3.1 The metric tensor

Consider the representation of two vectors $\vec{A}$ and $\vec{B}$ on the basis $\{\vec{e}_\alpha\}$ of some frame $\mathcal{O}$:

$$\vec{A} = A^\alpha \vec{e}_\alpha, \quad \vec{B} = B^\beta \vec{e}_\beta.$$ 

Their scalar product is

$$\vec{A} \cdot \vec{B} = (A^\alpha \vec{e}_\alpha) \cdot (B^\beta \vec{e}_\beta).$$

(Note the importance of using different indices $\alpha$ and $\beta$ to distinguish the first summation from the second.) Following Exer. 34, §2.9, we can rewrite this as

$$\vec{A} \cdot \vec{B} = A^\alpha B^\beta (\vec{e}_\alpha \cdot \vec{e}_\beta),$$

which, by Eq. (2.27), is

$$\vec{A} \cdot \vec{B} = A^\alpha B^\beta \eta_{\alpha\beta}.$$ \hspace{1cm} (3.1)

This is a frame-invariant way of writing

$$-A^0B^0 + A^1B^1 + A^2B^2 + A^3B^3.$$ 

The numbers $\eta_{\alpha\beta}$ are called ‘components of the metric tensor’. We will justify this name later. Right now we observe that they essentially give a ‘rule’ for associating with two vectors $\vec{A}$ and $\vec{B}$ a single number, which we call their scalar product. The rule is that the number is the double sum $A^\alpha B^\beta \eta_{\alpha\beta}$. Such a rule is at the heart of the meaning of ‘tensor’, as we now discuss.

3.2 Definition of tensors

We make the following definition of a tensor:

A tensor of type $\binom{0}{N}$ is a function of $N$ vectors into the real numbers, which is linear in each of its $N$ arguments.
Let us see what this definition means. For the moment, we will just accept the notation \( \mathbf{N} \); its justification will come later in this chapter. The rule for the scalar product, Eq. (3.1), satisfies our definition of a \((0|2)\) tensor. It is a rule which takes two vectors, \( \mathbf{A} \) and \( \mathbf{B} \), and produces a single real number \( \mathbf{A} \cdot \mathbf{B} \). To say that it is linear in its arguments means what is proved in Exer. 34, § 2.9. Linearity on the first argument means

\[
(\alpha \mathbf{A}) \cdot \mathbf{B} = \alpha (\mathbf{A} \cdot \mathbf{B}),
\]

and

\[
(\mathbf{A} + \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot \mathbf{C} + \mathbf{B} \cdot \mathbf{C},
\]

while linearity on the second argument means

\[
\mathbf{A} \cdot (\beta \mathbf{B}) = \beta (\mathbf{A} \cdot \mathbf{B}),
\]

\[
\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}.
\]

This definition of linearity is of central importance for tensor algebra, and the student should study it carefully.

To give concreteness to this notion of the dot product being a tensor, we introduce a name and notation for it. We let \( g \) be the metric tensor and write, by definition,

\[
g(\mathbf{A}, \mathbf{B}) := \mathbf{A} \cdot \mathbf{B}.
\]

Then we regard \( g(, ) \) as a function which can take two arguments, and which is linear in that

\[
g(\alpha \mathbf{A} + \beta \mathbf{B}, \mathbf{C}) = \alpha g(\mathbf{A}, \mathbf{C}) + \beta g(\mathbf{B}, \mathbf{C}),
\]

and similarly for the second argument. The value of \( g \) on two arguments, denoted by \( g(\mathbf{A}, \mathbf{B}) \), is their dot product, a real number.

Notice that the definition of a tensor does not mention components of the vectors. A tensor must be a rule which gives the same real number independently of the reference frame in which the vectors’ components are calculated. We showed in the previous chapter that Eq. (3.1) satisfies this requirement. This enables us to regard a tensor as a function of the vectors themselves rather than of their components, and this can sometimes be helpful conceptually.

Notice that an ordinary function of position, \( f(t, x, y, z) \), is a real-valued function of no vectors at all. It is therefore classified as a \((0|0)\) tensor.

**Aside on the usage of the term ‘function’**

The most familiar notion of a function is expressed in the equation

\[
y = f(x),
\]

where \( y \) and \( x \) are real numbers. But this can be written more precisely as: \( f \) is a ‘rule’ (called a mapping) which associates a real number (symbolically called \( y \), above) with another real number, which is the argument of \( f \) (symbolically called \( x \), above). The function itself is \textit{not} \( f(x) \), since \( f(x) \) is \( y \), which is a real number called the ‘value’ of the
function. The function itself is \( f \), which we can write as \( f(\ ) \) in order to show that it has one argument. In algebra this seems like hair-splitting since we unconsciously think of \( x \) and \( y \) as two things at once: they are, on the one hand, specific real numbers and, on the other hand, \( \textit{names} \) for general and arbitrary real numbers. In tensor calculus we will make this distinction explicit: \( \vec{A} \) and \( \vec{B} \) are \textit{specific} vectors, \( \vec{A} \cdot \vec{B} \) is a specific real number, and \( g \) is the name of the function that associates \( \vec{A} \cdot \vec{B} \) with \( \vec{A} \) and \( \vec{B} \).

### Components of a tensor

Just like a vector, a tensor has components. They are defined as:

The components in a frame \( O \) of a tensor of type \( (0, N) \) are the values of the function when its arguments are the basis vectors \( \{\vec{e}_\alpha\} \) of the frame \( O \).

Thus we have the notion of components as frame-dependent numbers (frame-dependent because the basis refers to a specific frame). For the metric tensor, this gives the components as

\[
g(\vec{e}_\alpha, \vec{e}_\beta) = \vec{e}_\alpha \cdot \vec{e}_\beta = \eta_{\alpha\beta}. \tag{3.5} \]

So the matrix \( \eta_{\alpha\beta} \) that we introduced before is to be thought of as an array of the components of \( g \) on the basis. In another basis, the components could be different. We will have many more examples of this later. First we study a particularly important class of tensors.

### 3.3 The \( (0, 1) \) tensors: one-forms

A tensor of the type \( (0, 1) \) is called a covector, a covariant vector, or a one-form. Often these names are used interchangeably, even in a single text-book or reference.

#### General properties

Let an arbitrary one-form be called \( \tilde{\eta} \). (We adopt the notation that \( \sim \) above a symbol denotes a one-form, just as \( \sim \) above a symbol denotes a vector.) Then \( \tilde{\eta} \), supplied with one vector argument, gives a real number: \( \tilde{\eta}(\vec{A}) \) is a real number. Suppose \( \tilde{\eta} \) is another one-form. Then we can define

\[
\begin{align*}
\tilde{s} &= \tilde{\eta} + \tilde{\eta} \\
\tilde{r} &= \alpha \tilde{\eta},
\end{align*} \tag{3.6a}
\]
to be the one-forms that take the following values for an argument \( \vec{A} \):

\[
\begin{align*}
\tilde{s}(\vec{A}) &= \tilde{p}(\vec{A}) + \tilde{q}(\vec{A}), \\
\tilde{r}(\vec{A}) &= \alpha \tilde{p}(\vec{A}).
\end{align*}
\]

(3.6b)

With these rules, the set of all one-forms satisfies the axioms for a vector space, which accounts for their other names. This space is called the ‘dual vector space’ to distinguish it from the space of all vectors such as \( \vec{A} \).

When discussing vectors we relied heavily on components and their transformations. Let us look at those of \( \tilde{p} \). The components of \( \tilde{p} \) are called \( p_\alpha \):

\[ p_\alpha := \tilde{p}(\vec{e}_\alpha). \]  

(3.7)

Any component with a single lower index is, by convention, the component of a one-form; an upper index denotes the component of a vector. In terms of components, \( \tilde{p}(\vec{A}) \) is

\[ \tilde{p}(\vec{A}) = \tilde{p}(A^\alpha \vec{e}_\alpha) = A^\alpha \tilde{p}(\vec{e}_\alpha), \]

\[ \tilde{p}(\vec{A}) = A^\alpha p_\alpha. \]  

(3.8)

The second step follows from the linearity which is the heart of the definition we gave of a tensor. So the real number \( \tilde{p}(\vec{A}) \) is easily found to be the sum \( A^0 p_0 + A^1 p_1 + A^2 p_2 + A^3 p_3 \). Notice that all terms have plus signs: this operation is called contraction of \( \vec{A} \) and \( \tilde{p} \), and is more fundamental in tensor analysis than the scalar product because it can be performed between any one-form and vector without reference to other tensors. We have seen that two vectors cannot make a scalar (their dot product) without the help of a third tensor, the metric.

The components of \( \tilde{p} \) on a basis \( \{\vec{e}_\beta\} \) are

\[ p_\beta := \tilde{p}(\vec{e}_\beta) = \tilde{p}(\Lambda^\alpha_\beta \vec{e}_\alpha) = \Lambda^\alpha_\beta \tilde{p}(\vec{e}_\alpha) = \Lambda^\alpha_\beta p_\alpha. \]  

(3.9)

Comparing this with

\[ \vec{e}_\beta = \Lambda^\alpha_\beta \vec{e}_\alpha, \]

we see that components of one-forms transform in exactly the same manner as basis vectors and in the opposite manner to components of vectors. By ‘opposite’, we mean using the inverse transformation. This use of the inverse guarantees that \( A^\alpha p_\alpha \) is frame independent for any vector \( \vec{A} \) and one-form \( \tilde{p} \). This is such an important observation that we shall prove it explicitly:

\[
A^\alpha p_\alpha = (\Lambda^\alpha_\beta A^\beta)(\Lambda^\mu_\alpha p_\mu),
\]

(3.10a)

\[ = \Lambda^\mu_\alpha \Lambda^\alpha_\beta A^\beta p_\mu, \]

(3.10b)

\[ = \delta^\mu_\beta A^\beta p_\mu, \]

(3.10c)

\[ = A^\beta p_\beta. \]  

(3.10d)
Tensor analysis in special relativity

(This is the same way in which the vector \( A^\alpha \vec{e}_\alpha \) is kept frame independent.) This inverse transformation gives rise to the word ‘dual’ in ‘dual vector space’. The property of transforming with basis vectors gives rise to the co in ‘covariant vector’ and its shorter form ‘covector’. Since components of ordinary vectors transform oppositely to basis vectors (in order to keep \( A^\beta \vec{e}_\beta \) frame independent), they are often called ‘contravariant’ vectors. Most of these names are old-fashioned; ‘vectors’ and ‘dual vectors’ or ‘one-forms’ are the modern names. The reason that ‘co’ and ‘contra’ have been abandoned is that they mix up two very different things: the transformation of a basis is the expression of new vectors in terms of old ones; the transformation of components is the expression of the same object in terms of the new basis. It is important for the student to be sure of these distinctions before proceeding further.

