Classical Fields
Part I:
Relativistic Covariance

Prof. J.J. Binney

Oxford University

Michaelmas Term 2005

Books: (i) *Introduction to Einstein's Relativity*, Ray d'Inverno, OUP; (ii) *The Classical Theory of Fields*, L.D. Landau & E.M. Lifschitz (Pergamon); (iii) *Gravitation and Cosmology*, S., Weinberg (Wiley)

Vacation work: Study §1 Relativistic Covariance and work the eight embedded Exercises

1 Relativistic Covariance

Observers who move relative to one another do not always agree about the values of quantities, such as speed, mass, energy etc, associated with the same physical system. The special theory of relativity tells us how we may predict the values measured by any observer once we know the values assigned by one particular observer, for example ourselves.

Special relativity teaches us to think of experience as being made up of 'events', each with a definite location in the four-dimensional continuum of spacetime. Any given observer assigns to each event a unique 4-tuple of numbers (t, x, y, z). Of course he can do this in many, many ways. But special relativity claims that there are certain specially favoured systems for assigning coordinates to events, the so-called inertial coordinate systems. O chooses one inertial system and another observer, O$'$, sets up a different one. But according to special relativity the coordinates (t', x', y', z') O$'$ assigns to any event can be related to O's coordinates (t, x, y, z) of the same event by

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} ct_0 \\ x_0 \\ y_0 \\ z_0 \end{pmatrix} + \mathbf{L} \cdot \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}, \tag{1.1}$$

where c is the speed of light and (t_0, x_0, y_0, z_0) is a set of numbers characteristic of the two observers, as is the 4×4 matrix $\mathbf{L}$.

Clearly, (t_0, x_0, y_0, z_0) are the coordinates O$'$ assigns to the event that marks the origin of O's coordinates. For simplicity we shall assume that $(t_0, x_0, y_0, z_0) = \mathbf{0}$. In general $\mathbf{L}$ can be represented as the product of matrices generating a rotation, a boost parallel to a coordinate direction and a second rotation: $\mathbf{L} = \mathbf{R}' \cdot \mathbf{L}_0 \cdot \mathbf{R}$, where $\mathbf{R}$ rotates the coordinate axes so as to align the boost direction with a coordinate direction, $\mathbf{L}_0$ effects the boost along the given axis and $\mathbf{R}'$ rotates the coordinates to any desired final orientation. If $\mathbf{R}$ is chosen such that the x-axis becomes the boost direction, $\mathbf{L}_0$ has the form

$$\mathbf{L}_0 = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{where} \quad \begin{array}{c} \beta \equiv v/c \\ \gamma \equiv 1/\sqrt{1-\beta^2} \end{array}. \tag{1.2}$$

For simplicity we confine ourselves to observers whose spatial coordinate systems are aligned, and whose relative motion lies along their (mutually parallel) x-axes. Then in (1.1) $\mathbf{L} = \mathbf{L}_0$ and we get the familiar equations of a Lorentz transformation:

$$\begin{aligned} t' &= \gamma t - \gamma v x/c^2 \\ x' &= \gamma x - \gamma v t \\ y' &= y \\ z' &= z \end{aligned} \tag{1.3}$$

4-vectors Lorentz transformations mix up space and time, so it is useful to define new coordinates which all have dimensions of length. We write $x^0 \equiv ct$, $x^1 \equiv x$, $x^2 \equiv y$, $x^3 \equiv z$, and refer to a general component of the 4-vector (x^0, x^1, x^2, x^3) as x^μ. (The reason for labelling the components with superscripts rather than subscripts will emerge shortly.) Then we write a Lorentz transformation as

$$x^\mu = \Lambda^\mu{}_\nu x^\nu, \tag{1.4a}$$

where

$$\mathbf{\Lambda} \equiv \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \qquad (1.4b$$

In (1.4a) the **Einstein summation convention** is being used in that the summation sign $\sum_{\nu=0}^{1}$ has been omitted for brevity. You know it's really there because ν appears twice on the right-hand side of the equation, once up and once down.

Why do we write the row index of $\mathbf{\Lambda}$ as a superscript and the column index as a subscript?

A key property of a Lorentz transformation is that $-(ct')^2 + x'^2 + y'^2 + z'^2 = -(ct)^2 + x^2 + y^2 + z^2$. This is analogous to the fact that if two vectors $\boldsymbol{a}$ and $\boldsymbol{a}'$ are related by a rotation matrix, then $a_x'^2 + a_y'^2 + a_z'^2 = a_x^2 + a_y^2 + a_z^2$. So a Lorentz transformation is a sort of modified, four-dimensional rotation. When we rotate a vector $\boldsymbol{a}$ we like to say that the length $|\boldsymbol{a}|$ is invariant (i.e., stays constant). Analogously we define the length of the 4-vector $\mathbf{x}$ to be

$$|\mathbf{x}| \equiv -(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2. \tag{1.5}$$

Notes:

(i) We don't extract a square root because we have no guarantee that $|\mathbf{x}| \geq 0$.

(ii) 4-vectors that have negative lengths are called **time-like**, while those with positive lengths are **space-like**. Vectors with zero length are said to be **null**.

(iii) Every book on relativity uses a different convention. The sign of the lengths of space-like vectors is called the "signature of the metric".

The lengths of 4-vectors are sufficiently important for it to be useful to have a way of writing them that does not involve writing out all the components explicitly. To achieve this we introduce this matrix, called the **Minkowski metric**:

$$\boldsymbol{\eta} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.6}$$

Then we have

$$|\mathbf{x}| = \mathbf{x} \cdot \boldsymbol{\eta} \cdot \mathbf{x}, \tag{1.7a}$$

or in component form

$$|\mathbf{x}| = x^\mu \eta_{\mu\nu} x^\nu. \qquad (1.7b$$

The Einstein convention is here being used to drop two summation signs. We write both of $\boldsymbol{\eta}$'s indices as subscripts so that each sum is over one up and one down index.

