A Review of Tensor Geometry

The most profound and most useful formulation of electromagnetism uses the language of tensor geometry. Jackson assumes far too much knowledge.

The present review (from Professor Robert Jaffe) presents enough background for our needs. The first section introduces the concept of tensors in the familiar world of 3-dimensional space (which we will need to treat the Maxwell stress tensor, for example). A second section (to come) extends these ideas to the 4-dimensional space-time of special relativity.

I. Geometry in 3-dimensions

1. Geometrical Objects

A. Rotations

We will always deal with good old, flat 3-dimensional space described by Cartesian coordinates. Distances are measured in the usual way. For two points \( \vec{x} = (x_1, x_2, x_3) \) and \( \vec{y} = (y_1, y_2, y_3) \) the distance between them is

\[
d^2(x, y) = \sum_{i=1}^{3} (x_i - y_i)^2
\]

Let's introduce a summation convention once and for all: Any repeated Latin index is understood to be summed from 1 to 3. So we rewrite \( d^2(\vec{x}, \vec{y}) \) as:

\[
d^2(\vec{x}, \vec{y}) = (x_i - y_i)(x_i - y_i)
\]

(1)

The objects in geometry are classified by their transformation properties under the changes of coordinates, in particular under rotations. Bear with this somewhat abstract notion to see why it's useful. So consider two coordinate systems \( S \) and \( S' \) related by a rotation. For convenience we'll write down a rotation about the \( z \)-axis but any rotation is allowed:

\[
\begin{align*}
x_1' &= x_1 \cos \theta + x_2 \sin \theta \\
x_2' &= -x_1 \sin \theta + x_2 \cos \theta \\
x_3' &= x_3
\end{align*}
\]

(2)
Using our summation convention

\[ x'_i = R_{ij}x_j \]  \hspace{1cm} (3)

where \( R_{ij} \) is the matrix

\[
R_{ij} = \begin{pmatrix}
    \cos \theta & \sin \theta & 0 \\
    -\sin \theta & \cos \theta & 0 \\
    0 & 0 & 1
\end{pmatrix} \]  \hspace{1cm} (4)

Coordinates are defined globally. Writing down \( \vec{x} \) always implies some choice of origin. It's easier to deal with things defined without reference to a specific origin. For example, coordinate differences -- or in the limit, the differentials \( dx_i' \). From eq.(3) we see

\[ dx'_i = R_{ij}dx_j \]  \hspace{1cm} (5)

or equivalently \[ \frac{\partial x'_i}{\partial x_j} = R_{ij} \]  \hspace{1cm} (6)

So the transformation of differentials is just the chain rule:

\[ dx'_i = \left( \frac{\partial x'_i}{\partial x_j} \right) dx_j \]  \hspace{1cm} (7)

The most important characteristic of rotations for us is that they preserve lengths. Rotate coordinates and you have not changed the distance between any two points. Mathematically

\[ (ds)^2 = dx_i dx_i \]

and

\[ (ds')^2 = dx'_i dx'_i \]

are equal. Thus from eq. (5):

\[ dx'_i dx'_i = R_{ij} R_{ik} dx_j dx_k = dx_j dx_j \]

so we find \[ R_{ij} R_{ik} = \delta_{jk} \]. That is, the rotation matrix is orthogonal (its transpose is its inverse)

\[ (R^T R)_{ij} = (R^T)_{im} R_{mj} = R_{mi} R_{mj} = \delta_{ij} = (I)_{ij} \]  \hspace{1cm} (8)

With these preliminaries we can launch on the classification of geometrical objects.

B. Scalars/Scalar fields

A scalar field is a function of coordinates whose value at a given point is independent of coordinate system. However, points in one coordinate system are labelled differently than points in
another, so the functional rule may look different. If \( \phi(x) \) is a scalar then \( \phi'(x') = \phi(x) \). Why is there a prime on \( \phi \) in the new coordinate system? Because the function is a different function of \( x' \) than \( \phi \) was of \( x \).

