1 Tensors

In order to describe more than particle motion in spacetime, we need to define objects more complicated than vectors, analogous to matrices and their generalizations. In order to construct meaningful physical quantities – invariants – we need to keep track of how these objects transform under Lorentz transformations. This leads us to define *Lorentz tensors* as those objects which transform *linearly and homogeneously* under Lorentz transformations. It is possible to define tensors for other groups of transformations as well. We will write our definition in a way that actually applies to the group of general coordinate transformations (or, in another guise, diffeomorphisms).

Our definition parallels our definition of a 4-vector as any set of four quantities transforming as

$$A'^{\alpha} = \sum_{\beta=0}^{3} M^{\alpha}{}_{\beta} A^{\beta}$$

where

$$x'^{\alpha} = \sum_{\beta=0}^{3} M^{\alpha}{}_{\beta} x^{\beta}$$

is a Lorentz transformation. Notice that the Jacobian matrix of partial derivatives,

$$\begin{aligned}
\frac{\partial x'^{\alpha}}{\partial x^{\mu}} &= \frac{\partial}{\partial x^{\mu}} \sum_{\beta=0}^{3} M^{\alpha}{}_{\beta} x^{\beta} \\
&= \sum_{\beta=0}^{3} M^{\alpha}{}_{\beta} \frac{\partial}{\partial x^{\mu}} x^{\beta} \\
&= \sum_{\beta=0}^{3} M^{\alpha}{}_{\beta} \delta^{\beta}_{\mu} \\
&= M^{\alpha}{}_{\mu}
\end{aligned}$$

To generalize to arbitrary coordinate transformations instead of just the constant Lorentz transformations, we simply replace $M^{\alpha}{}_{\mu}$ by the matrix $\frac{\partial x'^{\alpha}}{\partial x^{\mu}}$, where $x'^{\alpha} = x'^{\alpha}\left(x^{\beta}\right)$ may be any coordinate transformation. The transformation of a vector then becomes

$$A'^{\alpha} = \sum_{\beta=0}^{3} \frac{\partial x'^{\alpha}}{\partial x^{\beta}} A^{\beta}$$

Now suppose we have a vector equation of the form

$$V^{\alpha} = \sum_{\beta=0}^{3} T^{\alpha}{}_{\beta} S^{\beta}$$

for two vectors and a matrix. We say such an equation is *covariant* if it has the same form in any other coordinates. In order for this to be the case we require

$$\begin{aligned}
V'^{\alpha} &= \sum_{\mu=0}^{3} T'^{\alpha}{}_{\mu} S'^{\mu} \\
\sum_{\beta=0}^{3} \frac{\partial x'^{\alpha}}{\partial x^{\beta}} V^{\beta} &= \sum_{\mu=0}^{3} T'^{\alpha}{}_{\mu} \sum_{\nu=0}^{3} \frac{\partial x'^{\mu}}{\partial x^{\nu}} S^{\nu}
\end{aligned}$$

$$\begin{aligned}
\sum_{\beta=0}^{3}\frac{\partial x'^{\alpha}}{\partial x^{\beta}}\left(\sum_{\rho=0}^{3}T^{\beta}{}_{\rho}S^{\rho}\right) &= \sum_{\mu=0}^{3}\sum_{\nu=0}^{3}T'^{\alpha}{}_{\mu}\frac{\partial x'^{\mu}}{\partial x^{\nu}}S^{\nu} \\
\sum_{\beta=0}^{3}\frac{\partial x'^{\alpha}}{\partial x^{\beta}}\left(\sum_{\nu=0}^{3}T^{\beta}{}_{\nu}S^{\nu}\right) &= \sum_{\mu=0}^{3}\sum_{\nu=0}^{3}T'^{\alpha}{}_{\mu}\frac{\partial x'^{\mu}}{\partial x^{\nu}}S^{\nu} \\
\sum_{\nu=0}^{3}\left(\sum_{\beta=0}^{3}\frac{\partial x'^{\alpha}}{\partial x^{\beta}}T^{\beta}{}_{\nu}-\sum_{\mu=0}^{3}T'^{\alpha}{}_{\mu}\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right)S^{\nu} &= 0
\end{aligned}$$

