Physics 209
Fall 2002
Notes 4
More Tensor Analysis

Let $\{x^i, i = 1, \ldots, n\}$ be a set of *coordinates* on an n-dimensional space, for which we write x^i for short. Initially we make no assumptions about the nature of the space (it might be curved, have many dimensions, etc.), or the coordinates (they might be curvilinear, etc.), but we do assume that the coordinates provide a one-to-one labelling of the points of the space. Similarly, let $\{x'^i, i = 1, \ldots, n\}$, or x'^i for short, be a new coordinate system. Since it also labels points in a one-to-one manner, the coordinates x^i and x'^i are invertible functions of one another, $x^i = x^i(x')$, $x'^i = x'^i(x)$. This means that the *Jacobian matrices*,

$$\frac{\partial x'^i}{\partial x^j} \quad \text{and} \quad \frac{\partial x^i}{\partial x'^j} \tag{4.1}$$

exist, are nonsingular, and are inverses of each other,

$$\frac{\partial x'^i}{\partial x^j}\frac{\partial x^j}{\partial x'^k} = \delta^i_k. \tag{4.2}$$

Here and throughout we use the summation convention. It is customary to use a superscript index on the coordinate vectors x^i, x'^i, etc.

The different coordinates in a coordinate system need not have the same dimensions, for example (r, θ, ϕ) (one distance and two angles). The coordinates might be *curvilinear* even if the space is flat (for example, (r, θ, ϕ)), but if the space is curved (for example, the surface of a sphere), then the coordinates must be curvilinear. If the space is flat (a vector space), then the coordinates may be *rectilinear* (coordinate axes are straight lines), but rectilinear coordinates need not be orthogonal. In fact, orthogonal coordinates (in the usual sense) are only meaningful on a *Euclidean* space which has a Euclidean metric, so that lengths and angles are defined.

Now some definitions. A *scalar* is a quantity that has the same value in all coordinate systems. If $S(x)$ and $S'(x')$ are the values of a scalar field in two coordinate systems, then

$$S(x) = S'(x'), \tag{4.3}$$

where $x^i = x^i(x')$ is understood (x^i and x'^i refer to the same point). This statement (and others like it below) imply a physical or mathematical rule whereby S can be determined or measured or computed in different coordinate systems.

Next, a *contravariant vector* $\{A^i, i = 1, \ldots, n\}$, or A^i for short, is a set of n quantities that transform according to the rule,

$$A'^i = \frac{\partial x'^i}{\partial x^j} A^j, \tag{4.4}$$

where A^i and A'^i are the quantities measured or computed in the two coordinate systems x^i and x'^i. An upper (superscript) position is conventional for contravariant vectors. The most important example of a contravariant vector is a set of infinitesimal coordinate displacements dx^i connecting two nearby points. These transform by the chain rule, which is the same as the contravariant transformation law,

$$dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j. \tag{4.5}$$

If we divide this by dt we obtain a velocity vector, which also transforms as a contravariant vector,

$$\frac{dx'^i}{dt} = \frac{\partial x'^i}{\partial x^j} \frac{dx^j}{dt}, \quad \text{or} \quad v'^i = \frac{\partial x'^i}{\partial x^j} v^j. \tag{4.6}$$

Next, a *covariant vector* $\{B_i, i = 1, \ldots, n\}$, or B_i for short, is a set of n quantities that transform according to the rule,

$$B'_i = \frac{\partial x^j}{\partial x'^i} B_j. \tag{4.7}$$

Notice that it is the inverse Jacobian that is used to transform a covariant vector, in comparision to a contravariant vector (actually, the inverse transpose, see below). A lower (subscript) index is standard notation for a covariant vector. The most important example of a covariant vector is the gradient of a scalar, that is, if S is a scalar, then

$$B_i = \frac{\partial S}{\partial x^i} \tag{4.8}$$

transforms as a covariant vector. This follows from the chain rule,

$$\frac{\partial S}{\partial x'^i} = \frac{\partial x^j}{\partial x'^i} \frac{\partial S}{\partial x^j}. \tag{4.9}$$

The *scalar product* of a contravariant times a covariant vector is defined by

$$A^i B_i = A'^i B'_i = A \cdot B. \tag{4.10}$$

It is a simple exercise to show that it transforms as a scalar.

