Physics 221A
Academic Year 2021–22
Appendix E
Introduction to Tensor Analysis†

1. Introduction

These notes contain an introduction to tensor analysis as it is commonly used in physics, but mostly limited to the needs of this course. The presentation is based on how various quantities transform under coordinate transformations, and is fairly standard. It is also somewhat old-fashioned, but it has the advantage of at least creating a mental image of how measurements are actually made in physics. It also provides a good background if one wishes to learn more about the more modern and more abstract approaches to tensors.

Quantities of physical interest are often expressed in terms of tensors. When physical laws are expressed in this manner, they often become *covariant*, that is, they have the same form in all coordinate systems. For example, we may say that an equation is or is not covariant. In tensor analysis the word "covariant" is also used in a different sense, to characterize a type of vector or an index on a tensor, as explained below. Covariant equations are preferred for the formulations of fundamental physical laws, because the form-invariance implies a separation between the objective reality (the invariant form of the equations) and the constructions introduced by the observer to make measurements (the coordinate system).

2. Spaces and Coordinates

In these notes we will be mainly concerned with two spaces: ordinary, three-dimensional, physical space, and four-dimensional space-time. We usually think of physical space as a Euclidean space extending to infinity in all directions, and model it mathematically as $\mathbb{R}^3$ with a Euclidean metric. In reality physical space is not exactly Euclidean, and whether it extends to infinity is a cosmological question (it seems that it probably does, but looking deeper into space involves looking backwards in time, and we can only see a finite part of space). As for space-time, in special relativity we usually think of it as extending to infinity in all four directions (including time), and model it as $\mathbb{R}^4$ with the Minkowski metric. As with three-dimensional, physical space, the cosmological reality is more complicated, but in both cases the models ($\mathbb{R}^3$ and $\mathbb{R}^4$ with their respective metrics) are adequate for problems of limited space and time extent, which applies to all problems considered in this course.

---

† Links to the other sets of notes can be found at:
`http://bohr.physics.berkeley.edu/classes/221/2122/221.html`.

Nevertheless, we will be slightly more general at the beginning of these notes, and think of more general kinds of spaces. These include, for example, the configuration spaces and phase spaces of mechanical systems (see Sec. B.3 and B.18) and many others. We do this because the curvilinear coordinates that are necessary on general spaces are sometimes used also on vector spaces such as $\mathbb{R}^3$ and $\mathbb{R}^4$, and because the general context is useful conceptually.

Consider a space of some kind in a physical problem (this corresponds to what in mathematics is called a *differentiable manifold*). A *coordinate system* on a space is a way of uniquely labeling the points of the space by means of coordinates, that is, n-tuples of numbers. The number of coordinates required is the *dimensionality* of the space. If the space has the structure of a vector space, then the simplest coordinate systems are *rectilinear*, in which the coordinate axes are stright lines. Such spaces include ordinary three-dimensional, physical space (in the usual model), and space-time (in special relativity). In both these examples, the space possesses a metric (Euclidean or Minkowski), so it is possible to choose the coordinate axes to be orthogonal (in the Euclidean or Minkowski sense). The resulting coordinate systems (orthonormal, Euclidean coordinates or coordinates associated with a Lorentz frame) are the usual choices.

But on a vector space it is also possible to use *skew* coordinates, that is, rectilinear coordinates that are not orthonormal. In some cases the space has no metric so orthonormality is not defined.

Even if the space is a vector space, we may choose to use coordinates that are not rectilinear, for example, the spherical coordinates (r, θ, ϕ) in three-dimensional, physical space. Such coordinates are called "curvilinear," because the coordinate lines are not straight. (The coordinate lines are obtained by varying one of the coordinates, while holding all the others fixed.) But if the space is not a vector space (for example, the surface of a sphere), then rectlinear coordinates do not exist, and all coordinates are curvilinear.

In the general case, we will denote a system of coordinates by $\{x^i, i = 1, \ldots, n\}$, or x^i for short, where n is the dimensionality of the space. It is customary to use a superscript to label the coordinates. It is assumed that the coordinates provide a one-to-one labeling of the points of the space, that is, a set of coordinate values uniquely identifies a point and vice versa. In general this is only possible over some region of the space, and some corresponding region of coordinate values. That is, some spaces (mathematically, *manifolds*), require more than one coordinate patch (mathematically, a *chart*) to cover them. The different coordinates need not have the same physical dimensions, for example, in spherical coordinates (r, θ, ϕ), the radius r has dimensions of distance, while the other two coordinates, θ and ϕ, are dimensionless.

We will be interested in how various quantities transform under a change of coordinates. In physical applications a coordinate system represents how we measure a physical system and express those measurements as numbers, but otherwise the coordinate system is a construct introduced by the observer, one that is not intrinsic to the system itself. For example, the measurement of the velocity vector produces three numbers, which are referred to a set of coordinate axes, but the system does not know about the coordinates we have chosen. Therefore, to understand the intrinsic

physical properties of the system, we study how the results of our measurments change when the coordinate system is changed.