**Example:**
Temperature is a scalar. Let the temperature in a room vary linearly with height \( x_2 \): \( T(x) = \alpha x_2 \). In a new coordinate system the temperature is the same, requiring (using eq. (2))

\[
T'(x') = \alpha x_1' \sin \theta + \alpha x_2' \cos \theta = \alpha x_2 = T(x)
\]

You see, \( T' \) is a different function of \( x' \) than \( T \) is of \( x \).

**NOTE:** I am always considering what some books call passive transformations: The physical system sits still, only the coordinates are relabelled. The other possibility: the physical system is actually (actively) rotated is another way to do it. To avoid sign mistakes it's best to commit yourself entirely to one viewpoint.

C. **Vectors/Vector fields**

A vector field is defined to be a set of three functions of the coordinates \( v_i(x) \), \( i = 1, 2, 3 \), which transform under rotations in the same way as do the coordinate differentials \( dx_i \):

\[
v_i'(x') = \frac{\partial x_1'}{\partial x_j} v_j(x) = R_{ij} v_j(x)
\]  

(9)

The elementary concept of a vector as a direction plus a length is included in this definition. The important thing to realize is that \( v_i' \) and \( v_i \) represent the same vector but referred to different axes. Unlike a scalar the components of a vector actually change under rotation because they are defined along new axes.

**Example:** Consider the vector field \( v_i(x) = (1, 0, 0) \), i.e. to each point is assigned a unit vector in the \( \hat{e}_1 \) direction. Applying the rotation eq. (2) we find:

\[
v_1' = v_1 \cos \theta, \quad v_2' = -v_1 \sin \theta
\]

Under rotation of axes components of \( v \) transform into one another but the vector is the same. Examples of vectors abound: the velocity field of a fluid assigns to each point the local fluid velocity
vector; the electric and magnetic fields are vectors.

D. Tensors/Tensor fields

Tensors are a natural and simple generalization of vectors. Consider the "outer product" of two vectors, just the set of 9 numb $u_i v_j$ which can for convenience be thought of as a $3 \times 3$ matrix

$$ T_{ij} \equiv u_i v_j = \begin{pmatrix} u_1 v_1 & u_1 v_2 & u_1 v_3 \\ u_2 v_1 & u_2 v_2 & u_2 v_3 \\ u_3 v_1 & u_3 v_2 & u_3 v_3 \end{pmatrix} \quad (10) $$

It's easy to write down the transformation law for $T_{ij}$ knowing the law for vectors (eq.(9))

$$ T_{ij}^\prime = R_{ik} R_{jl} T_{kl} \quad (11) $$

Any object which transforms under rotations according to eq.(11) is called a rank 2 tensor. A rank 1 tensor is a vector while a scalar is a rank 0 tensor. The transformation law for a rank tensor is

$$ T_{i_1 i_2 \ldots i_n} = R_{i_1 j_1} R_{i_2 j_2} \ldots R_{i_n j_n} T_{j_1 j_2 \ldots j_n} \quad (12) $$

Not every tensor is a product of 2 vectors as in eq. (10). A very simple but illustrative example follows.

**Example:** The equation of an ellipsoid centered at the origin can be written

$$ x_i x_j A_{ij} = 1 \quad (13) $$

If the coordinate system is rotated the ellipsoid will be described by a similar formula in the new coordinates

$$ x'_i x'_j A'_{ij} = 1 \quad (14) $$

Let's prove $A_{ij}$ is a tensor. Using eq.(3) in eq.(14)

$$ R_{ik} x_k x'_j A'_{ij} = 1 $$

or $$ x_k x'_k (R_{ik} R_{jl} A'_{ij}) = 1 $$

Comparing with eq.(13) we see

$$ A_{kj} = R_{ik} R_{jl} A_{ij} $$
This looks backwards but we can use eq.(8) to fix that: first multiply eq.(15) by \( R_{mk} R_{nkl} \) and do the sums on \( k \) and \( l \):

\[
R_{mk} R_{nkl} A_{kl} = (R_{ik} R_{mk})(R_{jl} R_{nkl})A_{ij}'
\]
\[
= \delta_{im} \delta_{jn} A_{ij}'
\]
\[
= A_{mn}'
\]

This is the law given in eq.(11), so \( A_{ij}' \) forms a tensor. If you know the equation of an ellipsoid in any coordinate system you can find the equation in another system by the rule of eq.(11).