A *tensor* is an object with multiple indices, each of which takes on the values $1, \ldots, n$, some contravariant and some covariant, which transforms with the appropriate transformation law in each index. The *rank* of the tensor is the number of indices. For example, the third rank tensor $T^i{}_j{}^k$ transforms according to

$$T'^\ell{}_m{}^n = \frac{\partial x'^\ell}{\partial x^i} \frac{\partial x^j}{\partial x'^m} \frac{\partial x'^n}{\partial x^k} T^i{}_j{}^k, \tag{4.11}$$

where the first index is contravariant, the second covariant, and the third contravariant. Sometimes we use dots as place holders, for example, $T^{i\cdot k}_{\cdot j \cdot}$, to keep track of the order of the indices when some are contravariant and some covariant. A tensor with some contravariant and some covariant indices is said to be a *mixed* tensor.

The *Kronecker delta* δ^i_j is a mixed tensor. That is, suppose we define n^2 quantities δ^i_j in all coordinate systems to be 1 if $i = j$, and 0 otherwise. Then these quantities are related by the mixed transformation law,

$$\delta'^i_j = \frac{\partial x'^i}{\partial x^k}\,\frac{\partial x^\ell}{\partial x'^j}\,\delta^k_\ell = \frac{\partial x'^i}{\partial x^k}\,\frac{\partial x^k}{\partial x'^j} = \delta^i_j, \tag{4.12}$$

which follows from Eq. (4.2). But this only works for the mixed Kronecker delta. We can certainly define quantities δ_{ij} or δ^{ij} in all coordinate systems by the usual rules for the Kronecker delta, but the resulting object does not transform as a tensor (that is, a purely covariant or contravariant tensor of rank 2).

The *Levi-Civita symbol* is not generally a tensor. Suppose, for example, we attempt to interpret ϵ^{ijk} as a purely contravariant tensor on a 3-dimensional space, and we assign the usual values to this object in all coordinate systems. Then we find it does not transform as a tensor, since

$$\frac{\partial x'^i}{\partial x^\ell}\,\frac{\partial x'^j}{\partial x^m}\,\frac{\partial x'^k}{\partial x^n}\,\epsilon^{\ell m n} = \det\left(\frac{\partial x'^i}{\partial x^\ell}\right)\epsilon^{ijk}. \tag{4.13}$$

Thus, the Levi-Civita symbol behaves as a tensor only if the determinant of the Jacobian is 1. If, however, we restrict consideration to coordinate transformations that satisfy this condition, then the Levi-Civita symbol can be regarded as a tensor.

The product of any two tensors is a tensor, for example, $A_i F^j{}_k$ is a third-rank tensor.

In any tensor or product of tensors, a sum on one contravariant and one covariant index always produces a new tensor whose rank is two less than that of the original tensor. For example, we might have

$$A^i B_i, \qquad M^i{}_i, \qquad T^i{}_i{}^j, \qquad M^i{}_j A^j, \quad \text{etc.} \tag{4.14}$$

This is called a *contraction* of indices. Obviously the scalar product of a contravariant vector and a covariant vector is a special case. But a sum on two contravariant or two covariant indices, for example,

$$T^i{}_j{}^i, \qquad F_{ii}, \qquad A^i A^i, \quad \text{etc.} \tag{4.15}$$

is not a tensor. Things like this do occur, but only when we are stepping outside the bounds of usual tensor analysis. The point is that an object like (4.15) cannot have a meaning that is independent of coordinate system.

Let us write

$$J^i{}_j = \frac{\partial x'^i}{\partial x^j} \tag{4.16}$$

for the Jacobian matrix connecting two coordinate systems. It is not a tensor since it is not defined in all coordinate systems, rather it is an object that connects two specific coordinate systems. Nevertheless it is convenient to use upper and lower indices on the Jacobian to conform with the usual rules for index contraction. The Jacobian $J^i{}_j$ is the one that occurs in the transformation law for contravariant indices. The Jacobian that occurs in the covariant law is the inverse transpose of this one,

$$(J^{-1t})_i{}^j = \frac{\partial x^j}{\partial x'^i}. \tag{4.17}$$

Many spaces possess a *metric tensor*, which specifies the distance between two nearby points, with coordinates x^i and $x^i + dx^i$ in some coordinate system. The square of the distance is given by

$$ds^2 = g_{ij}\, dx^i dx^j. \tag{4.18}$$

It follows from this definition that g_{ij} transforms as a second rank, covariant tensor. The metric tensor is symmetric, $g_{ij} = g_{ji}$. If $ds^2 \geq 0$ for all dx^i, with $ds^2 = 0$ if and only if $dx^i = 0$, then the metric is *positive definite.* But relativity uses an indefinite metric (the Minkowski metric).