Let $\{x'^i, i = 1, \ldots, n\}$, or x'^i for short, be a new coordinate system, in addition to x^i. Since it also labels points in a one-to-one manner, the coordinates x^i and x'^i are invertible functions of one another, $x^i = x^i(x')$, $x'^i = x'^i(x)$. [An equation like $x'^i = x'^i(x)$ means $x'^i = x'^i(x^1, \ldots, x^n)$ for $i = 1, \ldots, n$.] This means that the *Jacobian matrices*,

$$\frac{\partial x'^i}{\partial x^j} \quad \text{and} \quad \frac{\partial x^i}{\partial x'^j} \tag{1}$$

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

$$\frac{\partial x'^i}{\partial x^j} \frac{\partial x^j}{\partial x'^k} = \delta^i_k. \tag{2}$$

Here we use the summation convention (see Sec. 3). For example, the two coordinate systems on physical space, (x, y, z) and (r, θ, ϕ), are related by the equations,

$$\begin{aligned}
x &= r \sin\theta \cos\phi, \\
y &= r \sin\theta \sin\phi, \\
z &= r \cos\theta,
\end{aligned} \tag{3}$$

or their inverses,

$$\begin{aligned}
\phi &= \tan^{-1}(y/x), \\
\theta &= \tan^{-1}(\sqrt{x^2 + y^2}/z), \\
r &= \sqrt{x^2 + y^2 + z^2}.
\end{aligned} \tag{4}$$

Notice that spherical coordinates (r, θ, ϕ) do not provide a one-to-one labeling of points at the north pole $(r > 0, \theta = 0)$, since the angle ϕ can take on any value. At such singular points of a coordinate system one of the Jacobian matrices (1) is singular, and the other does not exist. In the following discussion we assume that we are away from such singular points. Notice that a singularity of a coordinate system is not the same as a singularity of the space; a sphere has no singularity at the north pole, where it looks the same as anywhere else.

3. The Summation Convention

In Eq. (2) we use the *summation convention*, often attributed to Einstein, in which there is an implied summation over repeated indices. It is particularly useful in tensor analysis. Equation (2) could be written more explicitly as

$$\sum_{j=1}^{n} \frac{\partial x'^i}{\partial x^j} \frac{\partial x^j}{\partial x'^k} = \delta^i_k, \tag{5}$$

since the index j is repeated on the left-hand side of Eq. (2) but the indices i and k are not. If there are indices that occur more than twice, or indices that occur only once but are summed over, or

indices that occur twice that are not summed over, then the summation convention cannot be used. Many books state that the summation convention will be used unless explictly stated otherwise. This is close to how the summation convention is used in these notes.

When the summation convention is used, repeated indices are *dummies*, and can be replaced by any other index that is not used elsewhere, without changing the result. For example, we could replace j by n in Eq. (2) without changing the result. We cannot, however, replace j by k in Eq. (2), because k is an index that is already being used.

4. Scalars

A *scalar* is 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

$$S(x) = S'(x'), \tag{6}$$

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. For example, one may think of the temperature in a room; the thermometer reading depends on the position of the thermometer, but not on the coordinates used to describe that position.

5. Contravariant Vectors

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,

$$A'^i = \frac{\partial x'^i}{\partial x^j} A^j, \tag{7}$$

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,

$$dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j. \tag{8}$$

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

$$\frac{dx'^i}{dt} = \frac{\partial x'^i}{\partial x^j} \frac{dx^j}{dt}, \quad \text{or} \quad v'^i = \frac{\partial x'^i}{\partial x^j} v^j. \tag{9}$$

In this equation we have assumed that dt is the same in both coordinate systems; this is the usual assumption in nonrelativistic mechanics, but in relativity dt depends on the coordinates, too, and the velocity does not transform as a vector.

As an example, the contravariant components of the velocity of a particle in spherical coordinates are $v^i = (\dot{r}, \dot{\theta}, \dot{\phi})$. These are not to be confused with the usual components of the velocity with respect to the orthonormal frame, $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}})$,

$$
\begin{aligned}
v_r &= \hat{\mathbf{r}} \cdot \mathbf{v} = \dot{r}, \\
v_\theta &= \hat{\boldsymbol{\theta}} \cdot \mathbf{v} = r\,\dot{\theta}, \\
v_\phi &= \hat{\boldsymbol{\phi}} \cdot \mathbf{v} = r\sin\theta\,\dot{\phi},
\end{aligned}
\tag{10}
$$

which are often used in problems in mechanics. Equation (10) is equivalent to Eq. (D.27c).

Although we put an upper index on the coordinates, for example, x^i, these do not form a contravariant vector, at least not in general, curvilinear coordinates. That is because coordinate transformations are nonlinear in general, as exemplified by Eq. (3) or (4). For the special case of rectilinear coordinates see Sec. 13.

6. Covariant Vectors

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,

$$
B_i' = \frac{\partial x^j}{\partial x'^i}\, B_j.
\tag{11}
$$

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

$$
B_i = \frac{\partial S}{\partial x^i}
\tag{12}
$$

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

$$
\frac{\partial S}{\partial x'^i} = \frac{\partial x^j}{\partial x'^i}\frac{\partial S}{\partial x^j}.
\tag{13}
$$

For example, the covariant components of a scalar S in spherical coordinates are

$$
\left(\frac{\partial S}{\partial r}, \frac{\partial S}{\partial \theta}, \frac{\partial S}{\partial \phi} \right).
\tag{14}
$$

These are not to be confused with the components of the gradient referred to the orthonormal frame $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}})$,

$$
A_r = \hat{\mathbf{r}} \cdot \nabla S = \frac{\partial S}{\partial r},
\tag{15a}
$$

$$
A_\theta = \hat{\boldsymbol{\theta}} \cdot \nabla S = \frac{1}{r}\frac{\partial S}{\partial \theta},
\tag{15b}
$$

$$
A_\phi = \hat{\boldsymbol{\phi}} \cdot \nabla S = \frac{1}{r\sin\theta}\frac{\partial S}{\partial \phi},
\tag{15c}
$$

where $\mathbf{A} = \nabla S$. This equation is equivalent to Eq. (D.20).