Basis one-forms

Since the set of all one-forms is a vector space, we can use any set of four linearly independent one-forms as a basis. (As with any vector space, one-forms are said to be linearly independent if no nontrivial linear combination equals the zero one-form. The zero one-form is the one whose value on any vector is zero.) However, in the previous section we have already used the basis vectors \( \{ \vec{e}_\alpha \} \) to define the components of a one-form. This suggests that we should be able to use the basis vectors to define an associated one-form basis \( \{ \tilde{\omega}^\alpha, \alpha = 0, \ldots, 3 \} \), which we shall call the basis dual to \( \{ \vec{e}_\alpha \} \), upon which a one-form has the components defined above. That is, we want a set \( \{ \tilde{\omega}^\alpha \} \) such that

\[
\tilde{p} = p_\alpha \tilde{\omega}^\alpha. \tag{3.11}
\]

(Notice that using a raised index on \( \tilde{\omega}^\alpha \) permits the summation convention to operate.) The \( \{ \tilde{\omega}^\alpha \} \) are four distinct one-forms, just as the \( \{ \vec{e}_\alpha \} \) are four distinct vectors. This equation must imply Eq. (3.8) for any vector \( \vec{A} \) and one-form \( \tilde{p} \):

\[
\tilde{p}(\vec{A}) = p_\alpha A^\alpha. 
\]

But from Eq. (3.11) we get

\[
\tilde{p}(\vec{A}) = p_\alpha \tilde{\omega}^\alpha (\vec{A}) \\
= p_\alpha \tilde{\omega}^\alpha (A^\beta \vec{e}_\beta) \\
= p_\alpha A^\beta \tilde{\omega}^\alpha (\vec{e}_\beta). 
\]

(Notice the use of \( \beta \) as an index in the second line, in order to distinguish its summation from the one on \( \alpha \).) Now, this final line can only equal \( p_\alpha A^\alpha \) for all \( A^\beta \) and \( p_\alpha \) if

\[
\tilde{\omega}^\alpha (\vec{e}_\beta) = \delta^\alpha_\beta. \tag{3.12}
\]

Comparing with Eq. (3.7), we see that this equation gives the \( \beta \)th component of the \( \alpha \)th basis one-form. It therefore defines the \( \alpha \)th basis one-form. We can write out these components as
3.3 The \((\mathbf{i})\) tensors: one-forms

\[
\tilde{\omega}^0 \rightarrow (1, 0, 0, 0),
\]
\[
\tilde{\omega}^1 \rightarrow (0, 1, 0, 0),
\]
\[
\tilde{\omega}^2 \rightarrow (0, 0, 1, 0),
\]
\[
\tilde{\omega}^3 \rightarrow (0, 0, 0, 1).
\]

It is important to understand two points here. One is that Eq. (3.12) defines the basis \(\{\tilde{\omega}^\alpha\}\) in terms of the basis \(\{\tilde{e}^\beta\}\). The vector basis induces a unique and convenient one-form basis. This is not the only possible one-form basis, but it is so useful to have the relationship, Eq. (3.12), between the bases that we will always use it. The relationship, Eq. (3.12), is between the two bases, not between individual pairs, such as \(\tilde{\omega}^0\) and \(\tilde{e}^0\). That is, if we change \(\tilde{e}^0\), while leaving \(\tilde{e}^1\), \(\tilde{e}^2\), and \(\tilde{e}^3\) unchanged, then in general this induces changes not only in \(\tilde{\omega}^0\) but also in \(\tilde{\omega}^1\), \(\tilde{\omega}^2\), and \(\tilde{\omega}^3\). The second point to understand is that, although we can describe both vectors and one-forms by giving a set of four components, their geometrical significance is very different. The student should not lose sight of the fact that the components tell only part of the story. The basis contains the rest of the information. That is, a set of numbers \((0, 2, -1, 5)\) alone does not define anything; to make it into something, we must say whether these are components on a vector basis or a one-form basis and, indeed, which of the infinite number of possible bases is being used.

It remains to determine how \(\{\tilde{\omega}^\alpha\}\) transforms under a change of basis. That is, each frame has its own unique set \(\{\tilde{\omega}^\alpha\}\); how are those of two frames related? The derivation here is analogous to that for the basis vectors. It leads to the only equation we can write down with the indices in their correct positions:

\[
\tilde{\omega}^{\bar{\alpha}} = \Lambda^{\bar{\alpha}}_{\beta} \tilde{\omega}^\beta. \tag{3.13}
\]

This is the same as for components of a vector, and opposite that for components of a one-form.

**Picture of a one-form**

For vectors we usually imagine an arrow if we need a picture. It is helpful to have an image of a one-form as well. First of all, it is not an arrow. Its picture must reflect the fact that it maps vectors into real numbers. A vector itself does not automatically map another vector into a real number. To do this it needs a metric tensor to define the scalar product. With a different metric, the same two vectors will produce a different scalar product. So two vectors by themselves don’t give a number. We need a picture of a one-form which doesn’t depend on any other tensors having been defined. The one generally used by mathematicians is shown in Fig. 3.1. The one-form consists of a series of surfaces. The ‘magnitude’ of it is given by the spacing between the surfaces: the larger the spacing the smaller the magnitude. In this picture, the number produced when a one-form acts on a vector is the number of surfaces that the arrow of the vector pierces. So the closer their
Figure 3.1
(a) The picture of one-form complementary to that of a vector as an arrow. (b) The value of a one-form on a given vector is the number of surfaces the arrow pierces. (c) The value of a smaller one-form on the same vector is a smaller number of surfaces. The larger the one-form, the more ‘intense’ the slicing of space in its picture.

spacing, the larger the number (compare (b) and (c) in Fig. 3.1). In a four-dimensional space, the surfaces are three-dimensional. The one-form doesn’t define a unique direction, since it is not a vector. Rather, it defines a way of ‘slicing’ the space. In order to justify this picture we shall look at a particular one-form, the gradient.

### Gradient of a function is a one-form

Consider a scalar field $\phi(\vec{x})$ defined at every event $\vec{x}$. The world line of some particle (or person) encounters a value of $\phi$ at each on it (see Fig. 3.2), and this value changes from event to event. If we label (parametrize) each point on the curve by the value of proper time $\tau$ along it (i.e. the reading of a clock moving on the line), then we can express the coordinates of events on the curve as functions of $\tau$:

$$\begin{align*}
t &= t(\tau), \\
x &= x(\tau), \\
y &= y(\tau), \\
z &= z(\tau).
\end{align*}$$

The four-velocity has components

$$\vec{U} \rightarrow \left( \frac{dt}{d\tau}, \frac{dx}{d\tau}, \ldots \right).$$

Since $\phi$ is a function of $t, x, y,$ and $z$, it is implicitly a function of $\tau$ on the curve:

$$\phi(\tau) = \phi[t(\tau), x(\tau), y(\tau), z(\tau)],$$

and its rate of change on the curve is
### 3.3 The \( (0)^i \) tensors: one-forms

The tensors: one-forms

\[
\tau = 0 \\
\tau = 1 \\
\tau = 2 \\
\rightarrow U \\
\phi(\tau) = \phi(t(\tau), x(\tau), y(\tau), z(\tau))
\]

**Figure 3.2** A world line parametrized by proper time \( \tau \), and the values \( \phi(\tau) \) of the scalar field \( \phi(t, x, y, z) \) along it.

\[
\frac{d\phi}{d\tau} = \frac{\partial \phi}{\partial t} u^t + \frac{\partial \phi}{\partial x} u^x + \frac{\partial \phi}{\partial y} u^y + \frac{\partial \phi}{\partial z} u^z. \tag{3.14}
\]

It is clear from this that in the last equation we have devised a means of producing from the vector \( \vec{U} \) the number \( \frac{d\phi}{d\tau} \) that represents the rate of change of \( \phi \) on a curve on which \( \vec{U} \) is the tangent. This number \( \frac{d\phi}{d\tau} \) is clearly a linear function of \( \vec{U} \), so we have defined a one-form.

By comparison with Eq. (3.8), we see that this one-form has components \( (\partial \phi/\partial t, \partial \phi/\partial x, \partial \phi/\partial y, \partial \phi/\partial z) \). This one-form is called the **gradient** of \( \phi \), denoted by \( \vec{d}\phi \):

\[
\vec{d}\phi \to \left( \frac{\partial \phi}{\partial t}, \frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial y}, \frac{\partial \phi}{\partial z} \right). \tag{3.15}
\]

It is clear that the gradient fits our definition of a one-form. We will see later how it comes about that the gradient is usually introduced in three-dimensional vector calculus as a vector.