If this equation holds for all vectors S^{ν}, then the matrix in parentheses must vanish,

$$\sum_{\beta=0}^{3}\frac{\partial x'^{\alpha}}{\partial x^{\beta}}T^{\beta}{}_{\nu}-\sum_{\mu=0}^{3}T'^{\alpha}{}_{\mu}\frac{\partial x'^{\mu}}{\partial x^{\nu}}=0$$

Using the inverse transformation,

$$\begin{aligned}
\sum_{\mu=0}^{3}\frac{\partial x'^{\alpha}}{\partial x^{\mu}}\frac{\partial x^{\mu}}{\partial x'^{\beta}} &= \frac{\partial x'^{\alpha}}{\partial x'^{\beta}} \\
&= \delta^{\alpha}_{\beta}
\end{aligned}$$

we have

$$\begin{aligned}
0 &= \left(\sum_{\beta=0}^{3}\frac{\partial x'^{\alpha}}{\partial x^{\beta}}T^{\beta}{}_{\nu}-\sum_{\mu=0}^{3}T'^{\alpha}{}_{\mu}\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right)\frac{\partial x^{\nu}}{\partial x^{\sigma}} \\
&= \sum_{\nu=0}^{3}\sum_{\beta=0}^{3}\frac{\partial x'^{\alpha}}{\partial x^{\beta}}T^{\beta}{}_{\nu}\frac{\partial x^{\nu}}{\partial x^{\sigma}}-\sum_{\mu=0}^{3}T'^{\alpha}{}_{\mu}\sum_{\nu=0}^{3}\frac{\partial x'^{\mu}}{\partial x^{\nu}}\frac{\partial x^{\nu}}{\partial x^{\sigma}} \\
&= \sum_{\nu=0}^{3}\sum_{\beta=0}^{3}\frac{\partial x'^{\alpha}}{\partial x^{\beta}}T^{\beta}{}_{\nu}\frac{\partial x^{\nu}}{\partial x^{\sigma}}-\sum_{\mu=0}^{3}T'^{\alpha}{}_{\mu}\delta^{\mu}_{\sigma} \\
&= \sum_{\nu=0}^{3}\sum_{\beta=0}^{3}\frac{\partial x'^{\alpha}}{\partial x^{\beta}}T^{\beta}{}_{\nu}\frac{\partial x^{\nu}}{\partial x^{\sigma}}-T'^{\alpha}{}_{\sigma}
\end{aligned}$$

and we have the transformation law for the matrix,

$$T'^{\alpha}{}_{\sigma}=\sum_{\nu=0}^{3}\sum_{\beta=0}^{3}\frac{\partial x'^{\alpha}}{\partial x^{\beta}}T^{\beta}{}_{\nu}\frac{\partial x^{\nu}}{\partial x^{\sigma}}$$

Several things become evident from this calculation. First, we do not want to be writing all those summation symbols! Looking closely, we see that there are two types of index, raised and lowered. In every case where we have a sum, we have exactly two identical indices, and one is raised and one is lowered. From here on, we employ the Einstein summation convention, and omit the explicit $\sum$ symbol. Instead we sum automatically whenever we find a pair of matching indices with one raised and one lowered. The indices in such a summed pair are called dummy indices, because they can be changed at will. Other indices, which occur singly in each term, must always match in name and position with a corresponding index in each term. These are called free indices, and they tell us what type of object we have. For example, in the expression above, α and σ are free indices. They each occur once on each side of the equation, and α is raised and σ lowered. Using the summation convention, we rewrite:

$$T'^{\alpha}{}_{\sigma}=\frac{\partial x'^{\alpha}}{\partial x^{\beta}}T^{\beta}{}_{\nu}\frac{\partial x^{\nu}}{\partial x^{\sigma}}$$

and because β and ν are dummy indices, this means the same thing as

$$\begin{aligned} T'^{\alpha}{}_{\sigma} &= \frac{\partial x'^{\alpha}}{\partial x^{\mu}} T^{\mu}{}_{\lambda} \frac{\partial x^{\lambda}}{\partial x^{\sigma}} \\ &= \frac{\partial x'^{\alpha}}{\partial x^{\mu}} \frac{\partial x^{\lambda}}{\partial x^{\sigma}} T^{\mu}{}_{\lambda} \end{aligned}$$

Notice a second property of this transformation. $T^{\mu}{}_{\lambda}$ has one index of each type, and they transform differently. The raised index transforms just like a contravariant vector, with the Jacobian matrix $\frac{\partial x'^{\alpha}}{\partial x^{\mu}}$. However, the lowered index transforms with the inverse matrix, $\frac{\partial x^{\alpha}}{\partial x'^{\mu}}$. Raised indices are called contravariant and lowered indices are called covariant.