Tensors are useful because many physically interesting objects are tensors: For example:

The moment of inertia:

\[
I_{ij} \equiv \int d^3x \rho(x) (\delta_{ij} x_i^2 - x_i x_j)
\]

The quadrupole moment:

\[
Q_{ij} \equiv \int d^3x \rho(x) (3x_i x_j - \delta_{ij} x^2)
\]

The octupole moment:

\[
\Omega_{ijk} \equiv \frac{1}{2} \int d^3x \rho(x) (3x_i x_j x_k - \delta_{ij} x_k^2)
\]

(The octupole moment contributes to an electrostatic potential \( \Phi(x) = -\frac{x_i r_j x_k}{x^7} \Omega_{ijk} \))

It's easy to verify that all of these satisfy the transformation law eq.(11) using the same methods I used on \( A_{ij}' \).

The only difficulty about tensors is that you can't envision them as being little arrows like you can vectors.
E. Two Important Tensors:

i) THE KRONECKER DELTA: In the S coordinate system define

\[ \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \]

Now assume \( \delta_{ij} \) is a tensor and calculate its components in any other coordinate system \( S' \):

\[ \delta_{ij}' = R_{ij} R_{jk} \delta_{kl} = R_{ik} R_{jk} \]

but by eq(8) \( R_{ik} R_{jk} = \delta_{ij} \) so \( \delta_{ij}' = \delta_{ij} \). So this is a tensor whose components in all coordinate systems are identical. We conclude that the Kronecker \( \delta \) is a tensor and its components in any coordinate system are

\[ \delta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \]

ii) THE TOTALLY ANTISYMMETRIC TENSOR \( \varepsilon_{ijk} \)

The symbol \( \varepsilon_{ijk} \) is defined by

\[ \varepsilon_{123} = 1 \]

\[ \varepsilon_{ijk} = -\varepsilon_{jik} = -\varepsilon_{ikj} = -\varepsilon_{kji} \]

\( \varepsilon \) is odd under under interchanging any indices. From this it follows that \( \varepsilon \) is zero if any two indices are equal. Let's follow the same program as for \( \delta_{ij} \). Define a tensor which in one frame has the values of \( \varepsilon_{ijk} \) and calculate its components in another frame:

\[ \varepsilon_{ijk}' = R_{im} R_{jn} R_{kp} \varepsilon_{mnp} \]
Now I must appeal to a piece of linear algebra, where the following rule for the determinant of a 3x3 matrix is derived

\[ \epsilon_{ijk} \det M = M_{im} M_{jn} M_{kp} \epsilon_{mnp} \]  \hspace{1cm} (17)

so eq(16) becomes

\[ \epsilon'_{ijk} = \det R \epsilon_{ijk} \hspace{1cm} (17') \]

Referring back to eq.(8) we use the fact that \( \det(R) = \det(R^T) \) to conclude \( \det(RR^T) = (\det R)^2 = 1 \) and hence \( \det R = \pm 1 \). Those transformations with \( \det R = +1 \) are called \textit{proper} rotations while those with \( \det R = -1 \) are \textit{improper} and involve reflections \((x') = -x\). The composition of a reflection followed by a rotation (or vice versa) has \( \det R = \det R_{\text{ref}} \det R_{\text{rot}} = (-1)(+1) = -1 \) and so is improper. Now, looking back at (17'), definition (16) was not quite right if we want \( \epsilon_{ijk} \) to be the same in all reference frames. If \( \epsilon \) is to have the same values in all frames whether rotated or reflected

\[ \epsilon'_{ijk} = (\det R) R_{im} R_{jn} R_{kp} \epsilon_{mnp} \]  \hspace{1cm} (18)

is the correct transformation law. An object which transforms like a tensor except for an extra sign change under reflection (\( \det R = -1 \)) is a \textit{pseudotensor}. \( \epsilon_{ijk} \) is a rank 3 pseudotensor.