The gradient enables us to justify our picture of a one-form. In Fig. 3.3 we have drawn part of a topographical map, showing contours of equal elevation. If \( h \) is the elevation, then the gradient \( \vec{d}h \) is clearly largest in an area such as \( A \), where the lines are closest together, and smallest near \( B \), where the lines are spaced far apart. Moreover, suppose we wanted to know how much elevation a walk between two points would involve. We would lay out on the map a line (vector \( \Delta \vec{x} \)) between the points. Then the number of contours the line crossed would give the change in elevation. For example, line 1 crosses \( 1 \frac{1}{2} \) contours, while 2 crosses two contours. Line 3 starts near 2 but goes in a different direction, winding up only \( 1 \frac{1}{2} \) contour higher. But these numbers are just \( \Delta h \), which is the contraction of \( \vec{d}h \) with \( \Delta \vec{x} \): \( \Delta h = \sum_i (\partial h/\partial x^i) \Delta x^i \) or the **value** of \( \vec{d}h \) on \( \Delta \vec{x} \) (see Eq. (3.8)).
A topographical map illustrates the gradient one-form (local contours of constant elevation). The change of height along any trip (arrow) is the number of contours crossed by the arrow.

The value $\tilde{\omega}(\vec{V})$ is 2.5.

Therefore, a one-form is represented by a series of surfaces (Fig. 3.4), and its contraction with a vector $\vec{V}$ is the number of surfaces $\vec{V}$ crosses. The closer the surfaces, the larger $\tilde{\omega}$. Properly, just as a vector is straight, the one-form’s surfaces are straight and parallel. This is because we deal with one-forms at a point, not over an extended region: ‘tangent’ one-forms, in the same sense as tangent vectors.

These pictures show why we in general cannot call a gradient a vector. We would like to identify the vector gradient as that vector pointing ‘up’ the slope, i.e. in such a way that it crosses the greatest number of contours per unit length. The key phrase is ‘per unit length’. If there is a metric, a measure of distance in the space, then a vector can be associated with a gradient. But the metric must intervene here in order to produce a vector. Geometrically, on its own, the gradient is a one-form.

Let us be sure that Eq. (3.15) is a consistent definition. How do the components transform? For a one-form we must have

$$(\tilde{d}\phi)_{\bar{\alpha}} = \Lambda^\beta_{\bar{\alpha}} (\tilde{d}\phi)_\beta.$$  (3.16)

But we know how to transform partial derivatives:

$$\frac{\partial \phi}{\partial x^\alpha} = \frac{\partial \phi}{\partial x^\beta} \frac{\partial x^\beta}{\partial x^\alpha}.$$
which means
\[(\tilde{d}\phi)_{\alpha} = \frac{\partial x^\beta}{\partial x^{\alpha}}(\tilde{d}\phi)_\beta.\] (3.17)

Are Eqs. (3.16) and (3.17) consistent? The answer, of course, is yes. The reason: since
\[x^\beta = \Lambda^\beta_{\,\alpha} x^{\alpha},\]
and since \(\Lambda^\beta_{\,\alpha}\) are just constants, then
\[\frac{\partial x^\beta}{\partial \tilde{x}^{\alpha}} = \Lambda^\beta_{\,\alpha}.\] (3.18)

This identity is fundamental. Components of the gradient transform according to the inverse of the components of vectors. So the gradient is the ‘archetypal’ one-form.

**Notation for derivatives**

From now on we shall employ the usual subscripted notation to indicate derivatives:
\[\frac{\partial \phi}{\partial x} := \phi_x,\]
and, more generally,
\[\frac{\partial \phi}{\partial x^{\alpha}} := \phi_x^{\alpha}.\] (3.19)

Note that the index \(\alpha\) appears as a superscript in the denominator of the left-hand side of Eq. (3.19) and as a subscript on the right-hand side. As we have seen, this placement of indices is consistent with the transformation properties of the expression.

In particular, we have
\[x^{\alpha}_{\,\beta} \equiv \delta^{\alpha}_{\beta},\]
which we can compare with Eq. (3.12) to conclude that
\[\tilde{dx}^{\alpha} := \tilde{\omega}^{\alpha}.\] (3.20)

This is a useful result, that the basis one-form is just \(\tilde{dx}^{\alpha}\). We can use it to write, for any function \(f\),
\[\tilde{df} = \frac{\partial f}{\partial x^{\alpha}} \tilde{dx}^{\alpha}.\]

This looks very much like the physicist’s ‘sloppy-calculus’ way of writing differentials or infinitesimals. The notation \(\tilde{d}\) has been chosen partly to suggest this comparison, but this choice makes it doubly important for the student to avoid confusion on this point. The object \(\tilde{df}\) is a tensor, not a small increment in \(f\); it can have a small (‘infinitesimal’) value if it is contracted with a small vector.
Like the gradient, the concept of a normal vector – a vector orthogonal to a surface – is one which is more naturally replaced by that of a normal one-form. For a normal vector to be defined we need to have a scalar product: the normal vector must be orthogonal to all vectors tangent to the surface. This can be defined only by using the metric tensor. But a normal one-form can be defined without reference to the metric. A one-form is said to be normal to a surface if its value is zero on every vector tangent to the surface. If the surface is closed and divides spacetime into an ‘inside’ and ‘outside’, a normal is said to be an outward normal one-form if it is a normal one-form and its value on vectors which point outwards from the surface is positive. In Exer. 13, § 3.10, we prove that $\tilde{\delta}f$ is normal to surfaces of constant $f$.

### 3.4 The $(0^2)$ tensors

Tensors of type $(0^2)$ have two vector arguments. We have encountered the metric tensor already, but the simplest of this type is the product of two one-forms, formed according to the following rule: if $\tilde{p}$ and $\tilde{q}$ are one-forms, then $\tilde{p} \otimes \tilde{q}$ is the $(0^2)$ tensor which, when supplied with vectors $\tilde{A}$ and $\tilde{B}$ as arguments, produces the number $\tilde{p}(\tilde{A}) \tilde{q}(\tilde{B})$, i.e. just the product of the numbers produced by the $(1^0)$ tensors. The symbol $\otimes$ is called an ‘outer product sign’ and is a formal notation to show how the $(0^2)$ tensor is formed from the one-forms. Notice that $\otimes$ is not commutative: $\tilde{p} \otimes \tilde{q}$ and $\tilde{q} \otimes \tilde{p}$ are different tensors. The first gives the value $\tilde{p}(\tilde{A}) \tilde{q}(\tilde{B})$, the second the value $\tilde{q}(\tilde{A}) \tilde{p}(\tilde{B})$.

### Components

The most general $(0^2)$ tensor is not a simple outer product, but it can always be represented as a sum of such tensors. To see this we must first consider the components of an arbitrary $(0^2)$ tensor $f$:

$$f_{\alpha\beta} := f(\vec{e}_\alpha, \vec{e}_\beta). \quad (3.21)$$

Since each index can have four values, there are 16 components, and they can be thought of as being arrayed in a matrix. The value of $f$ on arbitrary vectors is

$$f(\tilde{A}, \tilde{B}) = f(A^\alpha \vec{e}_\alpha, B^\beta \vec{e}_\beta) = A^\alpha B^\beta f(\vec{e}_\alpha, \vec{e}_\beta) = A^\alpha B^\beta f_{\alpha\beta}. \quad (3.22)$$

(Again notice that two different dummy indices are used to keep the different summations distinct.) Can we form a basis for these tensors? That is, can we define a set of 16 $(0^2)$ tensors $\tilde{\omega}^{\alpha\beta}$ such that, analogous to Eq. (3.11),

**Normal one-forms**
3.4 The $\left(\frac{0}{2}\right)$ tensors

$$f = f_{\alpha\beta} \tilde{\omega}^{\alpha\beta}.$$  \hfill (3.23)

For this to be the case we would have to have

$$f_{\mu\nu} = f(\vec{e}_\mu, \vec{e}_\nu) = f_{\alpha\beta} \tilde{\omega}^{\alpha\beta}(\vec{e}_\mu, \vec{e}_\nu)$$

and this would imply, as before, that

$$\tilde{\omega}^{\alpha\beta}(\vec{e}_\mu, \vec{e}_\nu) = \delta^\alpha_\mu \delta^\beta_\nu.$$  \hfill (3.24)

But $\delta^\alpha_\mu$ is (by Eq. (3.12)) the value of $\tilde{\omega}^\alpha$ on $\vec{e}_\mu$, and analogously for $\delta^\beta_\nu$. Therefore, $\tilde{\omega}^{\alpha\beta}$ is a tensor the value of which is just the product of the values of two basis one-forms, and we therefore conclude

$$\tilde{\omega}^{\alpha\beta} = \tilde{\omega}^\alpha \otimes \tilde{\omega}^\beta.$$  \hfill (3.25)

So the tensors $\tilde{\omega}^\alpha \otimes \tilde{\omega}^\beta$ are a basis for all $\left(\frac{0}{2}\right)$ tensors, and we can write

$$f = f_{\alpha\beta} \tilde{\omega}^\alpha \otimes \tilde{\omega}^\beta.$$  \hfill (3.26)

This is one way in which a general $\left(\frac{0}{2}\right)$ tensor is a sum over simple outer-product tensors.

**Symmetries**

A $\left(\frac{0}{2}\right)$ tensor takes two arguments, and their order is important, as we have seen. The behavior of the value of a tensor under an interchange of its arguments is an important property of it. A tensor $f$ is called symmetric if

$$f(A, B) = f(B, A) \quad \forall A, B.$$  \hfill (3.27)

Setting $\tilde{A} = \vec{e}_\alpha$ and $\tilde{B} = \vec{e}_\beta$, this implies of its components that

$$f_{\alpha\beta} = f_{\beta\alpha}.$$  \hfill (3.28)

This is the same as the condition that the matrix array of the elements is symmetric. An arbitrary $\left(\frac{0}{2}\right)$ tensor $h$ can define a new symmetric $h_{(s)}$ by the rule

$$h_{(s)}(\tilde{A}, \tilde{B}) = \frac{1}{2} h(\tilde{A}, \tilde{B}) + \frac{1}{2} h(\tilde{B}, \tilde{A}).$$  \hfill (3.29)

Make sure you understand that $h_{(s)}$ satisfies Eq. (3.27) above. For the components this implies

$$h_{(s)}(\alpha\beta) = \frac{1}{2} (h_{\alpha\beta} + h_{\beta\alpha}).$$  \hfill (3.30)

This is such an important mathematical property that a special notation is used for it:

$$h_{(\alpha\beta)} := \frac{1}{2} (h_{\alpha\beta} + h_{\beta\alpha}).$$  \hfill (3.31)
Therefore, the numbers $h_{(\alpha\beta)}$ are the components of the symmetric tensor formed from $h$.