Suppose we have a covariant vector and a contravariant vector,

$$A_{\alpha}, B^{\beta}$$

We define the inner product or scalar product to be the sum

$$A_{\alpha} B^{\alpha} = B^{\alpha} A_{\alpha}$$

and this quantity is invariant,

$$\begin{aligned} A'_{\alpha} B'^{\alpha} &= \left(\frac{\partial x^{\mu}}{\partial x^{\alpha}} A_{\mu} \right) \left(\frac{\partial x'^{\alpha}}{\partial x^{\nu}} B^{\nu} \right) \\ &= \left(\frac{\partial x^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\alpha}}{\partial x^{\nu}} \right) A_{\mu} B^{\nu} \\ &= \delta^{\mu}_{\nu} A_{\mu} B^{\nu} \\ &= A_{\mu} B^{\mu} \end{aligned}$$

As we have seen, Lorentz transformations preserve the quadratic form

$$ds^2 = \left(dx^0\right)^2 - \left(dx^1\right)^2 - \left(dx^2\right)^2 - \left(dx^3\right)^2$$

We can write this as a double sum

$$ds^2 = g_{\alpha\beta} dx^{\alpha} dx^{\beta}$$

where

$$g_{\alpha\beta} = \eta_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Notice that there is a sign ambiguity here. We could equally well use the opposite sign of ds^2 to define the metric, in which case we would have

$$\eta_{\alpha\beta} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

This is actually the more common practice (see, e.g., Misner, Thorne & Wheeler, or Weinberg), but here we will stay with Jackson's convention. We will use the symbol $g_{\alpha\beta}$ to refer to a general metric which may have functions as entries, and reserve the symbol $\eta_{\alpha\beta}$ for this particular, constant, orthonormal matrix. For example, if we write the line element ds^2 in spherical coordinates,

$$ds^2 = \left(dx^0\right)^2 - \left(dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2\right)$$

then

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix}$$

Since $g_{\alpha\beta}$ has two covariant indices, we say it is type $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$. We define the inverse to $g_{\alpha\beta}$ with the symbol $g^{\alpha\beta}$ so that

$$g^{\alpha\beta}g_{\beta\mu} = \delta^{\alpha}_{\mu}$$

For the orthonormal metric $\eta_{\alpha\beta}$ we see that

$$\eta^{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

while for in spherical coordinates we write

$$g^{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -\frac{1}{r^2} & 0 \\ 0 & 0 & 0 & -\frac{1}{r^2\sin^2\theta} \end{pmatrix}$$

Now look at the inner product of two vectors,

$$\begin{aligned} A^0B^0 - \mathbf{A}\cdot\mathbf{B} &= A^0B^0 - A^1B^1 - A^2B^2 - A^3B^3 \\ &= \eta_{\alpha\beta}A^{\alpha}B^{\beta} \end{aligned}$$

This inner product is the same as the covariant-contravariant sum if we define

$$B_{\alpha} = \eta_{\alpha\beta}B^{\beta}$$

for then we have

$$\eta_{\alpha\beta}A^{\alpha}B^{\beta} = A^{\alpha}B_{\alpha}$$

Because $\eta_{\alpha\beta}$ is symmetric, we could equally well write

$$\begin{aligned} \eta_{\alpha\beta}A^{\alpha}B^{\beta} &= \eta_{\beta\alpha}A^{\alpha}B^{\beta} \\ &= A_{\beta}B^{\beta} \\ &= A_{\alpha}B^{\alpha} \end{aligned}$$

We say that contravariant vectors, A^{α}, are rank-1 tensors of type $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ while covariant vectors, A_{α}, are rank one tensors of type $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