Notice that in Eq. (12) we differentiate with respect to x^i (with an upper index), and obtain something (B_i) with a lower index. This is a general rule: when we differentiate with respect to something with indices, the positions of the indices are reversed in expressing the result.

7. Tensors

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

$$T'^{\ell}{}_m{}^n = \frac{\partial x'^{\ell}}{\partial x^i} \frac{\partial x^j}{\partial x'^m} \frac{\partial x'^n}{\partial x^k} T^i{}_j{}^k, \tag{16}$$

where the first index is contravariant, the second covariant, and the third contravariant. Sometimes we use dots as place holders, for example, $T^i{}_{.\,j}{}^k$, 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 product of any two tensors is a tensor, for example, $A_i F^j{}_k$ is a third-rank tensor.

A contravariant or covariant vector is considered a tensor of rank 1, while a scalar is considered a tensor of rank 0.

8. The Kronecker Delta

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,

$$\delta'^i_j = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^{\ell}}{\partial x'^j} \delta^k_{\ell} = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^k}{\partial x'^j} = \delta^i_j, \tag{17}$$

which follows from Eq. (2). But this only works if we define the Kronecker delta with mixed indices. If we attempt to define an object such as δ_{ij} or δ^{ij} as having the value 1 if $i = j$ and 0 if $i \neq j$, requiring these values in all coordinate systems, then the resulting object does not transform as a tensor (either a purely covariant or purely contravariant tensor of rank 2). But see however Sec. 14, for the special case of orthonormal coordinates on a Euclidean space, where δ_{ij} does transform as a tensor.

When writing the Kronecker delta as a mixed tensor, δ^i_j, we do not bother to place one index before the other (or to use dots as place holders), since the tensor is symmetric in i and j.

9. Scalar Products and Contractions

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

$$A^i B_i = A'^i B'_i = A \cdot B. \tag{18}$$

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

The scalar product can be generalized. 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

$$A^i B_i, \qquad M^i{}_i, \qquad T^i{}_i{}^j, \qquad M^i{}_j A^j, \quad \text{etc.} \tag{19}$$

Such a sum is called a *contraction* of indices. Notice that the scalar product of a contravariant vector and a covariant vector is a special case, as is taking the trace of a mixed tensor (thought of as a matrix), or multiplying a matrix (thought of as a mixed tensor) times a vector (thought of as a contravariant vector).

But a sum on two contravariant or two covariant indices, for example,

$$T^i{}_j{}^i, \qquad F_{ii}, \qquad A^i A^i, \quad \text{etc.} \tag{20}$$

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 (20) cannot have a meaning that is independent of coordinate system. Thus we have the rule, that in any covariant expression (taking the same form in all coordinate systems), any contractions of indices must take place between one contravariant and one covariant index.

10. The Jacobian

Let us write

$$J^i{}_j = \frac{\partial x'^i}{\partial x^j} \tag{21}$$

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 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,

$$(J^{-1t})_i{}^j = \frac{\partial x^j}{\partial x'^i}. \tag{22}$$

The Jacobian is an illustration of the fact that not everything with indices on it is a tensor.

11. The Metric Tensor

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

$$ds^2 = g_{ij}\, dx^i dx^j. \tag{23}$$

This defines the metric tensor g_{ij}. 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).

An expression of the form (23) is called the *line element* in the given coordinate system. The line element in various coordinate systems in three-dimensional, Euclidean space is given by Eqs. (D.26). Those equations can be used to read off the components of the metric tensor in the various coordinate systems.

If the metric is positive definite, then g_{ij}, regarded as a matrix of numbers, is invertible. Normally the metric tensor is invertible even when it is not positive definite, for example, in relativity theory. Its inverse is defined to be the metric tensor g^{ij} (with contravariant, rather than covariant, indices). That is, g^{ij} is defined by

$$g_{ij}\, g^{jk} = \delta_i^k. \tag{24}$$

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

12. Raising and Lowering Indices

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):

$$A_i = g_{ij}A^j. \tag{25}$$

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,

$$T_{ij}{}^k = g_{i\ell}\, T^\ell{}_j{}^k, \tag{26}$$

where the first index has been lowered.

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):

$$B^i = g^{ij}B_j. \tag{27}$$

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

$$T^{ijk} = g^{j\ell} \, T^i{}_\ell{}^k, \tag{28}$$

where the middle index has been raised. 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:

$$A^i g_{ij} B^j = A_j B^j = A \cdot B. \tag{29}$$

In this way we define the scalar product of two contravariant vectors. Similarly, we can form the scalar product of two covariant vectors:

$$A_i \, g^{ij} B_j = A^j B_j = A \cdot B. \tag{30}$$

The scalar product of two contravariant vectors or two covariant vectors is not defined unless we have a metric; but the scalar product of a contravariant times a covariant vector is always defined (it does not require a metric).