Similarly, a tensor $f$ is called antisymmetric if

$$f(\vec{A}, \vec{B}) = -f(\vec{B}, \vec{A}), \quad \forall \vec{A}, \vec{B}.$$  
\hspace{1cm} (3.32)

$$f_{\alpha\beta} = -f_{\beta\alpha}. \hspace{1cm} (3.33)$$

An antisymmetric ($0^2$) tensor can always be formed as

$$h_{(\alpha\beta)}(\vec{A}, \vec{B}) = \frac{1}{2} h(\vec{A}, \vec{B}) - \frac{1}{2} h(\vec{B}, \vec{A}),$$

$$h_{(\alpha\beta)}^{\alpha\beta} = \frac{1}{2} (h_{\alpha\beta} - h_{\beta\alpha}).$$

The notation here is to use square brackets on the indices:

$$h_{[\alpha\beta]} = \frac{1}{2} (h_{\alpha\beta} - h_{\beta\alpha}). \hspace{1cm} (3.34)$$

Notice that

$$h_{\alpha\beta} = \frac{1}{2} (h_{\alpha\beta} + h_{\beta\alpha}) + \frac{1}{2} (h_{\alpha\beta} - h_{\beta\alpha})$$

$$= h_{(\alpha\beta)} + h_{[\alpha\beta]}.$$  \hspace{1cm} (3.35)

So any ($0^2$) tensor can be split uniquely into its symmetric and antisymmetric parts.

The metric tensor $g$ is symmetric, as can be deduced from Eq. (2.26):

$$g(\vec{A}, \vec{B}) = g(\vec{B}, \vec{A}). \hspace{1cm} (3.36)$$

### 3.5 Metric as a mapping of vectors into one-forms

We now introduce what we shall later see is the fundamental role of the metric in differential geometry, to act as a mapping between vectors and one-forms. To see how this works, consider $g$ and a single vector $\vec{V}$. Since $g$ requires two vectorial arguments, the expression $g(\vec{V}, \ )$ still lacks one: when another one is supplied, it becomes a number. Therefore, $g(\vec{V}, \ )$ considered as a function of vectors (which are to fill in the empty ’slot’ in it) is a linear function of vectors producing real numbers: a one-form. We call it $\tilde{V}$:

$$g(\vec{V}, \ ) := \tilde{V}(\ ), \hspace{1cm} (3.37)$$

where blanks inside parentheses are a way of indicating that a vector argument is to be supplied. Then $\tilde{V}$ is the one-form that evaluates on a vector $\vec{A}$ to $\vec{V} \cdot \vec{A}$:

$$\tilde{V}(\vec{A}) := g(\vec{V}, \vec{A}) = \vec{V} \cdot \vec{A}. \hspace{1cm} (3.38)$$

Note that since $g$ is symmetric, we also can write

$$g(\ , \vec{V}) := \tilde{V}(\ ).$$
What are the components of $\tilde{V}$? They are

$$V_{\alpha} := V(\hat{e}_{\alpha}) = \tilde{V} \cdot \hat{e}_{\alpha} = \tilde{e}_{\alpha} \cdot \tilde{V} = \hat{e}_{\alpha} \cdot (V^\beta \tilde{e}_\beta) = (\tilde{e}_{\alpha} \cdot \tilde{e}_\beta) V^\beta$$

$$V_{\alpha} = \eta_{\alpha\beta} V^\beta. \quad (3.39)$$

It is important to notice here that we distinguish the components $V^\alpha$ of $\tilde{V}$ from the components $V^\beta$ of $\tilde{V}$ only by the position of the index: on a vector it is up; on a one-form, down.

Then, from Eq. (3.39), we have as a special case

$$V_0 = V^\beta \eta_{\beta 0} = V^0 \eta_{00} + V^1 \eta_{10} + \ldots$$
$$= V^0(-1) + 0 + 0 + 0$$
$$= -V^0, \quad (3.40)$$

$$V_1 = V^\beta \eta_{\beta 1} = V^0 \eta_{01} + V^1 \eta_{11} + \ldots$$
$$= +V^1, \quad (3.41)$$

etc. This may be summarized as:

$$\text{if } \tilde{V} \to (a, b, c, d),$$
then $\tilde{V} \to (-a, b, c, d). \quad (3.42)$

The components of $\vec{V}$ are obtained from those of $\tilde{V}$ by changing the sign of the time component. (Since this depended upon the components $\eta_{\alpha\beta}$, in situations we encounter later, where the metric has more complicated components, this rule of correspondence between $\vec{V}$ and $\tilde{V}$ will also be more complicated.)

**The inverse: going from $\tilde{A}$ to $\hat{A}$**

Does the metric also provide a way of finding a vector $\hat{A}$ that is related to a given one-form $\tilde{A}$? The answer is yes. Consider Eq. (3.39). It says that $\{V_{\alpha}\}$ is obtained by multiplying $\{V^\beta\}$ by a matrix $(\eta_{\alpha\beta})$. If this matrix has an inverse, then we could use it to obtain $\{V^\beta\}$ from $\{V_{\alpha}\}$. This inverse exists if and only if $(\eta_{\alpha\beta})$ has nonvanishing determinant. But since $(\eta_{\alpha\beta})$ is a diagonal matrix with entries $(-1, 1, 1, 1)$, its determinant is simply $-1$. An inverse does exist, and we call its components $\eta^{\alpha\beta}$. Then, given $\{A_{\beta}\}$ we can find $\{A^\alpha\}$:

$$A^\alpha := \eta^{\alpha\beta} A_{\beta}. \quad (3.43)$$
The use of the inverse guarantees that the two sets of components satisfy Eq. (3.39):

\[ A_\beta = \eta_{\beta\alpha} A^\alpha. \]

So the mapping provided by \( g \) between vectors and one-forms is one-to-one and invertible.

In particular, with \( \tilde{d}\phi \) we can associate a vector \( \tilde{\phi} \), which is the one usually associated with the gradient. We can see that this vector is orthogonal to surfaces of constant \( \phi \) as follows: its inner product with a vector in a surface of constant \( \phi \) is, by this mapping, identical with the value of the one-form \( \tilde{d}\phi \) on that vector. This, in turn, must be zero since \( \tilde{d}\phi(\tilde{V}) \) is the rate of change of \( \phi \) along \( \tilde{V} \), which in this case is zero since \( \tilde{V} \) is taken to be in a surface of constant \( \phi \).

It is important to know what \( \{\eta^{\alpha\beta}\} \) is. You can easily verify that

\[ \eta^{00} = -1, \quad \eta^{0i} = 0, \quad \eta^{ij} = \delta^{ij}, \tag{3.44} \]

so that \( (\eta^{\alpha\beta}) \) is identical to \( (\eta_{\alpha\beta}) \). Thus, to go from a one-form to a vector, simply change the sign of the time component.

Why distinguish one-forms from vectors?

In Euclidean space, in Cartesian coordinates the metric is just \( \{\delta_{ij}\} \), so the components of one-forms and vectors are the same. Therefore no distinction is ever made in elementary vector algebra. But in SR the components differ (by that one change in sign). Therefore, whereas the gradient has components

\[ \tilde{d}\phi \rightarrow \left( \frac{\partial \phi}{\partial t}, \frac{\partial \phi}{\partial x}, \ldots \right), \]

the associated vector normal to surfaces of constant \( \phi \) has components

\[ \tilde{\phi} \rightarrow \left( -\frac{\partial \phi}{\partial t}, \frac{\partial \phi}{\partial x}, \ldots \right). \tag{3.45} \]

Had we simply tried to define the ‘vector gradient’ of a function as the vector with these components, without first discussing one-forms, the reader would have been justified in being more than a little skeptical. The non-Euclidean metric of SR forces us to be aware of the basic distinction between one-forms and vectors: it can’t be swept under the rug.

As we remarked earlier, vectors and one-forms are dual to one another. Such dual spaces are important and are found elsewhere in mathematical physics. The simplest example is the space of column vectors in matrix algebra

\[ \begin{pmatrix} a \\ b \\ \vdots \end{pmatrix}, \]

whose dual space is the space of row vectors \((a \ b \ \cdots)\). Notice that the product

\[ (a \ b \ \cdots) \begin{pmatrix} p \\ q \\ \vdots \end{pmatrix} = ap + bq + \ldots \tag{3.46} \]
is a real number, so that a row vector can be considered to be a one-form on column vectors. The operation of finding an element of one space from one of the others is called the ‘adjoint’ and is 1–1 and invertible. A less trivial example arises in quantum mechanics. A wave-function (probability amplitude that is a solution to Schrödinger’s equation) is a complex scalar field $\psi(\vec{x})$, and is drawn from the Hilbert space of all such functions. This Hilbert space is a vector space, since its elements (functions) satisfy the axioms of a vector space. What is the dual space of one-forms? The crucial hint is that the inner product of any two functions $\phi(\vec{x})$ and $\psi(\vec{x})$ is not $\int \phi(\vec{x})\psi(\vec{x}) \, d^3x$, but, rather, is $\int \phi^*(\vec{x}) \psi(\vec{x}) \, d^3x$, the asterisk denoting complex conjugation. The function $\phi^*(\vec{x})$ acts like a one-form whose value on $\psi(\vec{x})$ is its integral with it (analogous to the sum in Eq. (3.8)). The operation of complex conjugation acts like our metric tensor, transforming a vector $\phi(\vec{x})$ (in the Hilbert space) into a one-form $\phi^*(\vec{x})$. The fact that $\phi^*(\vec{x})$ is also a function in the Hilbert space is, at this level, a distraction. (It is equivalent to saying that members of the set $(1, -1, 0, 0)$ can be components of either a vector or a one-form.) The important point is that in the integral $\int \phi^*(\vec{x}) \psi(\vec{x}) \, d^3x$, the function $\phi^*(\vec{x})$ is acting as a one-form, producing a (complex) number from the vector $\psi(\vec{x})$. This dualism is most clearly brought out in the Dirac ‘bra’ and ‘ket’ notation. Elements of the space of all states of the system are called $| \rangle$ (with identifying labels written inside), while the elements of the dual (adjoint with complex conjugate) space are called $\langle |$. Two ‘vectors’ $|1\rangle$ and $|2\rangle$ don’t form a number, but a vector and a dual vector $|1\rangle$ and $\langle 2|$ do: $\langle 2|1\rangle$ is the name of this number.