The metric tensor g_{ij} is normally invertible. Its inverse is defined to be the metric tensor g^{ij} (with contravariant, rather than covariant, indices):

$$g_{ij}\, g^{jk} = \delta_i^k. \tag{4.19}$$

This equation specifies g^{ij} in all coordinate systems. Given that g_{ij} transforms as a covariant tensor, it is easy to show that g^{ij} transforms as a contravariant tensor.

If A^i is a contravariant vector, then by contracting with g_{ij} we can create a covariant vector $g_{ij}A^j$. This is customarily denoted A_i (with the same symbol A, just a lower index):

$$A_i = g_{ij} A^j. \tag{4.20}$$

This is called *lowering an index.* In physical or geometrical applications, A^i and A_i are often thought of as two different representations of the same object (for example, a velocity v^i or v_i), even though the numerical values of the components are different. Similarly, we can lower any contravariant index in any tensor, for example,

$$T_{ij}{}^k = g_{i\ell}\, T^\ell{}_j{}^k. \tag{4.21}$$

Similarly, if B_i is a covariant vector, then by contracting with g^{ij} we can create a contravariant vector $g^{ij}B_j$. This is customarily denoted B^i (with the same symbol B, just an upper index):

$$B^i = g^{ij} B_j. \tag{4.22}$$

This is called *raising an index*. Similarly, we can raise any covariant index in any tensor, for example,

$$T^{ijk} = g^{j\ell} T^i{}_\ell{}^k. \tag{4.23}$$

By raising or lowering indices, any tensor can be written with any index in either an upper or lower position.

Once the index on a contravariant vector A^i has been lowered, we can form the scalar product with another contravariant vector:

$$A^i g_{ij} B^j = A_j B^j = A \cdot B. \tag{4.24}$$

Similarly, we can form the scalar product of two covariant vectors:

$$A_i g^{ij} B_j = A^j B_j = A \cdot B. \tag{4.25}$$

The scalar product of two contravariant vectors or two covariant vectors is not defined unless we have a metric.

Henceforth we restrict consideration to flat spaces (vector spaces) upon which rectilinear coordinates are used. At first we do not assume these are orthonormal, for example, we might have skew coordinates in plane.

Such coordinate systems on such spaces are related by linear transformations. This means that the Jacobian matrix is a constant matrix,

$$x'^i = J^i{}_j \, x^j, \qquad J^i{}_j = \frac{\partial x'^i}{\partial x^j} = \text{const.} \tag{4.26}$$

In rectilinear coordinate systems, the coordinates themselves (not just their differentials) transform as a contravariant vector.

In rectilinear coordinates, the coordinate derivative (or gradient) of a tensor field is another tensor field. For example, consider a contravariant vector field A^i, and define

$$B^i{}_j = \frac{\partial A^i}{\partial x^j}. \tag{4.27}$$

This defines quantities $B^i{}_j$ in all coordinate systems. Then it is easy to show that $B^i{}_j$ transforms as a tensor (a mixed tensor, in this case). The proof requires that the Jacobian matrix be a constant, so it only holds for rectilinear coordinate systems.

The gradient or coordinate derivative introduces one extra covariant index into a tensor, thus raising the rank by one. We may say that the gradient operator $\partial/\partial x^i$ transforms as a covariant vector. Notice that the superscript below in a partial derivative such as in Eq. (4.27) acts like a subscript above. Sometimes we write

$$\partial_i = \frac{\partial}{\partial x^i}, \tag{4.28}$$

for example, Eq. (4.27) can be written

$$B^i{}_j = \partial_j A^i. \tag{4.29}$$

Another way to write this is to use the "comma notation" for derivatives, for example,

$$F^{ij}{}_{,j} = \partial_j F^{ij} = \frac{\partial F^{ij}}{\partial x^j}. \tag{4.30}$$

This is the gradient operator. The divergence operator can be applied to any tensor with a contravariant index, for example,

$$\frac{\partial A^i}{\partial x^i} = \partial_i A^i = A^i{}_{,i}. \tag{4.31}$$

The divergence is just a contraction involving the gradient operator. It can be applied to tensors as well as vectors.

The curl in the usual sense requires the Levi-Civita symbol on three-dimensional, Euclidean space. In more general contexts, the curl is replaced by an antisymmetrized gradient, resulting in antisymmetric tensors.