If we have an contraction between a contravariant and covariant index, we can raise the covariant index with the metric and lower the contravariant index, without changing the result. For example,

$$M^i{}_j \, A^j = M^{ij} \, A_j. \tag{31}$$

13. Vector Spaces With Rectilinear Coordinates

The development up to this point has applied to general, curvilinear coordinate systems. But the case of vector spaces, upon which we use rectlinear coordinates, is common in practice. We now restrict to this case, where there are some simplifications. At first we do not assume that the coordinates are orthonormal, that is, we allow them to be skew.

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

$$x'^i = J^i{}_j \, x^j, \qquad J^i{}_j = \frac{\partial x'^i}{\partial x^j} = \text{const.} \tag{32}$$

In rectilinear coordinate systems, the coordinates x^i themselves (not just their differentials dx^i) 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

$$B^i{}_j = \frac{\partial A^i}{\partial x^j}. \tag{33}$$

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. (33) acts like a subscript above. Sometimes we write

$$\partial_i = \frac{\partial}{\partial x^i}, \tag{34}$$

for example, Eq. (33) can be written

$$B^i{}_j = \partial_j A^i. \tag{35}$$

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

$$F^{ij}{}_{,j} = \partial_j F^{ij} = \frac{\partial F^{ij}}{\partial x^j}. \tag{36}$$

The operator ∂_i can be thought of as the gradient operator, which can be applied to any tensor (not just scalars). The divergence is the gradient ∂_i followed by a contraction on the index i, for example,

$$\frac{\partial A^i}{\partial x^i} = \partial_i A^i = A^i{}_{,i}. \tag{37}$$

The divergence can be applied to any tensor with at least one contravariant index. As defined here, the divergence produces a tensor of rank one less than the original tensor, but this rule only works in rectilinear coordinates, where the derivatives of the Jacobian matrix vanish. A special case is the divergence of a vector field, which is a scalar.

The curl in the usual sense is only meaningful in three-dimensional space, where it requires the properties of the Levi-Civita symbol (see Sec. 15).

14. Three-Dimensional, Euclidean Space

We usually think of physical space as the Euclidean vector space $\mathbb{R}^3$. On this space there exists a rectilinear coordinate system in which the components of the metric are

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{38}$$

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

$$x'^{\,i} = R^i{}_j \, x^j \tag{39}$$

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 rules of tensor analysis simplify. This is because for an orthogonal matrix, $R^{-1t} = R$, so the transformation law for contravariant and covariant indices is identical. Moreover, since the components of the metric tensor form the identity matrix, raising or lowering an index does not change 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. Then we can write $x_i' = R_{ij}\, x_j$, contractions like $A_i A_i = A \cdot A$ are meaningful, the purely covariant Kronecker delta δ_{ij} transforms as a tensor, 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. See Notes 11 for more on orthogonal transformations and rotations.

15. The Levi-Civita Symbol in Three Dimensions

The Levi-Civita symbol (or permutation symbol) is a tensor-like object that can be defined on a space of any number of dimensions. We will consider the Levi-Civita symbol first in three-dimensional, Euclidean space where it is used to construct the cross product and the curl, and where its properties lead to a number of useful identities in vector algebra and vector calculus. We assume the use of orthonormal coordinates, and write lower indices throughout.

The Levi-Civita symbol in three dimensions is defined by

$$\epsilon_{ijk} = \begin{cases} 1, & \text{if } (ijk) \text{ is an even permutation of } (123); \\ -1, & \text{if } (ijk) \text{ is an odd permutation of } (123); \\ 0, & \text{otherwise.} \end{cases} \tag{40}$$

It has $3^3 = 27$ components, of which only 6 are nonzero. It follows directly from this definition that ϵ_{ijk} changes sign if any two of its indices are exchanged,

$$\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij} = -\epsilon_{jik} = -\epsilon_{ikj} = -\epsilon_{kji}. \tag{41}$$

The Levi-Civita symbol is convenient for expressing cross products and curls in tensor notation. Some relations involving the cross product are

$$(\mathbf{A} \times \mathbf{B})_i = \epsilon_{ijk}\, A_j B_k, \tag{42}$$

and

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix} = \epsilon_{ijk}\, A_i B_j C_k. \tag{43}$$

Equation (42) can be taken as the definition of the cross product. Likewise, the curl can be expressed or defined in terms of the Levi-Civita symbol,

$$(\nabla \times \mathbf{B})_i = \epsilon_{ijk}\, \frac{\partial B_k}{\partial x_j} = \epsilon_{ijk}\, \partial_j B_k. \tag{44}$$

This looks just like Eq. (42), with ∇ interpreted as the "vector" with components ∂_i.

Equation (43) shows a relationship between the Levi-Civita symbol and the determinant of a matrix. Let M_{ij} be a 3×3 matrix. Then other relations are

$$\det M = \epsilon_{ijk}\, M_{1i}M_{2j}M_{3k} = \frac{1}{6}\epsilon_{ijk}\,\epsilon_{\ell mn}\, M_{i\ell}M_{jm}M_{kn}, \tag{45}$$

and

$$\epsilon_{ijk}\, M_{i\ell}M_{jm}M_{kn} = (\det M)\,\epsilon_{\ell mn}. \tag{46}$$

Variations on these equations are obtained by replacing M by its transpose and noting that $\det M = \det M^t$.