Covariant and contravariant vectors We write the result of matrix multiplication of $\mathbf{x}$ by $\boldsymbol{\eta}$ as

$$x_\mu \equiv \eta_{\mu\nu} x^\nu .$$

We have $x_0 = -x^0 = -ct$, $x_1 = x^1$, $x_2 = x^2$ and $x_3 = x^3$. Thus the length of $\mathbf{x}$ is

$$x^\mu x_\mu = -c^2t^2 + x^2 + y^2 + z^2 .$$

Notice that here as everywhere else, we are summing over one up and one down index. In order to stick rigidly to this rule, we define

$$\eta^{\mu\nu} \equiv \eta_{\mu\nu} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{1.8}$$

Note:

We have $\eta^{\mu\gamma}\eta_{\gamma\nu} = \delta^\mu_\nu$, or in matrix form $\boldsymbol{\eta} \cdot \boldsymbol{\eta} = \mathbf{I}$, where $\mathbf{I}$ and δ^μ_ν are two ways of writing the 4×4 identity matrix. Also $\eta^{\mu\nu} = \eta^{\mu\gamma}\delta^\nu_\gamma$, so in a sense $\boldsymbol{\eta}$ is merely the up-up and down-down forms of the identity matrix.

From x_μ we can recover x^μ;

$$x^\mu = \eta^{\mu\nu} x_\nu . \tag{1.9}$$

x_μ is a 4-vector, but of a slightly different type than x^μ, because under a Lorentz transformation we have

$$\begin{aligned} x'_\mu &= \eta_{\mu\nu} x'^\nu = \eta_{\mu\nu}\Lambda^\nu{}_\kappa x^\kappa = \eta_{\mu\nu}\Lambda^\nu{}_\kappa \eta^{\kappa\lambda} x_\lambda \\ &= \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \equiv \Lambda_\mu{}^\nu x_\nu , \end{aligned} \tag{1.10}$$

where we have defined a new matrix

$$\Lambda_\mu{}^\lambda \equiv \eta_{\mu\nu}\Lambda^\nu{}_\kappa \eta^{\kappa\lambda} . \tag{1.11}$$

Notice that the transpose of $\Lambda_\mu{}^\nu$ is the inverse of $\Lambda^\mu{}_\nu$:

$$\Lambda^\mu{}_\kappa \Lambda_\mu{}^\nu = \delta^\nu_\kappa , \tag{1.12}$$

where we have again written the 4×4 identity matrix as δ^ν_κ.

Exercise (1):

Obtain (1.12) from the requirement that for any two vectors $\mathbf{x}$, $\mathbf{y}$, we have $x'_\mu y'^\mu = x_\mu y^\mu$.

Vectors with their indices below are called **covariant** (x_μ). Vectors with indices above are called **contravariant** (x^μ). I shall call them down and up vectors. The operation of setting two indices equal and summing from 0 to 3 is called **contraction**. In a contraction one index must be up and one down. Quantities like $\sum_\mu x_\mu x_\mu$ have nothing to do with physics. An important motivation for writing x^μ rather than $\mathbf{x}$ is to distinguish the up from the down form of $\mathbf{x}$. Often an expression is equally valid for up or down vectors provided the basic rules are obeyed, and then it is neater to use conventional vector notation than to stick in indices. For example, if $\mathbf{a}$ and $\mathbf{b}$ are vectors and $\mathbf{M}$ is a matrix, we can interpret $\mathbf{a} = \mathbf{M} \cdot \mathbf{b}$ as $a^\mu = M^{\mu\nu} b_\nu$, as $a_\mu = M_{\mu\nu} b^\nu$, or in yet other ways. But if you ever express a 4-vector in component form, you *must* come clean and say whether you're giving the up or the down vector, as in $x^\mu = (ct, x, y, z)$.

According to special relativity, all quantities of physical interest can be grouped into n-tuples.

1.1 1-tuples (4-scalars)

On some things all observers agree, for example the charge and total spin of the an electron. These quantities are called **4-scalars** or relativistic invariants. The length of a 4-vector is a 4-scalar.

1.2 4-tuples (4-vectors)

If O measures the wave-vector and frequency of a photon to be $\boldsymbol{k}$ and ω, then an observer O$'$ who moves at speed v along O's x-axis measures wave-vector $\boldsymbol{k}'$ and frequency ω' given by

$$\begin{pmatrix} \omega'/c \\ k'_x \\ k'_y \\ k'_z \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \omega/c \\ k_x \\ k_y \\ k_z \end{pmatrix}. \tag{1.13a}$$

The matrix form of this equation is

$$\begin{pmatrix} \omega'/c \\ \boldsymbol{k}' \end{pmatrix} = \mathbf{\Lambda} \cdot \begin{pmatrix} \omega/c \\ \boldsymbol{k} \end{pmatrix} \quad \text{where} \quad \mathbf{\Lambda} \equiv \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.13b}$$

Notes:

(i) The Lorentz transformation matrix $\mathbf{\Lambda}$ is dimensionless, so ω has to be divided by c to give the same dimensions as $\boldsymbol{k}$ before being put into the last place of a 4-vector with $\boldsymbol{k}$.

(ii) Vectors written in italic boldface ($\boldsymbol{k}$) are 3-vectors, while those written in Roman boldface ($\mathbf{k}$) are 4-vectors.