The spaces we are most interested in are three-dimensional, physical space and four-dimensional space-time. In the following we consider only rectilinear coordinates on such spaces. These spaces possess a metric tensor g_{ij} whose components are constants. This means that taking derivatives (the gradient operator) commutes with raising and lowering indices. Thus, we can define the contravariant gradient operator,

$$\partial^i = g^{ij}\,\partial_j, \tag{4.32}$$

which takes any tensor field and produces another one with one more contravariant index.

On three-dimensional, physical space the metric is Euclidean, that is, positive definite. This means that there exists a rectilinear coordinate system in which the components of the metric are

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.33}$$

that is, $g_{ij} = \delta_{ij}$. Such coordinates are *orthonormal* coordinates. This form of the metric is preserved under orthogonal transformations, that is, transformations of the form

$$x'^i = R^i{}_j \, x^j \tag{4.34}$$

in which R (the Jacobian matrix) is orthogonal, $R^t R = I$. All systems of orthonormal coordinates are connected by orthogonal transformations.

If we restrict consideration to orthonormal coordinates and orthogonal transformations, then the laws of tensor analysis simplify. This is because for an orthogonal matrix, $R^{-1t} = R$, so the transformation law for contravariant and covariant vectors is identical. Moreover, since the metric tensor forms an identity matrix, raising or lowering an index does not change any of the values of the components. Thus, there is no point in keeping track of the difference between contravariant and covariant indices, and we might as well use subscripts uniformly everywhere. Thus we can write $x'_i = R_{ij} \, x_j$, expressions like $A_i A_i = A \cdot A$ are meaningful, etc.

The orthogonality condition $R^t R = I$ implies $\det R = \pm 1$. Orthogonal transformations with $\det R = +1$ are called *proper*, and those with $\det R = -1$ are called *improper.* A proper orthogonal transformation perserves the sense of the axes (right-handed or left-handed), while an improper one reverses the sense.

On three-dimensional, physical space there exists one orthogonal coordinate system (hence a whole family) with the physical interpretation of being a right-handed coordinate system. The members of this family are connected by proper orthogonal transformations.

The Levi-Civita symbol behaves as a tensor if we restrict consideration to proper, orthogonal changes of coordinates, as shown by Eq. (4.11). It changes sign under improper orthogonal coordinate transformations.

On three-dimensional space we often distinguish *polar* or true vectors from *axial* or *pseudo*-vectors. A polar vector transforms as a vector under all orthogonal transformations, proper and improper, while an axial vector changes sign under improper orthogonal transformations. Axial vectors may be created by taking the cross product of two polar vectors or curl of a polar vector; the transformation properties of the Levi-Civita symbol then cause a change in sign under an improper orthogonal transformation.

Now we turn to four-dimensional space-time. We switch to Greek indices, which are popular for relativity. The indices run from 0 to 3. Space-time has an indefinite metric with one positive and three negative eigenvalues. This means that ds^2 in Eq. (4.18) can be positive, negative, or zero. It also means that there exists a rectilinear coordinate system in which the components of the metric tensor are given by

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{4.35}$$

Alternatively, we can define the metric with the opposite sign; this is a matter of convention. Note that if $g_{\mu\nu}$ has the form (4.35), then $g_{\mu\nu} = g^{\mu\nu}$ (the components of the covariant and contravariant metric tensors are identical). Not all coordinate systems in which the metric has the form (4.35) are inertial frames in the usual sense, for we must demand that the 0-axis be pointing in the positive time direction and the 1,2,3 axes be oriented in a right-handed sense. More on this in a moment.

The form (4.35) of the metric is preserved under a restricted class of linear transformations called *Lorentz transformations*. If we write

$$x'^{\mu} = L^{\mu}{}_{\nu}\, x^{\nu} \tag{4.36}$$

for a Lorentz transformation, then the transformation matrix L (the Jacobian) must satisfy

$$L^{\alpha}{}_{\mu}\, g_{\alpha\beta}\, L^{\beta}{}_{\nu} = g_{\mu\nu}, \tag{4.37}$$

which is just the transformation law for the components of a covariant tensor, except that we are setting $g'_{\mu\nu} = g_{\mu\nu}$. Alternatively, if we write L for the matrix with components $L^{\mu}{}_{\nu}$ (μ is the row index and ν is the column), and simply g for the matrix with components $g_{\mu\nu}$, then the condition on a Lorentz transformation is

$$L^t g L = g. \tag{4.38}$$

This may be compared to the condition $R^t I R = I$ for an orthogonal transformation; a Lorentz transformation preserves the Minkowski metric g, while an orthogonal transformation preserves the Euclidean metric I.