We are calling ϵ_{ijk} a "symbol" because it does not transform as a tensor, in general. To see this, consider two orthonormal coordinate systems x_i and x_i', connected by $x_i' = R_{ij}\,x_j$, where R is orthogonal. Then according to Eq. (46), we have

$$\epsilon_{ijk}' = R_{i\ell}R_{jm}R_{kn}\,\epsilon_{\ell mn} = (\det R)\,\epsilon_{ijk}. \tag{47}$$

Since $\det R = \pm 1$, we see that ϵ_{ijk} transforms as a tensor under proper orthogonal transformations, but not improper ones.

This is an illustration of the fact that some things are tensors (that is, transform as tensors) only under a restricted class of coordinate transformations. The Levi-Civita symbol ϵ_{ijk} is a tensor only under proper orthogonal transformations; δ_{ij} is a tensor under all orthogonal transformations; and the mixed Kronecker delta, δ_j^i, is a tensor under general 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.

16. The Levi-Civita Symbol and Vector Identities

The definitions (42) and (44) of the cross product and curl can be used to prove vector identities such as

$$\mathbf{A} \cdot (\mathbf{B}\times\mathbf{C}) = \mathbf{B} \cdot (\mathbf{C}\times\mathbf{A}) = \mathbf{C} \cdot (\mathbf{A}\times\mathbf{B}) \tag{48}$$

and

$$\nabla \cdot (\mathbf{A}\times\mathbf{B}) = \mathbf{B} \cdot (\nabla\times\mathbf{A}) - \mathbf{A} \cdot (\nabla\times\mathbf{B}), \tag{49}$$

which contain a single cross product or curl.

Any combination of an even number of Levi-Civita symbols (or an even number of cross products and curls) can be reduced to dot products with the following system of identities. Similarly, any combination of an odd number of Levi-Civita symbols (or an odd number of cross products and

curls) can be reduced to a single Levi-Civita symbol (or a single cross product or a curl) plus dot products. The first is the most general:

$$\epsilon_{ijk}\,\epsilon_{\ell mn} = \begin{vmatrix} \delta_{i\ell} & \delta_{im} & \delta_{in} \\ \delta_{j\ell} & \delta_{jm} & \delta_{jn} \\ \delta_{k\ell} & \delta_{km} & \delta_{kn} \end{vmatrix}. \tag{50}$$

Notice that the indices (ijk) label the rows, while (ℓmn) label the columns. If this is contracted on i and l, we obtain

$$\epsilon_{ijk}\,\epsilon_{imn} = \begin{vmatrix} \delta_{jm} & \delta_{jn} \\ \delta_{km} & \delta_{kn} \end{vmatrix} = \delta_{jm}\,\delta_{kn} - \delta_{jn}\,\delta_{km}. \tag{51}$$

By contracting Eq. (51) in j and m we obtain

$$\epsilon_{ijk}\,\epsilon_{ijn} = 2\,\delta_{kn}. \tag{52}$$

Finally, contracting on k and n we obtain

$$\epsilon_{ijk}\,\epsilon_{ijk} = 6. \tag{53}$$

The identity (51) is the one used most often, for boiling down two cross products that have one index in common. For example, it can be used to derive

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}(\mathbf{A}\cdot\mathbf{B}) \tag{54}$$

(the "back minus cab" identity), and

$$\mathbf{A}\times(\nabla\times\mathbf{B}) = (\nabla\mathbf{B})\cdot\mathbf{A} - \mathbf{A}\cdot\nabla\mathbf{B}. \tag{55}$$

In this identity we use a notation in which $\nabla\mathbf{B}$ is interpreted as a matrix, see Sec. 17. Other identities with two cross products or curls include

$$\nabla\times(\mathbf{A}\times\mathbf{B}) = \mathbf{B}\cdot\nabla\mathbf{A} + \mathbf{A}(\nabla\cdot\mathbf{B}) - \mathbf{A}\cdot\nabla\mathbf{B} - \mathbf{B}(\nabla\cdot\mathbf{A}), \tag{56}$$

$$\nabla\times(\nabla\times\mathbf{A}) = \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A}, \tag{57}$$

$$\nabla\cdot(\mathbf{A}\times\mathbf{B}) = \mathbf{B}\cdot(\nabla\times\mathbf{A}) - \mathbf{A}\cdot(\nabla\times\mathbf{B}), \tag{58}$$

and

$$(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{C}\times\mathbf{D}) = \begin{vmatrix} \mathbf{A}\cdot\mathbf{C} & \mathbf{A}\cdot\mathbf{D} \\ \mathbf{B}\cdot\mathbf{C} & \mathbf{B}\cdot\mathbf{D} \end{vmatrix}. \tag{59}$$

The identity (51) can also be used to convert expressions involving antisymmetric matrices or tensors in three-dimensions into the language of vectors. Let $A_{ij} = -A_{ji}$ be an antisymmetric, 3×3 tensor. It has 3 independent components that we can associate with a 3-vector $\mathbf{A}$, as follows:

$$A_{ij} = \begin{pmatrix} 0 & A_3 & -A_2 \\ -A_3 & 0 & A_1 \\ A_2 & -A_1 & 0 \end{pmatrix} = \epsilon_{ijk}\,A_k. \tag{60}$$