If we define $k^0 \equiv \omega/c$, then

$$\mathbf{k}' = \boldsymbol{\Lambda} \cdot \mathbf{k} \quad \text{i.e.,} \quad k'^{\mu} = \Lambda^{\mu}{}_{\nu} k^{\nu}. \tag{1.14}$$

Exercise (2):

Determine whether the photon is blue or red shifted between its emission by O and its detection by O′. Relate this to the question of whether O′ is approaching or receding from O.

The length of a photon's 4-vector is the scalar

$$|\mathbf{k}| \equiv -(k^0)^2 + (k^1)^2 + (k^2)^2 + (k^3)^2 = -\frac{\omega^2}{c^2} + |\boldsymbol{k}|^2 = 0.$$

One can prove that this really is a scalar by brute force:

$$\begin{aligned}
|\mathbf{k}'| &= -(k'^0)^2 + (k'^1)^2 + (k'^2)^2 + (k'^3)^2 \\
&= -\left(\gamma\frac{\omega}{c} - \beta\gamma k^1\right)^2 + \left(-\beta\gamma\frac{\omega}{c} + \gamma k^1\right)^2 + (k^2)^2 + (k^3)^2 \\
&= -\gamma^2\left(1-\beta^2\right)\frac{\omega^2}{c^2} + \gamma^2\left(1-\beta^2\right)(k^1)^2 + (k^2)^2 + (k^3)^2 \\
&= -(k^0)^2 + (k^1)^2 + (k^2)^2 + (k^3)^2.
\end{aligned}$$

Another familiar 4-tuple: if observer O measures energy E and momentum $\mathbf{p}$ for some particle, then O′ will measure E' and $\boldsymbol{p}'$ given by

$$\begin{pmatrix} E'/c \\ \boldsymbol{p}' \end{pmatrix} = \boldsymbol{\Lambda} \cdot \begin{pmatrix} E/c \\ \boldsymbol{p} \end{pmatrix}, \tag{1.15}$$

or setting $p^0 \equiv E/c$, we have $p'^{\mu} = \Lambda^{\mu}{}_{\nu} p^{\nu}$.

The length of the momentum-energy 4-vector of a particle of rest mass $m_0 \neq 0$ is just $-c^2$ times the square of its rest mass m_0. We show this by arguing that it doesn't matter in whose frame we evaluate a scalar. We choose the particle's rest frame. Then $\boldsymbol{p} = 0$ and $E = cp^0 = m_0 c^2$, so

$$-(p^0)^2 + (p^1)^2 + (p^2)^2 + (p^3)^2 = -m_0^2 c^2.$$

1.3 6-tuples (antisymmetric 2nd rank tensors)

If the electric and magnetic fields measured by O are arranged into the antisymmetric matrix $\mathbf{F}$,

$$F^{\mu\nu} \equiv \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{pmatrix} \quad \text{(SI units)}, \tag{1.16}$$

then O′ will measure E' and B' as

$$\begin{pmatrix} 0 & {E'}_x/c & {E'}_y/c & {E'}_z/c \\ -{E'}_x/c & 0 & {B'}_z & -{B'}_y \\ -{E'}_y/c & -{B'}_z & 0 & {B'}_x \\ -{E'}_z/c & {B'}_y & -{B'}_x & 0 \end{pmatrix} \equiv F'^{\mu\nu} = \Lambda^\mu{}_\kappa \Lambda^\nu{}_\lambda F^{\kappa\lambda}. \tag{1.17}$$

Note that $F^{\mu\nu}$ transforms *as if* it were the product $p^\mu p^\nu$ of two down-vectors (which it isn't). Objects that transform in this way are called second-rank tensors.

F is called the **Maxwell field tensor**.

Exercise (3):

Transform $F^{\kappa\lambda}$ with the matrix $\Lambda^\mu{}_\nu$ defined by (1.13b) to show that an observer who moves at speed v down the x-axis of an observer who sees fields $\boldsymbol{E} = (E_x, E_y, 0)$ and $\boldsymbol{B} = 0$, perceives fields $\boldsymbol{E}' = (E_x, \gamma E_y, 0)$ and $\boldsymbol{B}' = (0, 0, \gamma v E_y/c)$. [Hint: since $\mathbf{\Lambda}$ is symmetric, we can write $\mathbf{F}' = \mathbf{\Lambda}\cdot\mathbf{F}\cdot\mathbf{\Lambda}$.] Hence deduce the general rules $\boldsymbol{E}'_\| = \boldsymbol{E}_\|$, $\boldsymbol{E}'_\perp = \gamma(\boldsymbol{E}_\perp + \boldsymbol{v}\times\boldsymbol{B})$, $\boldsymbol{B}'_\| = \boldsymbol{B}_\|$, $\boldsymbol{B}_\perp = \gamma(\boldsymbol{B}_\perp - \boldsymbol{v}\times\boldsymbol{E}/c^2)$. Verify that $(B^2 - E^2/c^2) = ({B'}^2 - {E'}^2/c^2)$.

Some 6-tuples correspond to elements of area. This correspondence works as follows. With any two displacements, say **u** and **v**, we associate the parallelogram bounded by **u** and **v**. Information about the size and orientation of this parallelogram is conveyed by the antisymmetric tensor $S^{\alpha\beta} \equiv u^\alpha v^\beta - u^\beta v^\alpha$; in particular, if $\mathbf{u} = \mathbf{v}$, then $\mathbf{S} = 0$. **S** has fewer degrees of freedom than the eight numbers involved in **u** and **v** because we can add to **u** any multiple of **v** without affecting **S**, and vice versa for **v** and **u**.