We previously derived the condition (4.37) or (4.38) as a generalization of the condition that the speed of light is constant in all inertial frames, and we checked that the Lorentz transformations we derived from the postulates of relativity do satisfy it. Now we are turning the question around, and asking, are all transformations that satisfy Eq. (4.37) or (4.38) Lorentz transfomations in the sense of the boosts and rotations? The answer is not quite, since some Lorentz transformations involve a time-reversal or parity (space-reversal) operation. This is like the distinction between proper and improper rotations, except that now there is time-reversal also to worry about.

Taking the determinant of Eq. (4.38) we find $\det L = \pm 1$. Thus, Lorentz transformations are characterized by the sign of the determinant. They are also characterized by the sign of the component $L^0{}_0$, which never vanishes. This sign tells whether the Lorentz transformation reverses the direction of time. Usually we are interested in Lorentz transformations such that $\det L = +1$ and $L^0{}_0 > 0$, which are called the *proper* Lorentz transformations. These are interpreted as transformations between inertial frames that do not change the direction of time and that do maintain the sense (right- or left-handed) of the spatial axes.

As mentioned above, an inertial frame in the usual sense is a coordinate system in which the metric has the standard form (4.35) with the 0-axis pointing in the direction of increasing time and the 1,2,3 axes oriented in a right-handed manner. We now see that inertial frames are connected by proper Lorentz transformations. If we apply an improper Lorentz transformation to an inertial frame in the usual sense, we get a frame which has been subjected to either a time-reversal or space-reversal (parity) operator, or both. For most of this course we will deal only with inertial frames in the usual sense and with proper Lorentz transformations.

Under Lorentz transformations it is not true (as it was with orthogonal rotations in three-dimensional space) that the Jacobian is equal to its inverse transpose. Thus it is necessary to maintain the distinction between contravariant and covariant indices, and raising and lowering an index does change the numerical values of the components.

In the older literature on special relativity, it was popular to define $x_4 = ict$ instead of $x_0 = ct$. Then the metric tensor becomes Euclidean, $g_{\mu\nu} = \delta_{\mu\nu}$, at the expense of introducing an imaginary coordinate. The advantage of doing this is that one can avoid the formalism of contravariant and covariant indices. The trend nowadays is to face up to contravariant and covariant indices, in order to have real coordinates.

The transformation matrix for a covariant index under a Lorentz transformation can be worked out in terms of $L^\mu{}_\nu$ by using the metric to raise the index, then applying the contravariant law. That is, we have

$$B'_\mu = g_{\mu\nu} B'^\nu = g_{\mu\nu}\, L^\nu{}_\alpha B^\alpha = g_{\mu\nu}\, L^\nu{}_\alpha\, g^{\alpha\beta} B_\beta, \tag{4.39}$$

or

$$B'_\mu = L_\mu{}^\beta B_\beta, \tag{4.40}$$

where

$$L_\mu{}^\beta = g_{\mu\nu}\, L^\nu{}_\alpha\, g^{\alpha\beta}. \tag{4.41}$$

This notation just follows the usual rules for raising and lowering indices (even though $L^\mu{}_\nu$ is not a tensor). Alternatively, we can use Eq. (4.38) and matrix notation to write $L^{-1t} = gLg^{-1}$ for the matrix that transforms a covariant vector. Notice that if we agree to raise and lower indices on $L^\mu{}_\nu$ with $g_{\mu\nu}$ and $g^{\mu\nu}$, Eq. (4.37) can be written

$$L_{\beta\mu} L^{\beta\nu} = \delta^\nu_\mu. \tag{4.42}$$

The completely contravariant, four-dimensional Levi-Civita symbol $\epsilon^{\mu\nu\alpha\beta}$ is defined as $+1$ if $(\mu\nu\alpha\beta)$ is an even permutation of (0123), as -1 if it is an odd permutation, and 0 otherwise. It transforms as a tensor under proper Lorentz transformations. If we lower all four indices to get the completely covariant tensor, we find

$$\epsilon_{\mu\nu\alpha\beta} = g_{\mu\sigma}\, g_{\nu\tau}\, g_{\alpha\kappa}\, g_{\beta\lambda}\, \epsilon^{\sigma\tau\kappa\lambda} = \det(g)\epsilon^{\mu\nu\alpha\beta} = -\epsilon^{\mu\nu\alpha\beta}. \tag{4.43}$$

Thus the completely covariant Levi-Civita has the same numerical values as the completely contravariant symbol, except for a minus sign.