In such ways the concept of a dual vector space arises very frequently in advanced mathematical physics.

### Magnitudes and scalar products of one-forms

A one-form $\tilde{\rho}$ is defined to have the same magnitude as its associated vector $\tilde{p}$. Thus we write

$$\tilde{\rho}^2 = \tilde{p}^2 = \eta_{\alpha\beta} p^\alpha p^\beta. \quad (3.47)$$

This would seem to involve finding $\{p^\alpha\}$ from $\{p_\alpha\}$ before using Eq. (3.47), but we can easily get around this. We use Eq. (3.43) for both $p^\alpha$ and $p^\beta$ in Eq. (3.47):

$$\tilde{\rho}^2 = \eta_{\alpha\beta} (\eta^{\alpha\mu} p_\mu)(\eta^{\beta\nu} p_\nu). \quad (3.48)$$

(Notice that each independent summation uses a different dummy index.) But since $\eta_{\alpha\beta}$ and $\eta^{\beta\nu}$ are inverse matrices to each other, the sum on $\beta$ collapses:

$$\eta_{\alpha\beta} \eta^{\beta\nu} = \delta^\nu_\alpha. \quad (3.49)$$

Using this in Eq. (3.48) gives

$$\tilde{\rho}^2 = \eta^{\mu\nu} p_\mu p_\nu. \quad (3.50)$$
Thus, the inverse metric tensor can be used directly to find the magnitude of $\tilde{p}$ from its components. We can use Eq. (3.44) to write this explicitly as

$$\tilde{p}^2 = -(p_0)^2 + (p_1)^2 + (p_2)^2 + (p_3)^2.$$  

(3.51)

This is the same rule, in fact, as Eq. (2.24) for vectors. By its definition, this is frame invariant. One-forms are timelike, spacelike, or null, as their associated vectors are.

As with vectors, we can now define an inner product of one-forms. This is

$$\tilde{p} \cdot \tilde{q} := \frac{1}{2} \left[ (\tilde{p} + \tilde{q})^2 - \tilde{p}^2 - \tilde{q}^2 \right].$$  

(3.52)

Its expression in terms of components is, not surprisingly,

$$\tilde{p} \cdot \tilde{q} = -p_0 q_0 + p_1 q_1 + p_2 q_2 + p_3 q_3.$$  

(3.53)

### Normal vectors and unit normal one-forms

A vector is said to be normal to a surface if its associated one-form is a normal one-form. Eq. (3.38) shows that this definition is equivalent to the usual one that the vector be orthogonal to all tangent vectors. A normal vector or one-form is said to be a unit normal if its magnitude is $\pm 1$. (We can’t demand that it be $+1$, since timelike vectors will have negative magnitudes. All we can do is to multiply the vector or form by an overall factor to scale its magnitude to $\pm 1$.) Note that null normals cannot be unit normals.

A three-dimensional surface is said to be timelike, spacelike, or null according to which of these classes its normal falls into. (Exer. 12, § 3.10, proves that this definition is self-consistent.) In Exer. 21, § 3.10, we explore the following curious properties normal vectors have on account of our metric. An outward normal vector is the vector associated with an outward normal one-form, as defined earlier. This ensures that its scalar product with any vector which points outwards is positive. If the surface is spacelike, the outward normal vector points outwards. If the surface is timelike, however, the outward normal vector points inwards. And if the surface is null, the outward vector is tangent to the surface!

These peculiarities simply reinforce the view that it is more natural to regard the normal as a one-form, where the metric doesn’t enter the definition.

### 3.6 Finally: $(M_N)$ tensors

#### Vector as a function of one-forms

The dualism discussed above is in fact complete. Although we defined one-forms as functions of vectors, we can now see that vectors can perfectly well be regarded as linear functions that map one-forms into real numbers. Given a vector $\tilde{V}$, once we supply a one-form we get a real number:
3.6 Finally: \( (M^N) \) tensors

Finally:

\[ F_i \equiv M_{\alpha}^i = t \cdot V_{\alpha} = \langle \tilde{p}, \tilde{V} \rangle. \quad (3.54) \]

In this way we dethrone vectors from their special position as things ‘acted on’ by tensors, and regard them as tensors themselves, specifically as linear functions of single one-forms into real numbers. The last notation on Eq. (3.54) is new, and emphasizes the equal status of the two objects.

\( (M^0) \) tensors

Generalizing this, we define:

An \( (M^0) \) tensor is a linear function of \( M \) one-forms into the real numbers.

All our previous discussions of \( (0) \) tensors apply here. A simple \( (0) \) tensor is \( \tilde{V} \otimes \tilde{W} \), which, when supplied with two arguments \( \tilde{p} \) and \( \tilde{q} \), gives the number \( \tilde{V}(\tilde{p})\tilde{W}(\tilde{q}) := \tilde{p}(\tilde{V})\tilde{q}(\tilde{W}) = V^\alpha p_\alpha W^\beta q_\beta \). So \( \tilde{V} \otimes \tilde{W} \) has components \( V^\alpha W^\beta \). A basis for \( (0) \) tensors is \( \tilde{e}_\alpha \otimes \tilde{e}_\beta \). The components of an \( (M^0) \) tensor are its values when the basis one-form \( \tilde{\omega}^\alpha \) are its arguments. Notice that \( (M^0) \) tensors have components all of whose indices are superscripts.

\( (M^N) \) tensors

The final generalization is:

An \( (M^N) \) tensor is a linear function of \( M \) one-forms and \( N \) vectors into the real numbers.

For instance, if \( R \) is a \( (1^1) \) tensor, then it requires a one-form \( \tilde{p} \) and a vector \( \tilde{A} \) to give a number \( R(\tilde{p}; \tilde{A}) \). It has components \( R(\tilde{\omega}^\alpha; \tilde{e}_\beta) := R^\alpha_{\beta} \). In general, the components of a \( (M^N) \) tensor will have \( M \) indices up and \( N \) down. In a new frame,

\[ R'^{\bar{\alpha}}_{\bar{\beta}} = R(\Lambda^\alpha_{\mu} \tilde{\omega}^{\mu}; \Lambda^\nu_{\beta} \tilde{e}_\nu) = \Lambda^\alpha_{\mu} \Lambda^\nu_{\beta} R^\mu_{\nu}. \quad (3.55) \]

So the transformation of components is simple: each index transforms by bringing in a \( \Lambda \) whose indices are arranged in the only way permitted by the summation convention. Some old names that are still in current use are: upper indices are called ‘contravariant’ (because they transform contrary to basis vectors) and lower ones ‘covariant’. An \( (M^N) \) tensor is said to be ‘\( M \)-times contravariant and \( N \)-times covariant’.
At this point the student might worry that all of tensor algebra has become circular: one-forms were defined in terms of vectors, but now we have defined vectors in terms of one-forms. This ‘duality’ is at the heart of the theory, but is not circularity. It means we can do as physicists do, which is to identify the vectors with displacements \( \Delta \vec{x} \) and things like it (such as \( \vec{p} \) and \( \vec{v} \)) and then generate all \( \binom{N}{M} \) tensors by the rules of tensor algebra; these tensors inherit a physical meaning from the original meaning we gave vectors. But we could equally well have associated one-forms with some physical objects (gradients, for example) and recovered the whole algebra from that starting point. The power of the mathematics is that it doesn’t need (or want) to say what the original vectors or one-forms are. It simply gives rules for manipulating them. The association of, say, \( \vec{p} \) with a vector is at the interface between physics and mathematics: it is how we make a mathematical model of the physical world. A geometer does the same. He adds to the notion of these abstract tensor spaces the idea of what a vector in a curved space is. The modern geometer’s idea of a vector is something we shall learn about when we come to curved spaces. For now we will get some practice with tensors in physical situations, where we stick with our (admittedly imprecise) notion of vectors ‘like’ \( \Delta \vec{x} \).