\[ \text{F. Some Consequences of Tensor Geometry} \]

1. The cross product of two vectors is a vector. Usually you "prove" this by drawing the cross product as a length and direction. Here's a proof

Let \( w_i = (\bar{u} \times \bar{v})_i = \epsilon_{ijk} u_j v_k \) in system \( S \).
Then in another coordinate system

\[ w'_i = \varepsilon'_{ijk} u'_j v'_k \quad \text{but} \quad \varepsilon'_{ijk} = \varepsilon_{mnpr} R_{jm} R_{kn} R_{kp} \det R \]

and

\[ u'_j = R_{jq} u_q \]

\[ v'_k = R_{kr} v_r \]

so

\[ w'_i = \varepsilon_{mnpr} R_{jm} R_{jq} (R_{kp} R_{kr}) u_q v_r \det R \]

\[ = \varepsilon_{mnpr} \delta_{nq} \delta_{pr} u_q v_r \det R \]

\[ = R_{im} (\varepsilon_{mnp} u_n v_p) \det R \]

\[ = (\det R) R_{im} w_m \]

Since \( w'_i \) transforms with an extra factor of \( \det R \) it is a pseudovector and has opposite properties under reflection as a vector.

Example \( \mathbf{L} = \mathbf{\hat{r}} \times \mathbf{\hat{p}} \) defines angular momentum. Under reflection \( \mathbf{\hat{r}} \rightarrow -\mathbf{\hat{r}}, \mathbf{\hat{p}} \rightarrow -\mathbf{\hat{p}} \) but \( \mathbf{L} \rightarrow -\mathbf{L} \).

II. The components of an antisymmetric rank 2 tensor form a pseudovector. (ie a cross product)

Let \( T_{ij} = -T_{ji} \) then it's of the form

\[ T_{ij} = \begin{pmatrix}
0 & t_3 & -t_2 \\
-t_3 & 0 & t_1 \\
t_2 & -t_1 & 0
\end{pmatrix} \]

Note it only has 3 independent, non-zero entries. But we can rewrite this as

\[ T_{ij} = \varepsilon_{ijk} t_k \quad \text{(19)} \]

(check: \( T_{12} = \varepsilon_{123} t_3 = t_3 \)

\( T_{32} = \varepsilon_{321} t_1 = -t_1 \) etc.)
I claim the known transformation properties of $\epsilon_{ijk}$ and $T_{ij}$ make $t_k$ a pseudovector:

$$\epsilon_{ijk}\epsilon_{ijl} = 2\delta_{kl}$$
(prove this?)

Then eq(19) becomes:

$$t_k = \frac{1}{2} \epsilon_{kij} T_{ij}$$
(check this?)

Now rotate:

$$t'_k = \frac{1}{2} \epsilon'_{kij} T'_{ij}$$

$$= \frac{1}{2} \epsilon_{mnp} R_{km} R_{in} R_{jp} R_{is} R_{jr} T_{sr} (\det R)$$

$$= \frac{1}{2} \epsilon_{mnp} R_{km} \delta_{ns} \delta_{pr} T_{sr} (\det R)$$

$$= (\det R) R_{km} (\frac{1}{2} \epsilon_{mnp} T_{np})$$

$$= (\det R) R_{km} t_m$$

III. Any rank two tensor can be thought of as a combination of an ellipsoid and a pseudovector.

First note that any tensor can be written as the sum of an antisymmetric and a symmetric tensor:

$$T_{ij} = \frac{1}{2} (T_{ij} + T_{ji}) + \frac{1}{2} (T_{ij} - T_{ji})$$

$$= T^S_{ij} + T^A_{ij} \quad \text{where } T^S_{ij} = T^S_{ji}$$

$$T^A_{ij} = -T^A_{ji}$$

We've already shown $T^A_{ij}$ is equivalent to a pseudovector. Now consider the ellipsoid

$$x_i x_j T_{ij} = 1$$
Only $T_{ij}^S$ contributes to this

$$x_i x_j T_{ij}^A = -x_i x_j T_{ji}^A = -x_j x_i T_{ij}^A = 0$$

so the symmetric part of $T_{ij}$ defines an ellipsoid.