Exercise (4):

Consider transformation $\mathbf{u} \to \mathbf{u}' = a\mathbf{u} + b\mathbf{v}$, $\mathbf{v} \to \mathbf{v}' = c\mathbf{u} + d\mathbf{v}$ with the corresponding mapping $\mathbf{S} \to \mathbf{S}'$. Show that the equation $\mathbf{S}' = \mathbf{S}$ imposes one constraint on the numbers a, b, c, d. Hence only $8 - 3 = 5$ numbers are needed to specify **S**. Give a geometrical interpretation of this result.

In three-space the size and orientation of a parallelogram may be specified by giving the magnitude and direction of the normal. Hence in three-space full information about an antisymmetric 2nd rank tensor can be packed into the three components of the 3-vector which we call the cross-product of the parallelogram's sides. In four-dimensional spacetime each parallelogram has a magnitude and two mutually perpendicular normals, requiring five numbers for its full specification. Consequently there is no direct analogue of the cross product and we must represent areas directly with antisymmetric tensors.

Exercise (5):

Relate the above statements to the number of independent components of an antisymmetric $n \times n$ matrix for $n = 2, 3, 4$.

A physically interesting 6-tuple that describes an area is the tensor $(x^\mu p^\nu - x^\nu p^\mu)$ formed from the space-time coordinate vector $x^\mu = (ct, x, y, z)$ and the 4-momentum of a particle. If the angular momentum about the origin is $\boldsymbol{L}$, we have

$$H^{\mu\nu} \equiv (x^\mu p^\nu - x^\nu p^\mu) = \begin{pmatrix} 0 & \ddots & \ddots & \\ c(xE/c^2 - tp_x) & 0 & \ddots & \ddots \\ c(yE/c^2 - tp_y) & -L_z & 0 & \ddots \\ c(zE/c^2 - tp_z) & L_y & -L_x & 0 \end{pmatrix}, \tag{1.18}$$

where the diagonal dots stand for minus the quantities in the lower left triangle of the matrix. The numbers in the first column of this matrix give mc times the particle's initial position vector.

With every 6-tuple we get two free scalars. If the 6-tuple is of the form $(u^\alpha v^\beta - u^\beta v^\alpha)$, then one of these is twice the squared magnitude of the corresponding parallelogram:

$$\begin{aligned} S^{\mu\nu}(\eta_{\mu\kappa}\eta_{\nu\lambda}S^{\kappa\lambda}) &\equiv S^{\mu\nu}S_{\mu\nu} = -\operatorname{Tr}\mathbf{S}\cdot\mathbf{S} \\ &= (u^\mu v^\nu - u^\nu v^\mu)(u_\mu v_\nu - u_\nu v_\mu) = 2[|\mathbf{u}||\mathbf{v}| - (\mathbf{u}\cdot\mathbf{v})^2]. \end{aligned}$$

Note:

Here by $\operatorname{Tr}\mathbf{M}$ we mean $M_\alpha{}^\alpha = M^\alpha{}_\alpha$. That is, the sum implied by Tr must always be over one up and one down index.

Evaluation in the particle's rest frame shows that the scalar $\frac{1}{2}H_{\mu\nu}H^{\mu\nu} = [|\mathbf{x}||\mathbf{p}| - (\mathbf{x}\cdot\mathbf{p})^2] = -(m_0 c r_0)^2$, where r_0 is the distance (in the rest frame) between the particle and the origin at $t = 0$.

It is interesting to evaluate this same scalar for the Maxwell field tensor. Straightforward matrix multiplication shows that the down-down shadow of $F^{\mu\nu}$ is[1]

$$F_{\mu\nu} \equiv \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{pmatrix} \quad \text{(SI units)}, \tag{1.19}$$

Multiplying each element of $F_{\mu\nu}$ by the corresponding element of $F^{\mu\nu}$ we find

$$\begin{aligned} m &\equiv \tfrac{1}{2}F_{\mu\nu}F^{\mu\nu} = -\tfrac{1}{2}\operatorname{Tr}\mathbf{F}\cdot\mathbf{F} \\ &= \tfrac{1}{2}(\text{each element of } F_{\mu\nu}) \times (\text{corresponding element of } F^{\mu\nu}) \\ &= (B^2 - E^2/c^2). \end{aligned} \tag{1.20}$$

To extract another scalar from a 6-tuple we need to introduce the **Levi-Civita symbol**:

$$\epsilon^{\alpha\beta\gamma\delta} = \begin{cases} +1 & \text{if } \alpha\beta\gamma\delta \text{ is an even permutation of 0123} \\ -1 & \text{if } \alpha\beta\gamma\delta \text{ is an odd permutation of 0123} \\ 0 & \text{otherwise.} \end{cases} \tag{1.21}$$

[1] It is worth remembering that in special relativity the lowering operation only *changes the sign of the mixed space-time components.*

Note:

Whereas when n is odd, the cyclic interchange $i_1 \to i_2 \to \ldots \to i_{n-1} \to i_n \to i_1$ is an even permutation of the i_k, when n is even, this permutation is odd. (To prove this exchange i_1 and i_n and then make $n-2$ exchanges to work i_1 back to the second place.) So whereas for 3-dimensional tensors $\epsilon_{jki} = \epsilon_{ijk}$, we now have $\epsilon^{\beta\gamma\delta\alpha} = -\epsilon^{\alpha\beta\gamma\delta}$.