### 3.7 Index ‘raising’ and ‘lowering’

In the same way that the metric maps a vector \( \vec{V} \) into a one-form \( \tilde{V} \), it maps an \( \binom{N}{M} \) tensor into an \( \binom{N-1}{M+1} \) tensor. Similarly, the inverse maps an \( \binom{N}{M} \) tensor into an \( \binom{N+1}{M-1} \) tensor. Normally, these are given the same name, and are distinguished only by the positions of their indices. Suppose \( T^{\alpha\beta\gamma} \) are the components of a \( \binom{3}{1} \) tensor. Then

\[
T^{\alpha\beta\gamma} := \eta_{\beta\mu}T^{\alpha\mu\gamma}\tag{3.56}
\]

are the components of a \( \binom{1}{2} \) tensor (obtained by mapping the second one-form argument of \( T^{\alpha\beta\gamma} \) into a vector), and

\[
T_{\alpha\beta\gamma} := \eta_{\alpha\mu}T^{\mu\beta\gamma}\tag{3.57}
\]

are the components of another (inequivalent) \( \binom{1}{2} \) tensor (mapping on the first index), while

\[
T^{\alpha\beta\gamma} := \eta^{\gamma\mu}T^{\alpha\mu\beta}\tag{3.58}
\]

are the components of a \( \binom{3}{0} \) tensor. These operations are, naturally enough, called index ‘raising’ and ‘lowering’. Whenever we speak of raising or lowering an index we mean this map generated by the metric. The rule in SR is simple: when raising or lowering a ‘0’ index, the sign of the component changes; when raising or lowering a ‘1’ or ‘2’ or ‘3’ index (in general, an ‘\( i \)’ index) the component is unchanged.
Mixed components of metric

The numbers \( \{ \eta_{\alpha\beta} \} \) are the components of the metric, and \( \{ \eta^{\alpha\beta} \} \) those of its inverse. Suppose we raise an index of \( \eta_{\alpha\beta} \) using the inverse. Then we get the ‘mixed’ components of the metric,

\[
\eta^{\alpha}_{\beta} \equiv \eta^{\alpha\mu} \eta_{\mu\beta}.
\]  

But on the right we have just the matrix product of two matrices that are the inverse of each other (readers who aren’t sure of this should verify the following equation by direct calculation), so it is the unit identity matrix. Since one index is up and one down, it is the Kronecker delta, written as

\[
\eta^{\alpha}_{\beta} \equiv \delta^{\alpha}_{\beta}.
\]  

By raising the other index we merely obtain an identity, \( \eta^{\alpha\beta} = \eta^{\alpha\beta} \). So we can regard \( \eta^{\alpha\beta} \) as the components of the \( (2\,0) \) tensor, which is mapped from the \( (0\,2) \) tensor \( g \) by \( g^{-1} \). So, for \( g \), its ‘contravariant’ components equal the elements of the matrix inverse of the matrix of its ‘covariant’ components. It is the only tensor for which this is true.

Metric and nonmetric vector algebras

It is of some interest to ask why the metric is the one that generates the correspondence between one-forms and vectors. Why not some other \( (0\,2) \) tensor that has an inverse? We’ll explore that idea in stages.

First, why a correspondence at all? Suppose we had a ‘nonmetric’ vector algebra, complete with all the dual spaces and \( (M\,N) \) tensors. Why make a correspondence between one-forms and vectors? The answer is that sometimes we do and sometimes we don’t. Without one, the inner product of two vectors is undefined, since numbers are produced only when one-forms act on vectors and vice-versa. In physics, scalar products are useful, so we need a metric. But there are some vector spaces in mathematical physics where metrics are not important. An example is phase space of classical and quantum mechanics.

Second, why the metric and not another tensor? If a metric were not defined but another symmetric tensor did the mapping, a mathematician would just call the other tensor the metric. That is, he would define it as the one generating a mapping. To a mathematician, the metric is an added bit of structure in the vector algebra. Different spaces in mathematics can have different metric structures. A Riemannian space is characterized by a metric that gives positive-definite magnitudes of vectors. One like ours, with indefinite sign, is called pseudo-Riemannian. We can even define a ‘metric’ that is antisymmetric: a two-dimensional space called spinor space has such a metric, and it turns out to be of fundamental importance in physics. But its structure is outside the scope of this book. The point here is that we don’t have SR if we just discuss vectors and tensors. We get SR when
we say that we have a metric with components $\eta_{\alpha\beta}$. If we assigned other components, we might get other spaces, in particular the curved spacetime of GR.

### 3.8 Differentiation of Tensors

A function $f$ is a $(0,0)$ tensor, and its gradient $\tilde{d}f$ is a $(0,1)$ tensor. Differentiation of a function produces a tensor of one higher (covariant) rank. We shall now see that this applies as well to differentiation of tensors of any rank.

Consider a $(1,1)$ tensor $T$ whose components $\{T^\alpha_\beta\}$ are functions of position. We can write $T$ as

$$T = T^\alpha_\beta \tilde{\omega}^\beta \otimes \tilde{e}_\alpha. \quad (3.61)$$

Suppose, as we did for functions, that we move along a world line with parameter $\tau$, proper time. The rate of change of $T$,

$$\frac{dT}{d\tau} = \lim_{\Delta\tau \to 0} \frac{T(\tau + \Delta\tau) - T(\tau)}{\Delta \tau}, \quad (3.62)$$

is not hard to calculate. Since the basis one-forms and vectors are the same everywhere (i.e. $\tilde{\omega}^\alpha(\tau + \Delta\tau) = \tilde{\omega}^\alpha(\tau)$), it follows that

$$\frac{dT}{d\tau} = \left(\frac{dT^\alpha_\beta}{d\tau}\right) \tilde{\omega}^\beta \otimes \tilde{e}_\alpha, \quad (3.63)$$

where $dT^\alpha_\beta/d\tau$ is the ordinary derivative of the function $T^\alpha_\beta$ along the world line:

$$\frac{dT^\alpha_\beta}{d\tau} = T^\alpha_\beta, \quad (3.64)$$

Now, the object $dT/d\tau$ is a $(1,1)$ tensor, since in Eq. (3.62) it is defined to be just the difference between two such tensors. From Eqs. (3.63) and (3.64) we have, for any vector $\tilde{U}$,

$$\frac{dT}{d\tau} = (T^\alpha_\beta, \tilde{\omega}^\beta \otimes \tilde{e}_\alpha) \tilde{U}^\gamma, \quad (3.65)$$

from which we can deduce that

$$\nabla T := (T^\alpha_\beta, \tilde{\omega}^\beta \otimes \tilde{\omega}^\gamma \otimes \tilde{e}_\alpha) \quad (3.66)$$

is a $(1,2)$ tensor. This tensor is called the gradient of $T$.

We use the notation $\nabla T$ rather than $\tilde{d}T$ because the latter notation is usually reserved by mathematicians for something else. We also have a convenient notation for Eq. (3.65):

$$\frac{dT}{d\tau} = \nabla_{\tilde{U}} T, \quad (3.67)$$

$$\nabla_{\tilde{U}} T \rightarrow \{T^\alpha_\beta, \tilde{U}^\gamma\}. \quad (3.68)$$

This derivation made use of the fact that the basis vectors (and therefore the basis one-forms) were constant everywhere. We will find that we can’t assume this in the curved spacetime of GR, and taking this into account will be our entry point into the theory!
Our approach to tensor analysis stresses the geometrical nature of tensors rather than the transformation properties of their components. Students who wish amplification of some of the points here can consult the early chapters of Misner et al. (1973) or Schutz (1980b). See also Bishop and Goldberg (1981).

Most introductions to tensors for physicists outside relativity confine themselves to ‘Cartesian’ tensors, i.e. to tensor components in three-dimensional Cartesian coordinates. See, for example, Bourne and Kendall (1991) or the chapter in Mathews and Walker (1965).

A very complete reference for tensor analysis in the older style based upon coordinate transformations is Schouten (1990). See also Yano (1955). Books which develop that point of view for tensors in relativity include Adler et al. (1975), Landau and Lifshitz (1962), and Stephani (2004).

### 3.10 Exercises

1. (a) Given an arbitrary set of numbers \( \{ M_{\alpha\beta}; \alpha = 0, \ldots, 3; \beta = 0, \ldots, 3 \} \) and two arbitrary sets of vector components \( \{ A^\mu, \mu = 0, \ldots, 3 \} \) and \( \{ B^\nu, \nu = 0, \ldots, 3 \} \), show that the two expressions

\[
M_{\alpha\beta} A^\alpha B^\beta := \sum_{\alpha=0}^{3} \sum_{\beta=0}^{3} M_{\alpha\beta} A^\alpha B^\beta
\]

and

\[
\sum_{\alpha=0}^{3} M_{\alpha\alpha} A^\alpha B^\alpha
\]

are not equivalent.

(b) Show that

\[
A^\alpha B^\beta \eta_{\alpha\beta} = -A^0 B^0 + A^1 B^1 + A^2 B^2 + A^3 B^3.
\]

2. Prove that the set of all one-forms is a vector space.

3. (a) Prove, by writing out all the terms, the validity of the following

\[
\tilde{p}(A^\alpha e_\alpha) = A^\alpha \tilde{p}(\tilde{e}_\alpha).
\]

(b) Let the components of \( \tilde{p} \) be \((-1, 1, 2, 0)\), those of \( \tilde{A} \) be \((2, 1, 0, -1)\) and those of \( \tilde{B} \) be \((0, 2, 0, 0)\). Find (i) \( \tilde{p}(\tilde{A}) \); (ii) \( \tilde{p}(\tilde{B}) \); (iii) \( \tilde{p}(\tilde{A} - 3\tilde{B}) \); (iv) \( \tilde{p}(\tilde{A}) - 3\tilde{p}(\tilde{B}) \).