Thus we see that the study of a rank two tensor is only as complicated as the study of the geometry of an ellipsoid. The pseudovector part of $T_{ij}$ can be treated using vector notation—it doesn't require tensors at all.

You may remember that in 8.06 the moment of inertia (a symmetric tensor) was described in terms of an inertial ellipsoid. This is the justification. Also you may have heard reference to the angular momentum tensor as opposed to the vector $\mathbf{I}$ you know and love. It's the tensor

$$\left( \begin{array}{ccc} 0 & L_z & -L_y \\ -L_z & 0 & L_x \\ L_y & -L_x & 0 \end{array} \right)$$

which is being referred to.

G. The Gradient: Two Types of Vectors

Consider the gradient of a scalar field $\phi(x)$

$$\partial_i \phi = \frac{\partial \phi}{\partial x_i}$$

Using the chain rule for the change of coordinates $x'_i = R_{ij} x_j$ (3) we find

$$\frac{\partial \phi'}{\partial x'_i} = \frac{\partial \phi'}{\partial x_j} \frac{\partial x_j}{\partial x'_i}$$

but $\phi'(x') = \phi(x)$ for a scalar

so

$$\partial_i' \phi' = \frac{\partial x_j}{\partial x'_i} \partial_j \phi$$

(20)

From eq(3) and $R^T R = 1$ we find

$$x_j = R^T_{ji} x'_i = R_{ij} x'_i$$

so

$$\frac{\partial x_j}{\partial x'_i} = R_{ij}$$

and eq(20) becomes

$$\partial_i' \phi' = R_{ij} \partial_j \phi$$

(21)
Thus the gradient of a scalar transforms exactly like a vector.
It is a vector.
However we should be aware that eqs (9) and (20) are different.
Write them side by side
\[
v_i^j = \left( \frac{\partial x_i}{\partial x_j} \right) v_j
\]  
(22)
\[
a_i^j \phi = \left( \frac{\partial x_i}{\partial x^j} \right) a_j \phi
\]  
(23)
In general \( \frac{\partial x_i}{\partial x_j} \) is not equal to \( \frac{\partial x_i}{\partial x^j} \), but for rotations in Euclidian space they are equal. In 4-dimensional space-time they aren't quite equal, as we shall see.
To be specific, an object which transforms according to eq (22) is called a contravariant vector. One which transforms according to eq (23) is a covariant vector. The coordinates form a contravariant vector.

Apparently the partial derivatives with respect to a contravariant vector themselves form a covariant vector.
To reiterate: in 3-dimensional space so far as rotations are concerned there is no difference between contravariant and covariant vectors (that's why you never encountered them before).

2. METRIC GEOMETRY IN 3-DIMENSIONS

A. The Metric
A metric is the rank two tensor which converts coordinate differentials into the rotationally invariant length element:
\[
ds^2 = g_{ij} dx_i dx_j
\]
For cartesian coordinates in Euclidean space \( g_{ij} = \delta_{ij} \) the Kronecker \( \delta \)-symbol. In other coordinate systems it may change. For example in spherical coordinates: \( dx_1 = dr \), \( dx_2 = d\theta \); \( dx_3 = d\phi \) then
\[
ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2
\]
so
\[
g_{11} = 1, \ g_{22} = r^2, \ g_{33} = r^2 \sin^2 \theta\]
We'll mainly confine our attention to Cartesian coordinates. Since $g_{ij} = \delta_{ij}$ we've already proven it's a tensor and is (in fact) the same in all coordinate systems.

**B. Contraction**

If any tensor is multiplied by the metric tensor and the indices are summed then the resulting object transforms as though the summed indices were absent. For example consider $T_{ijk}$, a rank-3 tensor, then

$$\lambda_i = g_{jk} T_{ijk} = T_{ijj}$$

is a rank-1 tensor, a vector.

