, \nu = 0, 1, 2, 3$ has components which can be identified with the electric and magnetic fields.

The four Maxwell's equations combine to form two equations. The first equation is

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c} J^\nu , \tag{16}$$

where $\partial_\mu = \eta_{\mu\nu}\partial^\mu$, with $\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$. Identify which of the two Maxwell's equations are obtained from the above equation. The other equation is an identity known as the **Bianchi identity**:

$$\epsilon_{\mu\nu\rho\sigma}\partial^\mu F^{\nu\rho} = 0 . \tag{17}$$

$\epsilon_{\mu\nu\rho\sigma}$ is the totally antisymmetric Levi-Civita tensor with $\epsilon_{0ijk} = \epsilon_{ijk}$ and i, j, k run over the spatial indices. Prove the Bianchi identity after identifying which of the two Maxwell's equations is obtained from it.

[1] In this subsection, we will explicit write the index rather than use the underbar notation. The index 0 is used for "time" and 1, 2, 3 for "space".