4. Given the following vectors in \( \mathcal{O} \):

\[
\tilde{A} \rightarrow (2, 1, 1, 0), \tilde{B} \rightarrow (1, 2, 0, 0), \tilde{C} \rightarrow (0, 0, 1, 1), \tilde{D} \rightarrow (-3, 2, 0, 0),
\]

(a) show that they are linearly independent;

(b) find the components of \( \tilde{p} \) if
\[ \tilde{p}(\vec{A}) = 1, \tilde{p}(\vec{B}) = -1, \tilde{p}(\vec{C}) = -1, \tilde{p}(\vec{D}) = 0; \]

(c) find the value of \( \tilde{p}(\vec{E}) \) for \( \vec{E} \to (1, 1, 0, 0) \);

(d) determine whether the one-forms \( \tilde{p}, \tilde{q}, \tilde{r}, \) and \( \tilde{s} \) are linearly independent if \( \tilde{q}(\vec{A}) = \tilde{q}(\vec{B}) = 0, \tilde{q}(\vec{C}) = 1, \tilde{q}(\vec{D}) = -1, \tilde{r}(\vec{A}) = 2, \tilde{r}(\vec{B}) = \tilde{r}(\vec{C}) = \tilde{r}(\vec{D}) = 0, \tilde{s}(\vec{A}) = -1, \tilde{s}(\vec{B}) = -1, \tilde{s}(\vec{C}) = \tilde{s}(\vec{D}) = 0. \]

5 Justify each step leading from Eqs. (3.10a) to (3.10d).

6 Consider the basis \( \{\vec{e}_\alpha\} \) of a frame \( \mathcal{O} \) and the basis \( (\tilde{\lambda}_0, \tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3) \) for the space of one-forms, where we have

\[
\begin{align*}
\tilde{\lambda}_0 & \to (1, 1, 0, 0), \\
\tilde{\lambda}_1 & \to (1, -1, 0, 0), \\
\tilde{\lambda}_2 & \to (0, 0, 1, -1), \\
\tilde{\lambda}_3 & \to (0, 0, 1, 1).
\end{align*}
\]

Note that \( \{\tilde{\lambda}^\beta\} \) is not the basis dual to \( \{\vec{e}_\alpha\} \).

(a) Show that \( \tilde{p} \neq \tilde{p}(\vec{e}_\alpha)\tilde{\lambda}^\alpha \) for arbitrary \( \tilde{p} \).

(b) Let \( \tilde{p} \to \mathcal{O} (1, 1, 1, 1) \). Find numbers \( l_\alpha \) such that

\[ \tilde{p} = l_\alpha \tilde{\lambda}^\alpha. \]

These are the components of \( \tilde{p} \) on \( \{\tilde{\lambda}^\alpha\} \), which is to say that they are the values of \( \tilde{p} \) on the elements of the vector basis dual to \( \{\tilde{\lambda}^\alpha\} \).

7 Prove Eq. (3.13).

8 Draw the basis one-forms \( \tilde{dt} \) and \( \tilde{dx} \) of a frame \( \mathcal{O} \).

9 Fig. 3.5 shows curves of equal temperature \( T \) (isotherms) of a metal plate. At the points \( \mathcal{P} \) and \( \mathcal{Q} \) as shown, estimate the components of the gradient \( \tilde{d}T \). (Hint: the components
are the contractions with the basis vectors, which can be estimated by counting the number of isotherms crossed by the vectors.)

10. (a) Given a frame $\mathcal{O}$ whose coordinates are $\{x^\alpha\}$, show that

$$\frac{\partial x^\alpha}{\partial x^\beta} = \delta^\alpha_\beta.$$ 

(b) For any two frames, we have, Eq. (3.18):

$$\frac{\partial x^\beta}{\partial \tilde{x}_\alpha} = \Lambda^\beta_{\tilde{\alpha}}.$$ 

Show that (a) and the chain rule imply

$$\Lambda^\beta_{\tilde{\alpha}} \Lambda_{\tilde{\mu}} = \delta^\beta_\mu.$$ 

This is the inverse property again.

11. Use the notation $\partial \phi / \partial x^\alpha = \phi_{,\alpha}$ to re-write Eqs. (3.14), (3.15), and (3.18).

12. Let $S$ be the two-dimensional plane $x = 0$ in three-dimensional Euclidean space. Let $\tilde{n} \neq 0$ be a normal one-form to $S$.

(a) Show that if $\tilde{V}$ is a vector which is not tangent to $S$, then $\tilde{n}(\tilde{V}) \neq 0$.

(b) Show that if $\tilde{n}(\tilde{V}) > 0$, then $\tilde{n}(\tilde{W}) > 0$ for any $\tilde{W}$, which points toward the same side of $S$ as $\tilde{V}$ does (i.e. any $\tilde{W}$ whose $x$ components has the same sign as $V^x$).

(c) Show that any normal to $S$ is a multiple of $\tilde{n}$.

(d) Generalize these statements to an arbitrary three-dimensional surface in four-dimensional spacetime.

13. Prove, by geometric or algebraic arguments, that $\tilde{df}$ is normal to surfaces of constant $f$.

14. Let $\tilde{p} \to_\mathcal{O} (1, 1, 0, 0)$ and $\tilde{q} \to_\mathcal{O} (-1, 0, 1, 0)$ be two one-forms. Prove, by trying two vectors $\tilde{A}$ and $\tilde{B}$ as arguments, that $\tilde{p} \otimes \tilde{q} \neq \tilde{q} \otimes \tilde{p}$. Then find the components of $\tilde{p} \otimes \tilde{q}$.

15. Supply the reasoning leading from Eq. (3.23) to Eq. (3.24).

16. (a) Prove that $\mathbf{h}_{(s)}$ defined by

$$\mathbf{h}_{(s)}(\tilde{A}, \tilde{B}) = \frac{1}{2} \mathbf{h}(\tilde{A}, \tilde{B}) + \frac{1}{2} \mathbf{h}(\tilde{B}, \tilde{A})$$

is an symmetric tensor.

(b) Prove that $\mathbf{h}_{(A)}$ defined by

$$\mathbf{h}_{(A)}(\tilde{A}, \tilde{B}) = \frac{1}{2} \mathbf{h}(\tilde{A}, \tilde{B}) - \frac{1}{2} \mathbf{h}(\tilde{B}, \tilde{A})$$

is an antisymmetric tensor.

(c) Find the components of the symmetric and antisymmetric parts of $\tilde{p} \otimes \tilde{q}$ defined in Exer. 14.

(d) Prove that if $\mathbf{h}$ is an antisymmetric $\binom{0}{2}$ tensor,

$$\mathbf{h}(\tilde{A}, \tilde{A}) = 0$$

for any vector $\tilde{A}$.

(e) Find the number of independent components $\mathbf{h}_{(s)}$ and $\mathbf{h}_{(A)}$ have.

17. (a) Suppose that $\mathbf{h}$ is a $\binom{0}{2}$ tensor with the property that, for any two vectors $\tilde{A}$ and $\tilde{B}$ (where $\tilde{B} \neq 0$)

$$\mathbf{h}(\ , \tilde{A}) = \alpha \mathbf{h}(\ , \tilde{B}),$$

for any vector $\tilde{A}$. 


where \( \alpha \) is a number which may depend on \( \vec{A} \) and \( \vec{B} \). Show that there exist one-forms \( \tilde{p} \) and \( \tilde{q} \) such that
\[
h = \tilde{p} \otimes \tilde{q}.
\]

(b) Suppose \( T \) is a \((1,1)\) tensor, \( \tilde{\omega} \) a one-form, \( \vec{v} \) a vector, and \( T(\tilde{\omega}; \vec{v}) \) the value of \( T \) on \( \tilde{\omega} \) and \( \vec{v} \). Prove that \( T(\tilde{\omega}; \vec{v}) \) is a vector and \( T(\tilde{\omega}; \vec{v}) \) is a one-form, i.e. that a \((1,1)\) tensor provides a map of vectors to vectors and one-forms to one-forms.

18 (a) Find the one-forms mapped by the metric tensor from the vectors
\[
\vec{A} \rightarrow O(1,0,-1,0), \quad \vec{B} \rightarrow O(0,1,1,0), \quad \vec{C} \rightarrow O(-1,0,-1,0), \quad \vec{D} \rightarrow O(0,0,1,1).
\]

(b) Find the vectors mapped by the inverse of the metric tensor from the one-form \( \tilde{p} \rightarrow O(3,0,-1,-1), \quad \tilde{q} \rightarrow O(1,-1,1,1), \quad \tilde{r} \rightarrow O(0,-5,-1,0), \quad \tilde{s} \rightarrow O(-2,1,0,0).

19 (a) Prove that the matrix \( \{\eta^{\alpha\beta}\} \) is inverse to \( \{\eta_{\alpha\beta}\} \) by performing the matrix multiplication.

(b) Derive Eq. (3.53).

20 In Euclidean three-space in Cartesian coordinates, we don’t normally distinguish between vectors and one-forms, because their components transform identically. Prove this in two steps.

(a) Show that
\[
A^{\bar{\alpha}} = \Lambda^{\bar{\alpha}}_{\beta} A^{\beta}
\]
and
\[
P_{\bar{\alpha}} = \Lambda^{\alpha}_{\bar{\beta}} P_{\alpha}
\]
are the same transformation if the matrix \( \{\Lambda^{\bar{\alpha}}_{\beta}\} \) equals the transpose of its inverse. Such a matrix is said to be orthogonal.

(b) The metric of such a space has components \( \{\delta_{ij}, i,j = 1, \ldots, 3\} \). Prove that a transformation from one Cartesian coordinate system to another must obey
\[
\delta_{ij} = \Lambda^{k}_{i} \Lambda^{l}_{j} \delta_{kl}
\]
and that this implies \( \{\Lambda^{k}_{j}\} \) is an orthogonal matrix. See Exer. 32 for the analog of this in SR.

21 (a) Let a region of the \( t-x \) plane be bounded by the lines \( t = 0, \ t = 1, \ x = 0, \ x = 1 \).
Within the \( t-x \) plane, find the unit outward normal one-forms and their associated vectors for each of the boundary lines.

(b) Let another region be bounded by the straight lines joining the events whose coordinates are \((1,0), (1,1), \) and \((2,1)\). Find an outward normal for the null boundary and find its associated vector.