The inverse of this is

$$A_{ij} = \frac{1}{2}\epsilon_{ijk}\,A_k. \tag{61}$$

Using thes identities, the multiplication of an antisymmetric matrix times a vector can be reexpressed in terms of a cross product. That is, if

$$X_i = A_{ij}\,Y_j \tag{62}$$

then

$$\mathbf{X} = \mathbf{Y}{\times}\mathbf{A}. \tag{63}$$

Similarly, if **A** and **B** are two vectors, then

$$A_i B_j - A_j B_i = \epsilon_{ijk}\,(\mathbf{A}{\times}\mathbf{B})_k, \tag{64}$$

and

$$\frac{\partial B_j}{\partial x_i} - \frac{\partial B_i}{\partial x_j} = \epsilon_{ijk}\,(\nabla{\times}\mathbf{B})_k. \tag{65}$$

17. Notation for $\nabla\mathbf{B}$ as a Matrix

In Eq. (55), the term $(\nabla\mathbf{B})\cdot\mathbf{A}$ can be interpreted as the vector resulting from multiplying the matrix $\nabla\mathbf{B}$ times the vector $\mathbf{A}$, where $\nabla\mathbf{B}$ has the components

$$(\nabla\mathbf{B})_{ij} = \partial_i B_j = \frac{\partial B_j}{\partial x_i}. \tag{66}$$

The parentheses in $(\nabla\mathbf{B})\cdot\mathbf{A}$ are used to indicate that the ∇ acts only on **B**, and to distinguish this expression from $\nabla(\mathbf{B}\cdot\mathbf{A})$, the gradient of a scalar.

If we multiply the vector **A** on the other side of the matrix $\nabla\mathbf{B}$, we get

$$\mathbf{A}\cdot\nabla\mathbf{B}, \tag{67}$$

which is fairly standard notation for the directional derivative of a vector, although some books would write $(\mathbf{A}\cdot\nabla)\mathbf{B}$ to make it clear that the **A** is contracted with the ∇.

On the other hand, the notation $(\nabla\mathbf{B})\cdot\mathbf{A}$ is nonstandad, and is not used outside these notes, as far as I know. Many books write

$$\mathbf{A}{\times}(\nabla{\times}\mathbf{B}) + \mathbf{B}\cdot\nabla\mathbf{A} \tag{68}$$

for this quantity, making use of the identity (55). In our notation the gradient of a dot product can be expanded,

$$\nabla(\mathbf{A}\cdot\mathbf{B}) = (\nabla\mathbf{A})\cdot\mathbf{B} + (\nabla\mathbf{B})\cdot\mathbf{A}. \tag{69}$$

18. Lorentz Transformations

We turn now to tensor analysis in special relativity. In special relativity we conceive of space-time as the space $\mathbb{R}^4$, with coordinates

$$x^\mu = \begin{pmatrix} ct \\ \mathbf{x} \end{pmatrix}. \tag{70}$$

where $\mu = 0, 1, 2, 3$. Here we follow the common custom of using Greek indices for relativity theory. The factor of c on the first coordinate $x^0 = ct$ makes all four coordinates have dimensions of distance (and there are dimensional simplifications if we choose units such that $c = 1$).

The physical meaning of these coordinates is the following. The spatial origin of the coordinates $x = y = z = 0$ coincides with an observer or physical system upon which no forces act (an unaccelerated observer). The time coordinate is measured by a clock carried by this observer. Clocks at different locations are synchronized by a certain procedure. The spatial coordinates are distances measured relative to a set of mutually orthogonal, spatial axes, which we assume are right-handed. It is assumed that free particle orbits are straight lines with constant velocity with respect to this system of coordinates. This excludes rotating spatial axes. This coordinate system constitutes what is called a *Lorentz frame.*

Different unaccelerated observers and/or different choices of spatial axes constitute different Lorentz frames. For simplicity we only consider the case that the space-time origins ($x = y = z = ct = 0$) of the Lorentz frames coincide. Then any two systems of coordinates x^μ and x'^μ, corresponding to two Lorentz frames, are related by a linear transformation,

$$x'^\mu = \Lambda^\mu{}_\nu\, x^\nu, \tag{71}$$

where $\Lambda^\mu{}_\nu$ is constant (independent of space or time). This transformation (or the matrix $\Lambda^\mu{}_\nu$) is called a *Lorentz transformation.* The form of the matrix $\Lambda^\mu{}_\nu$ can be determined from a set of assumptions (the constancy of the speed of light, arguments of homogeneity and isotropy, etc). Then $\Lambda^\mu{}_\nu$ turns out to depend on a set of parameters (the velocity connecting the two Lorentz frames and the Euler angles of a rotation connecting the spatial axes).

19. The Minkowski Metric

Lorentz transformations derived in this way have an important property, that the square interval between two events,

$$c^2\,\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2, \tag{72}$$

is an invariant. Here $\Delta x^\mu = x_2^\mu - x_1^\mu$, where x_1^μ and x_2^μ are the coordinates of the two events. That is, all observers, using their own Lorentz frames and corresponding coordinate systems, agree on the value of the quantity (72). Although we call this quantity a "square interval," numerically it can have any sign, positive, negative or zero. Intervals for which this quantity is positive, negative or zero are called time-like, space-like or light-like, respectively.