$\epsilon^{\alpha\beta\gamma\delta}$ allows us to form the **dual** $\overline{\mathbf{F}}$ of $\mathbf{F}$:

$$\begin{aligned}\overline{F}^{\alpha\beta} &\equiv \tfrac{1}{2}\epsilon^{\alpha\beta\gamma\delta}F_{\gamma\delta} \\ &= \begin{pmatrix} 0 & B_x & B_y & B_z \\ -B_x & 0 & -E_z/c & E_y/c \\ -B_y & E_z/c & 0 & -E_x/c \\ -B_z & -E_y/c & E_x/c & 0 \end{pmatrix}.\end{aligned} \tag{1.22}$$

$\overline{\mathbf{F}}$ can be obtained from $\mathbf{F}$ by the transformation $\boldsymbol{E} \to \boldsymbol{B}$, $\boldsymbol{B} \to -\boldsymbol{E}$. The other scalar is the trace of the product of $\mathbf{F}$ with its dual:

$$\begin{aligned} f &\equiv \mathrm{Tr}\,\mathbf{F}\cdot\overline{\mathbf{F}} \\ &= -(\text{each element of } F_{\alpha\beta}) \times (\text{corresponding element of } \overline{F}^{\alpha\beta}) \\ &= \frac{4}{c}\boldsymbol{E}\cdot\boldsymbol{B}. \end{aligned} \tag{1.23}$$

Exercise (6):

Show that with $S_{\mu\nu} = u_\mu v_\nu - u_\nu v_\mu$, $\mathrm{Tr}\,\mathbf{S}\cdot\overline{\mathbf{S}} = 0$. This result explains why $\mathbf{S}$ has only 5 degrees of freedom (Exercise 4).

1.4 10-tuples (symmetric 2nd rank tensors)

Imagine that we move some charges around. Then the rate at which we do work *on* the e.m. field is

$$\begin{aligned}\dot{\mathcal{E}} &= -\int \boldsymbol{E}\cdot\boldsymbol{j}\,\mathrm{d}^3\boldsymbol{x} \\ &= -\frac{1}{\mu_0}\int \boldsymbol{E}\cdot\Big(\nabla\times\boldsymbol{B} - \frac{1}{c^2}\frac{\partial\boldsymbol{E}}{\partial t}\Big)\,\mathrm{d}^3\boldsymbol{x}\end{aligned} \tag{1.24}$$

But $\nabla\cdot(\boldsymbol{E}\times\boldsymbol{B}) = \boldsymbol{B}\cdot(\nabla\times\boldsymbol{E}) - \boldsymbol{E}\cdot(\nabla\times\boldsymbol{B})$, so (1.24) can be rewritten

$$\begin{aligned}\dot{\mathcal{E}} &= \frac{1}{\mu_0}\int \nabla\cdot(\boldsymbol{E}\times\boldsymbol{B})\,\mathrm{d}^3\boldsymbol{x} + \frac{1}{\mu_0}\int\Big(-\boldsymbol{B}\cdot(\nabla\times\boldsymbol{E}) + \frac{1}{c^2}\boldsymbol{E}\cdot\frac{\partial\boldsymbol{E}}{\partial t}\Big)\,\mathrm{d}^3\boldsymbol{x} \\ &= \frac{1}{\mu_0}\oint(\boldsymbol{E}\times\boldsymbol{B})\cdot\mathrm{d}^2\boldsymbol{S} + \frac{1}{2\mu_0}\int\frac{\partial}{\partial t}\big(B^2 + E^2/c^2\big)\,\mathrm{d}^3\boldsymbol{x}.\end{aligned} \tag{1.25}$$

If energy is to be conserved, the energy we deploy moving the charges has to go somewhere. According to (1.25) energy will be conserved if we interpret the **Poynting vector**

$$\boldsymbol{N} \equiv \frac{1}{\mu_0}\boldsymbol{E}\times\boldsymbol{B} \tag{1.26}$$

as the flux of e.m. energy, and

$$\frac{1}{2\mu_0}\left(B^2+E^2/c^2\right) \tag{1.27}$$

as the density of e.m. energy.

How do the Poynting vector and the e.m. energy-density fit into the scheme of n-tuples? From $\mathbf{F}$ we can construct the following important tensor:

$$\begin{aligned} T^{\mu\nu} &= \frac{1}{\mu_0}[-\tfrac{1}{4}(F_{\delta\gamma}F^{\delta\gamma})\eta^{\mu\nu} - F^{\mu}{}_{\gamma}F^{\gamma\nu}];\\ \mathbf{T} &= \frac{1}{\mu_0}\left[\tfrac{1}{4}\,\mathrm{Tr}(\mathbf{F}\cdot\mathbf{F})\boldsymbol{\eta} - \mathbf{F}\cdot\mathbf{F}\right], \end{aligned} \tag{1.28}$$

where $\mathbf{F}$ is, as usual, the Maxwell field tensor (1.16). It's easy to see that $\mathrm{Tr}\,\mathbf{T}=0$. A little slog shows that in terms of $\boldsymbol{E}$ and $\boldsymbol{B}$ the tensor $\mathbf{T}$ is

$$T^{\mu\nu} = \begin{pmatrix} \frac{1}{2\mu_0}(B^2+E^2/c^2) & N_x/c & N_y/c & N_z/c \\ N_x/c & & & \\ N_y/c & & P_{ij} & \\ N_z/c & & & \end{pmatrix}, \tag{1.29}$$

where

$$P_{ij} \equiv \frac{1}{\mu_0}\left[\tfrac{1}{2}\delta_{ij}\left(B^2+\frac{E^2}{c^2}\right) - \left(B_iB_j + \frac{E_iE_j}{c^2}\right)\right] \quad (i,j=1,2,3). \tag{1.30}$$

Thus the energy density in the e.m. field is the 00 component of $\mathbf{T}$ and the Poynting vector occupies the mixed space-time components of $\mathbf{T}$. It turns out that the 3×3 matrix P_{ij} describes the flux of the three kinds of momentum: P_{ix} = flux of x-momentum etc.

Exercise (7):

Show that a uniform magnetic field parallel to the z-axis is associated with tension (negative pressure) along the axis, and pressure in the perpendicular directions.