**Proof:**

$$\lambda_i = g'_{jk} T'_{ijk} \quad \text{use } g'_{jk} = \delta'_{jk} = \delta_{jk}$$

$$= \delta_{jk} R_{im} R_{jn} R_{kp} T_{mnp}$$

$$= R_{im} (R_{jn} R_{jp}) T_{mnp}$$

$$= R_{im} \delta_{np} T_{mnp}$$

$$= R_{im} (g_{np} T_{mnp})$$

$$= R_{im} \lambda_m$$

The process of multiplying by the metric tensor to convert tensors to objects of lower rank is called contracting. For our very simple case contracting is equivalent to setting indices equal and summing.

**EXAMPLES:**

**SCALAR PRODUCT:** $$(v \cdot w) = v_i w_j g_{ij}$$ is a scalar

**TRACE:** The trace of a rank two tensor is not changed by rotation

$$\text{tr } T = T_{jj} = g_{jk} T_{jk}$$

$$T_{jj} = T_{jj}$$

(This is useful because it tells us that the trace of a symmetric tensor is the sum of its eigenvalues.)
C. Invariants

A very useful application of contraction techniques is the construction of scalars out of tensors of higher rank. These invariants are useful because they transform very simply under coordinate rotations. The scalar product and the trace of a rank two tensor are simple invariants. Other examples are the totally contracted product of two tensors \( S_{ij} T_{ij} \), which may be thought of as the trace of the rank two tensor \( S_{ik} T_{kj} \); or the "triple scalar product" of three vectors \( \varepsilon_{ijk} u^i v_j w_k \). This latter is not quite invariant since it changes sign under reflections as you can convince yourself

\[
\varepsilon_{ijk} u^i v^j w_k = (\det R) \varepsilon_{ijk} u^i v^j w_k
\]

3. COVARIANCE AND SYMMETRY

Finally we are set to apply these geometrical ideas to physics. Even in 3-dimensions we can learn something useful. Suppose we believe that the laws of physics do not depend on the orientation of one's coordinate system. This means that the form of the laws must be the same in all coordinate systems related to one another by rotations. This way no experiment could detect a preferred reference frame since all phenomena conceivable in one frame are conceivable in all frames.

This will be so if and only if all the equations of physics equate objects of the same geometrical type. Vectors are equated to vectors, tensors to tensors, etc. Consider for example Newton's second law: \( F_i = m x_i \). Viewed from some rotated coordinate system we find

\[
F'_i = m' x'_i
\]

where \( F_i \) and \( x_i \) are vectors:

\[
F'_i = R_{ij} F_j; \quad x'_i = R_{ij} x_j
\]

and \( m \) is a scalar:

\[
m' = m.
\]
Equations where left and right hand sides transform the same way are called **covariant** (not related to covariant versus contravariant vectors). So the **invariance** or **symmetry** of a theory under rotations implies the **covariance** of the laws of physics under such transformations.

If we include reflections too (det $R = -1$) then we further require that the equations of physics equate vectors to vectors but not to pseudovectors, and so on.

**Example**

The magnetic field $\mathbf{B}$ is a pseudovector (we will see this later). So the force law

$$\mathbf{F}_i = e\mathbf{B}_i$$

violates reflection symmetry (charge is a scalar), because under reflection $\mathbf{F}'_i = -\mathbf{F}_i$ but $\mathbf{B}'_i = -(\text{det } R)\mathbf{B}_i = +\mathbf{B}_i$, so

$$\mathbf{F}'_i = -e\mathbf{B}'_i$$, a different force in the reflected coordinate system. On the other hand, the correct force law is OK:

$$\mathbf{F}_i = e(\nabla \times \mathbf{B})_i$$ because the cross product of a vector and a pseudovector is again a vector.

(The famous discovery of "parity" or reflection symmetry violation in weak interactions in 1957 involves just such an equation. The equations of motion of weak interactions change when you reflect the coordinate system!)