If the two events are infinitesimally close then we can replace Δx^μ by dx^μ for the infinitesimal coordinate increments, and write

$$g_{\mu\nu}\, dx^\mu dx^\nu \tag{73}$$

for the invariant quantity, where

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{74}$$

Since the quantity (72) is the same in all Lorentz frames, we take $g_{\mu\nu}$ to have the same values (74) in all Lorentz frames. Then it turns out that $g_{\mu\nu}$ transforms as a tensor under Lorentz transformations. For the invariance of the square interval implies

$$g_{\mu\nu}\, dx'^\mu dx'^\nu = g_{\alpha\beta}\, dx^\alpha dx^\beta, \tag{75}$$

where x^μ and x'^μ are connected by the Lorentz transformation (71). But by the chain rule we have

$$dx'^\mu = \frac{\partial x'^\mu}{\partial x^\alpha}\, dx^\alpha, \tag{76}$$

so Eq. (75) becomes

$$g_{\mu\nu} \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x'^\nu}{\partial x^\beta} dx^\alpha dx^\beta = g_{\alpha\beta}\, dx^\alpha dx^\beta. \tag{77}$$

But since dx^α is arbitrary, this implies

$$g'_{\alpha\beta} = g_{\mu\nu} \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x'^\nu}{\partial x^\beta} = g_{\alpha\beta}, \tag{78}$$

where $g'_{\alpha\beta}$ is the value determined by the transformation law for covariant tensors. It agrees with the value assigned by Eq. (74). Notice incidentally that the Jacobian matrix is the same as the matrix defining the Lorentz transformation,

$$\frac{\partial x'^\mu}{\partial x^\nu} = \Lambda^\mu{}_\nu. \tag{79}$$

Let us write

$$ds^2 = g_{\mu\nu}\, dx^\mu dx^\nu \tag{80}$$

for the invariant quantity (73). This has the same mathematical form as Eq. (23), the square of the distance between two nearby points on a manifold. For this reason we call $g_{\mu\nu}$ the metric tensor in special relativity, or the *Minkowski metric* to be more specific. However, if ds is the distance in the usual sense between two points, then $ds^2 \geq 0$, but we have seen that the invariant (80) can have any sign. This is obviously because the spatial and time coordinates enter with opposite signs in Eq. (72). This should not be surprising; after all, space is physically different from time. What is surprising is that space-time has a metric at all. In a sense, the fact that it does is the great discovery of special relativity.

Given the covariant metric tensor (74), we can compute the contravariant metric tensor as the inverse matrix. It is

$$g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{81}$$

That is, the components of the contravariant metric tensor are the same, numerically speaking, as the components of the covariant metric tensor. This is similar to the situation on a Euclidean vector space in orthonormal coordinates, where $g^{ij} = g_{ij} = \delta_{ij}$. However, unlike the case of a Euclidean vector space, the Minkowski metric is not the identity matrix. Therefore the contravariant and covariant components of a vector or tensor are not numerically equal, and one must maintain the distinction between contravariant and covariant components. Switching between the two, in the case of a contravariant or covariant vector, amounts to simply changing the sign of the spatial parts of a vector.

20. The Sign of the Metric, and a Notational Problem

If the quantity (72) is an invariant, then so is its negative. There is no physical significance to the choice of the overall sign of this quantity; the only thing that matters is the opposite signs with which the spatial and time coordinates enter. In these Notes we will use the metric given by Eq. (74), but many books would define $g_{\mu\nu}$ with an opposite sign. Our sign convention follows that of Jackson, *Classical Electrodynamics*, 3rd edition, and is commonly used in the physics literature on special relativity. Unfortunately, it gives rise to some notational difficulties.

As pointed out in Sec. 14, in three-dimensional, Euclidean space with orthonormal coordinates, there is no need to maintain the distinction between contravariant and covariant indices, and we normally just use lower indices everywhere, for example, in expressing the components p_i of the usual momentum vector $\mathbf{p}$. But $\mathbf{p}$ is also the spatial part of the contravariant 4-momentum vector p^μ in special relativity, that is,

$$p^\mu = \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix}, \tag{82}$$

where the 4-vector is shown as a column vector. Thus the covariant form of the 4-momentum is

$$p_\mu = \begin{pmatrix} E/c \\ -\mathbf{p} \end{pmatrix}. \tag{83}$$

Now setting $\mu = i = 1, 2, 3$ in this to select the spatial components of p_μ, and using the usual custom in 3-dimensional space for expressing the components of the momentum $\mathbf{p}$, we obtain $p_i = -p_i$. A similar problem arises with the 4-vector potential A^μ in electromagnetism, for which

$$A^\mu = \begin{pmatrix} \Phi \\ \mathbf{A} \end{pmatrix}, \tag{84}$$

where Φ and $\mathbf{A}$ are the usual scalar and vector potentials.