As an example of $\mathbf{T}$ consider a plane e.m. wave running along $\hat{\mathbf{i}}$ polarized parallel to $\hat{\mathbf{j}}$. Then

$$\begin{aligned} \boldsymbol{E} &= (0,E,0)\cos(\omega t-kx)\\ \boldsymbol{B} &= (0,0,B)\cos(\omega t-kx). \end{aligned}$$

E and B are related by $-\partial\boldsymbol{B}/\partial t = \nabla\times\boldsymbol{E} \quad\Rightarrow\quad B=kE/\omega=E/c$. Hence

$$\boldsymbol{N} = (E^2/\mu_0 c,0,0)\cos^2(\omega t-kx).$$

The first term in our expression (1.30) is non-zero only on the diagonal. The second term is non-zero only in the yy and zz slots and there cancels the first term. So $\boldsymbol{P}$ is

$$P_{ij} = \begin{pmatrix} 1&0&0\\0&0&0\\0&0&0 \end{pmatrix}\frac{E^2}{\mu_0c^2}\cos^2(\omega t-kx),$$

and finally

$$T^{\mu\nu} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \frac{E^2}{\mu_0 c^2} \cos^2(\omega t - kx). \tag{1.31}$$

The stress tensor $\boldsymbol{P}$ has only an entry in the xx slot because our wave is engaged in the business of carrying x-type momentum in the x-direction; the wave would push back a mirror placed in a plane $x =$ constant. Clearly the Poynting vector is also directed along the x axis, which accounts for the off-diagonal units in $\mathbf{T}$. In proper relativistic units the wave employs unit energy density ("capital employed") to carry unit fluxes of energy and momentum ("turnover"). Notice that the wave's phase is the scalar $-\mathbf{k}\cdot\mathbf{x}$.

1.5 Derivatives of tensors

Derivatives with respect to any system of coordinates can be expressed in terms of derivatives w.r.t. any other system by use of the chain rule:

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu}. \tag{1.32}$$

If the primed and unprimed systems are linked by a Lorentz transformation,

$$x'^\nu = \Lambda^\nu{}_\mu x^\mu, \tag{1.33}$$

we have on multiplying by $\Lambda_\nu{}^\kappa$ and summing over ν,

$$\Lambda_\nu{}^\kappa x'^\nu = \Lambda_\nu{}^\kappa \Lambda^\nu{}_\mu x^\mu = x^\kappa,$$

where the last step follows by (1.12). Differentiating we get

$$\frac{\partial x^\kappa}{\partial x'^\nu} = \Lambda_\nu{}^\kappa. \tag{1.34}$$

Thus

$$\frac{\partial}{\partial x'^\mu} = \Lambda_\mu{}^\nu \frac{\partial}{\partial x^\nu}, \tag{1.35}$$

and we see that

$$\partial_\mu \equiv \partial/\partial x^\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \tag{1.36}$$

transforms like a down vector.

Notes:

(i)

$$\frac{\partial}{\partial x^\mu} \text{ operates on scalars to produce vectors:} \quad G_\mu \equiv \frac{\partial \phi}{\partial x^\mu} \equiv \partial_\mu \phi \equiv \phi,_\mu$$

$$\frac{\partial}{\partial x^\mu} \text{ operates on vectors to produce } 2^{\text{nd}} \text{ rank tensors:}$$

$$G_{\mu\nu} \equiv \frac{\partial A_\nu}{\partial x^\mu} \equiv \partial_\mu A_\nu \equiv A_\nu,_\mu$$

$$\frac{\partial}{\partial x^\mu} \text{ operates on tensors to produce higher-rank tensors:}$$

$$G_{\mu\lambda\nu} \equiv \frac{\partial B_{\lambda\nu}}{\partial x^\mu} \equiv \partial_\mu B_{\lambda\nu} \equiv B_{\lambda\nu},_\mu$$

The operand's indices can be either up or down: $G_\mu{}^\nu = \partial_\mu A^\nu$.

(ii) If we contract the tensor produced by operating on a vector, we get a scalar, the 4-divergence $\psi = \partial_\mu A^\mu$.

(iii) We can reduce the number of indices on a higher-rank tensor by contraction: $A^\nu = \partial_\mu G^{\mu\nu}$.

(iv) The 4-analogue of taking the curl of a vector is to antisymmetrize the tensor formed by operating on a vector: $F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu)$. If $A_\nu = \partial_\nu \phi$, then $F_{\mu\nu} = 0$ because partial derivatives commute.

(v) A natural generalization of the divergence theorem reads

$$\int_V \mathrm{d}^4\mathbf{x}\, \frac{\partial T_{\alpha\ldots}}{\partial x^\mu} = \oint_S (\mathrm{d}^3\mathbf{x})_\mu T_{\alpha\ldots}, \tag{1.37}$$

where S is the boundary of the 4-d region V. Notice that $\mathbf{T}$ may have as many indices as it pleases and that one of them may be contracted with μ if you wish.

Example:

In e.m. the usual vector potential $\boldsymbol{A}$ and the electrostatic potential ϕ form the four components of an up vector

$$A^\mu = (\phi/c, A_x, A_y, A_z) \quad [\Rightarrow \quad A_\mu = (-\phi/c, A_x, A_y, A_z)]. \tag{1.38}$$

Our old friend the Maxwell field tensor $\mathbf{F}$ is then

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.39}$$

Thus $F_{12} = \dfrac{\partial A_y}{\partial x} - \dfrac{\partial A_x}{\partial y} = B_z$ and $F_{01} = \dfrac{\dot{A}_x}{c} + \dfrac{1}{c}\dfrac{\partial \phi}{\partial x} = -E_x/c$.