22 Suppose that instead of defining vectors first, we had begun by defining one-forms, aided by pictures like Fig. 3.4. Then we could have introduced vectors as linear real-valued functions of one-forms, and defined vector algebra by the analogs of Eqs. (3.6a) and (3.6b) (i.e. by exchanging arrows for tildes). Prove that, so defined, vectors form a vector space. This is another example of the duality between vectors and one-forms.
23 (a) Prove that the set of all \( \binom{M}{N} \) tensors for fixed \( M, N \) forms a vector space. (You must define addition of such tensors and their multiplication by numbers.)
(b) Prove that a basis for this space is the set
\[
\{ e_\alpha \otimes e_\beta \otimes \ldots \otimes e_\gamma \otimes \tilde{\omega}^\mu \otimes \tilde{\omega}^\nu \otimes \ldots \otimes \tilde{\omega}^\lambda \}.
\]
(You will have to define the outer product of more than two one-forms.)

24 (a) Given the components of a \( \binom{2}{0} \) tensor \( M^{\alpha\beta} \) as the matrix
\[
\begin{pmatrix}
0 & 1 & 0 & 0 \\
1 & -1 & 0 & 2 \\
2 & 0 & 0 & 1 \\
1 & 0 & -2 & 0
\end{pmatrix},
\]
find:
(i) the components of the symmetric tensor \( M^{(\alpha\beta)} \) and the antisymmetric tensor \( M^{[\alpha\beta]} \);
(ii) the components of \( M^{\alpha\beta} \);
(iii) the components of \( M_{\alpha\beta} \);
(iv) the components of \( M_{[\alpha\beta]} \).
(b) For the \( \binom{1}{1} \) tensor whose components are \( M^{\alpha\beta} \), does it make sense to speak of its symmetric and antisymmetric parts? If so, define them. If not, say why.
(c) Raise an index of the metric tensor to prove
\[
\eta^{\alpha}_{\beta} = \delta^{\alpha}_{\beta}.
\]

25 Show that if \( A \) is a \( \binom{2}{0} \) tensor and \( B \) a \( \binom{0}{2} \) tensor, then
\[
A^{\alpha\beta}B_{\alpha\beta}
\]
is frame invariant, i.e. a scalar.

26 Suppose \( A \) is an antisymmetric \( \binom{2}{0} \) tensor, \( B \) a symmetric \( \binom{0}{2} \) tensor, \( C \) an arbitrary \( \binom{0}{2} \) tensor, and \( D \) an arbitrary \( \binom{2}{0} \) tensor. Prove:
(a) \( A^{\alpha\beta}B_{\alpha\beta} = 0 \);
(b) \( A^{\alpha\beta}C_{\alpha\beta} = A^{\alpha\beta}C^{[\alpha\beta]} \);
(c) \( B_{\alpha\beta}D^{\alpha\beta} = B_{\alpha\beta}D^{[\alpha\beta]} \).

27 (a) Suppose \( A \) is an antisymmetric \( \binom{2}{0} \) tensor. Show that \( \{A_{\alpha\beta}\} \), obtained by lowering indices by using the metric tensor, are components of an antisymmetric \( \binom{0}{2} \) tensor.
(b) Suppose \( V^\alpha = W^\alpha \). Prove that \( V_\alpha = W_\alpha \).

28 Deduce Eq. (3.66) from Eq. (3.65).

29 Prove that tensor differentiation obeys the Leibniz (product) rule:
\[
\nabla (A \otimes B) = (\nabla A) \otimes B + A \otimes \nabla B.
\]

30 In some frame \( O \), the vector fields \( \vec{U} \) and \( \vec{D} \) have the components:
\[
\vec{U} \rightarrow (1 + t^2, t^2, \sqrt{2} t, 0),
\vec{D} \rightarrow (x, 5tx, \sqrt{2} t, 0),
\]
and the scalar $\rho$ has the value

$$\rho = x^2 + t^2 - y^2.$$ 

(a) Find $\vec{U} \cdot \vec{U}, \vec{U} \cdot \vec{D}, \vec{D} \cdot \vec{D}$. Is $\vec{U}$ suitable as a four-velocity field? Is $\vec{D}$?

(b) Find the spatial velocity $v$ of a particle whose four-velocity is $\vec{U}$, for arbitrary $t$. What happens to it in the limits $t \to 0, t \to \infty$?

(c) Find $U_\alpha$ for all $\alpha$.

(d) Find $U^\alpha, \beta$ for all $\alpha, \beta$.

(e) Show that $U_\alpha U^\alpha, \beta = 0$ for all $\beta$. Show that $U^\alpha U_\alpha, \beta = 0$ for all $\beta$.

(f) Find $D^\beta, \beta$.

(g) Find $(U^\alpha D^\beta), \beta$ for all $\alpha$.

(h) Find $U_\alpha (U^\alpha D^\beta), \beta$ and compare with (f) above. Why are the two answers similar?

(i) Find $\rho, \alpha$ for all $\alpha$. Find $\rho^\alpha$ for all $\alpha$. (Recall that $\rho^\alpha := \eta^{\alpha\beta} \rho_\beta$.) What are the numbers $\{\rho^\alpha\}$ the components of?

(j) Find $\nabla_\alpha \vec{U}, \nabla_\alpha \vec{D}, \nabla_\beta \rho, \nabla_\beta \vec{U}$.

Consider a timelike unit four-vector $\vec{U}$, and the tensor $P$ whose components are given by

$$P_{\mu\nu} = \eta_{\mu\nu} + U_\mu U_\nu.$$ 

(a) Show that $P$ is a projection operator that projects an arbitrary vector $\vec{V}$ into one orthogonal to $\vec{U}$. That is, show that the vector $V^\alpha_\perp$ whose components are

$$V^\alpha_\perp = P^\alpha, \beta V^\beta = (\eta^\alpha, \beta + U^\alpha U_\beta) V^\beta$$

is

(i) orthogonal to $\vec{U}$,

and

(ii) unaffected by $P$:

$$V^\alpha_\perp \perp = P^\alpha, \beta V^\beta = V^\alpha_\perp.$$ 

(b) Show that for an arbitrary non-null vector $\vec{q}$, the tensor that projects orthogonally to it has components

$$\eta_{\mu\nu} - q_\mu q_\nu / (q^\alpha q_\alpha).$$

How does this fail for null vectors? How does this relate to the definition of $P$?

(c) Show that $P$ defined above is the metric tensor for vectors perpendicular to $\vec{U}$:

$$P(\vec{V}_\perp, \vec{W}_\perp) = g(\vec{V}_\perp, \vec{W}_\perp)$$

$$= \vec{V}_\perp \cdot \vec{W}_\perp.$$ 

From the definition $f_{\mu\nu} = \tilde{f}(\tilde{e}_\alpha, \tilde{e}_\beta)$ for the components of a $^{(0)}_2$ tensor, prove that the transformation law is

$$f_{\tilde{\alpha}\tilde{\beta}} = \Lambda^{\mu, \alpha} \Lambda^{\nu, \beta} f_{\mu\nu}$$

and that the matrix version of this is

$$(\tilde{f}) = (\Lambda)^T (f)(\Lambda),$$

where $(\Lambda)$ is the matrix with components $\Lambda^{\mu, \alpha}$. 

(b) Since our definition of a Lorentz frame led us to deduce that the metric tensor has components $\eta_{\alpha\beta}$, this must be true in all Lorentz frames. We are thus led to a more general definition of a Lorentz transformation as one whose matrix $\Lambda^\mu_\alpha$ satisfies

$$\eta_{\bar{\alpha}\bar{\beta}} = \Lambda^\mu_\alpha \Lambda^\nu_\beta \eta_{\mu\nu}. \quad (3.71)$$

Prove that the matrix for a boost of velocity $v \bar{e}_x$ satisfies this, so that this new definition includes our older one.

(c) Suppose $(\Lambda)$ and $(L)$ are two matrices which satisfy Eq. (3.71), i.e. $(\eta) = (\Lambda)^T (\eta) (\Lambda)$ and similarly for $(L)$. Prove that $(\Lambda)(L)$ is also the matrix of a Lorentz transformation.

33 The result of Exer. 32c establishes that Lorentz transformations form a group, represented by multiplication of their matrices. This is called the Lorentz group, denoted by $L(4)$ or $0(1,3)$.

(a) Find the matrices of the identity element of the Lorentz group and of the element inverse to that whose matrix is implicit in Eq. (1.12).

(b) Prove that the determinant of any matrix representing a Lorentz transformation is $\pm 1$.

(c) Prove that those elements whose matrices have determinant $+1$ form a subgroup, while those with $-1$ do not.

(d) The three-dimensional orthogonal group $O(3)$ is the analogous group for the metric of three-dimensional Euclidean space. In Exer. 20b, we saw that it was represented by the orthogonal matrices. Show that the orthogonal matrices do form a group, and then show that $0(3)$ is (isomorphic to) a subgroup of $L(4)$.

34 Consider the coordinates $u = t - x, v = t + x$ in Minkowski space.

(a) Define $\bar{e}_u$ to be the vector connecting the events with coordinates $\{u = 1, v = 0, y = 0, z = 0\}$ and $\{u = 0, v = 0, y = 0, z = 0\}$, and analogously for $\bar{e}_v$. Show that $\bar{e}_u = (\bar{e}_t - \bar{e}_x)/2, \bar{e}_v = (\bar{e}_t + \bar{e}_x)/2$, and draw $\bar{e}_u$ and $\bar{e}_v$ in a spacetime diagram of the $t - x$ plane.

(b) Show that $\{\bar{e}_u, \bar{e}_v, \bar{e}_y, \bar{e}_z\}$ are a basis for vectors in Minkowski space.

(c) Find the components of the metric tensor on this basis.

(d) Show that $\bar{e}_u$ and $\bar{e}_v$ are null and not orthogonal. (They are called a null basis for the $t - x$ plane.)

(e) Compute the four one-forms $\tilde{du}, \tilde{dv}, g(\bar{e}_u), g(\bar{e}_v)$ in terms of $\tilde{dt}$ and $\tilde{dx}$. 