Obviously we must distinguish the spatial parts of the covariant components of a 4-vector from the usual Cartesian components of a 3-vector, expressed with the usual custom of using lower indices. In these Notes we will use the following notation to deal with this problem. If we have a 4-vector such as p_μ, and we set a Greek index to a Latin one to indicate the spatial components, we will write something like $p_{\mu \to i}$ for the result. On the other hand, the usual components of the standard 3-vector will be denoted by something like p_i. For example, in the case of the momentum 4-vector, we can write,

$$p^{\mu \to i} = p_i, \qquad p_{\mu \to i} = -p_i. \tag{85}$$

21. The $x_4 = ict$ Convention

In older references on special relativity it is common to use a coordinate system x_μ on space-time, in which x_i, $i = 1, 2, 3$ are the usual rectangular coordinates (x, y, z) on space, while $x_4 = ict$. The purely imaginary time coordinate x_4 replaces the real coordinate $x^0 = ct$ used in these notes. Thus the invariant interval can be written in Euclidean form,

$$\Delta \mathbf{x}^2 - c^2 \Delta t^2 = \sum_{\mu=1}^{4} \Delta x_\mu^2. \tag{86}$$

That is, $g_{\mu\nu} = \delta_{\mu\nu}$ in these coordinates. This has the advantage that there is no need to distinguish between contravariant and covariant indices, and lower indices can be used throughout. With this convention, Lorentz transformations, which preserve the Euclidean form of the metric, are orthogonal, but they are not real, in general. This is because boosts turn out to be rotations by an imaginary angle.

The disadvantage of this convention is that physical time is real, not imaginary, and it is awkward to keep track of the factors of i that appear everywhere. Nowadays the trend is to use the real Minkowski metric (74) (with some choice of overall sign), and to maintain the distinction between contravariant and covariant indices. This distinction is necessary anyway in general relativity, which is becoming more central and important in physics. Sakurai uses the $x_4 = ict$ convention in his book *Advanced Quantum Mechanics* as does Jackson in *Classical Electrodynamics*, 1st ed, but the real metric is used in the 3rd edition of Jackson, as well as by Bjorken and Drell, *Relativistic Quantum Mechanics* and most other modern references. These notes use the real metric (74) throughout.

22. Electromagnetism in Special Relativity

We summarize here our notation and conventions for electromagnetism in special relativity.

The 4-vector potential is given in terms of the scalar potential Φ and the 3-vector potential $\mathbf{A}$ by

$$A^\mu = \begin{pmatrix} \Phi \\ \mathbf{A} \end{pmatrix}, \qquad A_\mu = \begin{pmatrix} \Phi \\ -\mathbf{A} \end{pmatrix}. \tag{87}$$

The 4-current is given in terms of the charge density ρ and the 3-vector current $\mathbf{J}$ by

$$J^\mu = \begin{pmatrix} c\rho \\ \mathbf{J} \end{pmatrix}, \qquad J_\mu = \begin{pmatrix} c\rho \\ -\mathbf{J} \end{pmatrix}. \tag{88}$$

In these definitions we insert factors of c to make all the components have the same physical dimensions, as is necessary if they are to transform as 4-vectors under Lorentz transformations. The same was done with the definitions of x^μ and p^μ [see Eq. (70) and (B.53)]. The continuity equation (conservation of charge) is

$$\partial_\mu J^\mu = \frac{\partial J^\mu}{\partial x^\mu} = 0. \tag{89}$$

The field tensor is defined by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}, \tag{90}$$

where the first index μ on $F_{\mu\nu}$ labels the rows of the matrix and the second the columns. Raising one index gives the mixed tensor,

$$F^\mu{}_\nu = \begin{pmatrix} 0 & E_x & E_y & E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}, \tag{91}$$

while raising both indices gives

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{92}$$

The inhomogeneous Maxwell's equations are

$$\partial_\mu F^{\mu\nu} = \frac{\partial F^{\mu\nu}}{\partial x^\mu} = \frac{4\pi}{c} J^\nu, \tag{93}$$

while the homogeneous ones are

$$\partial_\mu F_{\alpha\beta} + \partial_\alpha F_{\beta\mu} + \partial_\beta F_{\mu\alpha} = 0. \tag{94}$$

Finally, the equations of motion of a particle of charge q in the field given by $F^{\mu\nu}$ are

$$m\frac{d^2 x^\mu}{d\tau^2} = \frac{dp^\mu}{d\tau} = \frac{q}{c} F^\mu{}_\nu \frac{dx^\nu}{d\tau}. \tag{95}$$

Here p^μ is the kinetic momentum defined in Eq. (82) and τ is the proper time.

Problems

1 Show that the contraction of a contravariant vector times a covariant vector transforms as a scalar. See Eq. (18).

2. Some problems on the metric tensor.

(a) Show that the metric tensor g_{ij}, defined by Eq. (23) in all coordinate systems, transforms as a second-rank, purely covariant tensor.

(b) Given that g_{ij} transforms as a second-rank, purely covariant tensor, show that g^{ij}, defined by Eq. (24), transforms as a second-rank, purely contravariant tensor.

(c) Tensors g_{ij} and g^{ij} are regarded as the purely covariant and purely contravariant versions of the metric tensor. What do we get if we raise the index on one of the components of g_{ij}, to obtain a mixed version of the metric tensor? What do we get if we raise both indices on g_{ij}?

(d) Prove Eq. (31).

3. Let A^i be a contravariant vector, and let $B_i{}^j$ be defined by Eq. (33). Show that under general (curvilinear coordinate) transformations, $B_i{}^j$ does not transform as a (mixed) tensor, but that under linear transformations it does.