Derivatives with respect to proper time The history of a particle defines a curve in space-time. Let λ be a parameter which labels points on the curve in a continuous way. Then the coordinates x^μ of points on the curve are continuous functions $x^\mu(\lambda)$. For $\delta\lambda \ll 1$ the small vector

$$\delta\mathbf{x} \equiv \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\lambda}\,\delta\lambda$$

almost joins two points on the curve. Hence it is time-like and $|\delta\mathbf{x}| < 0$. For any two points A and B on the curve, we define

$$\tau \equiv \frac{1}{c}\int_{\mathrm{A}}^{\mathrm{B}} \sqrt{-\left|\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\lambda}\right|}\,\mathrm{d}\lambda \tag{1.40}$$

to be the **proper time** difference between A and B along the curve. If the curve is a straight line, we may transform to the coordinate system in which $x^\mu = (ct, 0, 0, 0)$ at all points on the curve, and then

$$\tau = \frac{1}{c}\int_{\mathrm{A}}^{\mathrm{B}} \sqrt{-\frac{\mathrm{d}ct}{\mathrm{d}\lambda}\frac{\mathrm{d}(-ct)}{\mathrm{d}\lambda}}\,\mathrm{d}\lambda = [t_{\mathrm{B}} - t_{\mathrm{A}}]. \tag{1.41}$$

Hence the name. We regard the coordinates x^μ of events along the trajectory as functions $x^\mu(\tau)$ of the proper time. Differentiating w.r.t. τ and multiplying through by the rest mass m_0 we obtain a 4-vector, the momentum

$$\mathbf{p} \equiv m_0\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\tau}. \tag{1.42}$$

From the zeroth component of the up version of this equation we have $\mathrm{d}t = \gamma\mathrm{d}\tau$; the hearts of passengers on a fast train (they mark off units of τ) appear to beat slowly to a medic on the station platform (whose watch keeps t).

1.6 Laws of e.m. and mechanics in tensor form

The relativistic generalization of Newton's second law is

$$m_0\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}\tau^2} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(m_0\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\tau}\right) = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}\tau} = \mathbf{f}, \tag{1.43}$$

where $\mathbf{f}$ is the **4-force**. The last three components of f^μ are just the Newtonian force components f_i. With $\mu = 0$ equation (1.43) states that the zeroth component of f^μ is to $1/c$ times the rate of change of the particle's energy cp^0; hence physically f^0 is $1/c$ times the rate of working of the force w. In summary

$$f^\mu = (w/c, f_x, f_y, f_z). \tag{1.44}$$

The divergence of (1.16) consists of these four equations:

$$F^{\mu\nu}{}_{,\nu} = \begin{pmatrix} \dfrac{1}{c}\dfrac{\partial E_x}{\partial x} + \dfrac{1}{c}\dfrac{\partial E_y}{\partial y} + \dfrac{1}{c}\dfrac{\partial E_z}{\partial z} \\ \partial B_z/\partial y - \partial B_y/\partial z - \frac{1}{c^2}\partial E_x/\partial t \\ -\partial B_z/\partial x + \partial B_x/\partial z - \frac{1}{c^2}\partial E_y/\partial t \\ \partial B_y/\partial x - \partial B_x/\partial y - \frac{1}{c^2}\partial E_z/\partial t \end{pmatrix} = \begin{pmatrix} \frac{1}{c}\nabla\cdot\boldsymbol{E} \\ \nabla\times\boldsymbol{B} - \dfrac{1}{c^2}\dfrac{\partial\boldsymbol{E}}{\partial t} \end{pmatrix}. \tag{1.45}$$

The zeroth component is by Poisson's equation equal to $\rho/(c\epsilon_0) = c\mu_0\rho$, where ρ is the charge density. By Ampere's law, the last three of these equations are equal to $\mu_0\boldsymbol{j}$, where $\boldsymbol{j}$ is the current density. Hence if we form a 4-vector

$$j^\mu = (c\rho, j_x, j_y, j_z), \tag{1.46}$$

we may write four of Maxwell's equations as

$$F^{\mu\nu}{}_{,\nu} = \mu_0 j^\mu. \tag{1.47}$$

It is straightforward to verify that Maxwell's other four equations can be written

$$F_{\mu\nu},_\lambda + F_{\lambda\mu},_\nu + F_{\nu\lambda},_\mu = 0 \quad (\mu \neq \nu \neq \lambda). \tag{1.48}$$

Exercises (8):

(i) Show that when λ, μ and ν equal 1, 2 and 3 respectively, (1.48) becomes $\nabla \cdot \boldsymbol{B} = 0$.

(ii) Show that with equation (1.22) equation (1.48) may also be written $\overline{F}^{\mu\nu}{}_{,\nu} = 0$.

Charge conservation is expressed as

$$\mu_0 \partial \cdot \mathbf{j} = \mu_0 j^\mu{}_{,\mu} = F^{\mu\nu}{}_{,\nu\mu} = 0, \tag{1.49}$$

where the last step follows by the antisymmetry of $\mathbf{F}$.

The natural definition of the 4-current associated with a particle of charge q is

$$\mathbf{J} = q\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\tau}. \tag{1.50}$$

Since the force exerted on a charged particle by an e.m. field has to be linear in q, the fields represented by $\mathbf{F}$, and the particle's velocity vector, a suitable 4-vector to try as the force is

$$\mathbf{f} = \mathbf{F} \cdot \mathbf{J}. \tag{1.51}$$

Tentatively inserting this into (1.43) and multiplying through by $\mathrm{d}\tau/\mathrm{d}t = 1/\gamma$ to obtain the acceleration as measured in the laboratory frame, we get

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = q\mathbf{F} \cdot \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}. \tag{1.52}$$

It is straightforward to check that the last three components of the up form of this vector are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(m_0\gamma\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}\right) = q(\boldsymbol{v} \times \boldsymbol{B} + \boldsymbol{E}),$$

while the zeroth component is

$$\frac{\mathrm{d}(m_0 c\gamma)}{\mathrm{d}t} = \frac{q}{c}\boldsymbol{E} \cdot \boldsymbol{v},$$

or, in words, "the rate of change of the particle's energy mc^2 is equal to the rate of working of the Lorentz force."