We can now define rank two tensors of types $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ as 4×4 matrices which tranform as follows:

$$\begin{aligned} \tilde{R}^{\alpha\beta} &= R^{\mu\nu}\tfrac{\partial\tilde{x}^{\alpha}}{\partial x^{\mu}}\tfrac{\partial\tilde{x}^{\beta}}{\partial x^{\nu}} \qquad \textit{type} \begin{pmatrix} 2 \\ 0 \end{pmatrix} \\ \tilde{S}^{\alpha}{}_{\beta} &= S^{\mu}{}_{\nu}\tfrac{\partial\tilde{x}^{\alpha}}{\partial x^{\mu}}\tfrac{\partial x^{\nu}}{\partial\tilde{x}^{\beta}} \qquad \textit{type} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \tilde{T}_{\alpha\beta} &= T_{\mu\nu}\tfrac{\partial x^{\mu}}{\partial\tilde{x}^{\alpha}}\tfrac{\partial x^{\nu}}{\partial\tilde{x}^{\beta}} \qquad \textit{type} \begin{pmatrix} 0 \\ 2 \end{pmatrix} \end{aligned}$$

Every raised index transforms with a factor of $\frac{\partial \tilde{x}^\alpha}{\partial x^\mu}$, while every lowered index transforms with the inverse matrix, $\frac{\partial x^\mu}{\partial \tilde{x}^\alpha}$. This means that whenever we sum over one raised and one lowered index, the transformation matrices cancel, leaving the sum invariant. For example, even though

$$J_\alpha T^{\alpha\beta}$$

has three indices, the two α-indices are dummy indices and the sum over them is invariant. Therefore, the quantity transforms as a $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ tensor, and not a $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ tensor:

$$\tilde{J}_\alpha \tilde{T}^{\alpha\beta} = \frac{\partial \tilde{x}^\beta}{\partial x^\mu} \left(J_\alpha T^{\alpha\mu} \right)$$

We could go on to describe higher rank tensors of type $\begin{pmatrix} p \\ q \end{pmatrix}$ with $n = p + q$ indices, but we will not need them for electrodynamics.

A particularly important $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ tensor is the partial derivative operator, $\partial_\alpha \equiv \frac{\partial}{\partial x^\alpha} = \left(\frac{\partial}{\partial x^0} \cdot \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3} \right) = \left(\frac{\partial}{\partial x^0}, \boldsymbol{\nabla} \right)$. The chain rule shows that this is a covariant vector,

$$\frac{\partial}{\partial \tilde{x}^\alpha} = \frac{\partial x^\beta}{\partial \tilde{x}^\alpha} \frac{\partial}{\partial x^\beta}$$

This also an operator; acting on a contravariant vector, the result is a $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ tensor,

$$\begin{aligned} T_\alpha^{\ \ \beta} &= \frac{\partial A^\beta}{\partial x^\alpha} \\ &= \partial_\alpha A^\beta \end{aligned}$$

Setting $\alpha = \beta$ and summing gives an invariant, the divergence,

$$\begin{aligned} \partial_\alpha A^\alpha &= \partial^\alpha A_\alpha \\ &= \frac{\partial A^0}{\partial x^0} + \boldsymbol{\nabla} \cdot \mathbf{A} \end{aligned}$$

We can write the same operator in contravariant form,

$$\begin{aligned} \partial^\alpha &= \eta^{\alpha\beta} \partial_\beta \\ \frac{\partial}{\partial x^\alpha} &= \eta^{\alpha\beta} \frac{\partial}{\partial x^\beta} \\ &= \left(\frac{\partial}{\partial x^0}, -\boldsymbol{\nabla} \right) \end{aligned}$$

and using the two we may form the Lorentz invariant operator,

$$\begin{aligned} \square &\equiv \partial^\alpha \partial_\alpha \\ &= \left(\frac{\partial}{\partial x^0}, -\boldsymbol{\nabla} \right) \cdot \left(\frac{\partial}{\partial x^0}, \boldsymbol{\nabla} \right) \\ &= \frac{\partial^2}{\partial x^{02}} - \boldsymbol{\nabla}^2 \end{aligned}$$

This familiar wave operator is called the d'Alembertian.