Gauge invariance At a classical (i.e. non-quantum level) only $\boldsymbol{E}$ and $\boldsymbol{B}$ are physically meaningful—$\boldsymbol{A}$ is just an abstraction from which $\boldsymbol{E}$ and $\boldsymbol{B}$ can be calculated via $F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu)$. So nothing physical changes if we replace $\mathbf{A}$ by

$$\mathbf{A}' \equiv \mathbf{A} + \partial\Lambda, \tag{1.53}$$

where $\Lambda(\mathbf{x})$ is any scalar-valued function of space-time coordinates. The change (1.53) in $\mathbf{A}$ is called a **gauge transformation**.

Gauge transformations can be used to ensure that $\mathbf{A}$ satisfies an additional equation. In particular, given $\mathbf{A}$ we can choose Λ s.t. $\mathbf{A}'$ satisfies one of these gauge conditions:

(i) **Lorentz gauge**:[2]

$$\partial \cdot \mathbf{A}' = 0 \quad \Rightarrow \quad \Box\Lambda = \partial \cdot \mathbf{A} \tag{1.54}$$

The Lorentz condition (1.54) does not uniquely specify $\mathbf{A}'$ since many non-trivial functions satisfy $\Box\phi = 0$ and so given one Λ satisfying the 2^{nd} of eqs (1.54), we can construct many others $\Lambda' = \Lambda + \phi$.

(ii) **Coulomb** or **radiation** or transverse gauge

$$\nabla \cdot \boldsymbol{A}' = 0 \quad \Rightarrow \quad \nabla^2\Lambda = \nabla \cdot \boldsymbol{A} \tag{1.55}$$

In this gauge the 0^{th} eqn of the set $\partial^\nu F_{\mu\nu} = \mu_0 j_\mu$ reads

$$\begin{aligned}
\frac{\rho}{c\epsilon_0} &= -\mu_0 j_0 = -\partial^\nu(\partial_0 A_\nu - \partial_\nu A_0) \\
&= -\partial_0\partial^\nu A_\nu + \partial^\nu\partial_\nu A_0 \\
&= -\partial_0\partial^0 A_0 + \partial^\nu\partial_\nu A_0 \\
&= \partial^i\partial_i A_0 \\
&= -\nabla^2\phi/c
\end{aligned} \tag{1.56}$$

i.e., in this gauge the electrostatic potential satisfies Poisson's eqn, which explains the gauge's name.

1.7 Summary

The special theory of relativity requires that any physical quantity must fit into an n-tuple of numbers, where $n = 1, 4, 6, 10, \ldots$. Physical laws must be expressed as equations connecting the n-tuples associated with different physical quantities. These equations must be constructed in accordance with the rules of tensor calculus, which permit only:

(i) the multiplication of n-tuples to form either higher-rank n-tuples (as in $H_{\mu\nu} = x_\mu p_\nu - x_\nu p_\mu$) or lower-rank n-tuples (as in $f_\mu = F_\mu{}^\nu J_\nu$), or

[2] We denote the d'Alembertian opertor by $\Box \equiv \partial_\mu\partial^\mu$ by analogy with the notation $\triangle \equiv \nabla^2 = \partial_i\partial^i$ for the Laplacian operator.

(ii) the addition of n-tuples *of the same rank.*

In particular, both sides of every acceptable equation always form valid n-tuples of the same kind.

Rest-mass, electric charge and total spin are scalars (1-tuples). The most important 4-vectors (4-tuples) include any particle's energy-momentum $\mathbf{p}$, e.m. current $\mathbf{J}$ or acceleration $\mathrm{d}\mathbf{p}/\mathrm{d}\tau$, and the potential $\mathbf{A}$ of the e.m. field. Important 6-tuples include any particle's angular momentum $\mathbf{H}$ and the Maxwell field tensor $\mathbf{F}$. An important 10-tuple is the density $\mathbf{T}$ of the energy-momentum due to the e.m. field.

In 4-vector notation the key equation of mechanics and e.m. are

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\tau} \quad ; \quad \mathbf{p} = m_0\mathbf{v} \quad ; \quad \mathbf{J} = q\mathbf{v}$$

$$\mathbf{f} = \mathbf{F}\cdot\mathbf{J} \quad ; \quad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}\tau} = \mathbf{f}$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad ; \quad F^{\mu\nu}{}_{,\nu} = \mu_0 j_\mu \quad ; \quad \overline{F}^{\mu\nu}{}_{,\nu} = 0,$$

where $F^{\mu\nu} \equiv \eta^{\mu\gamma}\eta^{\nu\delta}F_{\gamma\delta}$ and $\overline{F}^{\mu\nu} \equiv \frac{1}{2}\epsilon^{\mu\nu\gamma\delta}F_{\gamma\delta}$. The energy-monentum tensor of the e.m. field is

$$T^{\mu\nu} = \frac{1}{\mu_0}\left[\tfrac{1}{4}\,\mathrm{Tr}(\mathbf{F}\cdot\mathbf{F})\eta^{\mu\nu} - F^{\mu}{}_{\gamma}F^{\gamma\nu}\right].$$