2. Notes on Tensors (Spring 2017)

References

There is a chapter on tensors in Boas. For cartesian tensors with many applications to physics see chapter 31 of The Feynman Lectures on Physics (volume 2). For the application of tensors to Special Relativity see 'Introduction to Special Relativity' by Wolfgang Rindler.

Vectors

A vector $\mathbf{V}$ is a geometrical object with a it magnitude and direction. It is convenient to describe a vector through components with respect to a set of *basis vectors:*

$$\mathbf{V} = V_1\mathbf{e}_1 + V_2\mathbf{e}_2 + V_3\mathbf{e}_3$$

The basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are linearly independent vectors and the components V_1, V_2, V_3 are real numbers.

For example one can take $\mathbf{e}_1 = \mathbf{i}$, $\mathbf{e}_2 = \mathbf{j}$, $\mathbf{e}_3 = \mathbf{k}$, that is unit vectors in the x, y and z directions. Other choices are possible. For now assume that the basis vectors are orthonormal, that is

$$\mathbf{e}_1 \cdot \mathbf{e}_2 = \mathbf{e}_2 \cdot \mathbf{e}_3 = \mathbf{e}_3 \cdot \mathbf{e}_1 = 0$$

and $|\mathbf{e}_1| = |\mathbf{e}_2| = |\mathbf{e}_3| = 1$. This can be expressed more compactly using the Kronecker delta:

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}.$$

The vector $\mathbf{V}$ can be written

$$\mathbf{V} = \sum_{i=1}^{3} V_i\mathbf{e}_i.$$

The summation sign can be omitted by using *Einstein's summation convention.* This convention is the understanding that any repeated indices are summed over. With this convention we can write

$$\mathbf{V} = V_i\mathbf{e}_i$$

Here i is a dummy index; as it is summed over it does not matter what letter is used $\mathbf{V} = V_i\mathbf{e}_i = V_j\mathbf{e}_j$.

The *dot product* of two vectors $\mathbf{U}$ and $\mathbf{V}$ can be written

$$\mathbf{U} \cdot \mathbf{V} = U_1V_1 + U_2V_2 + U_3V_3 = U_iV_i$$

using the Einstein summation convention. Note that this assumes that the basis vectors are orthonormal.

The *cross product* of two vectors is [1]

$$\mathbf{U}\times\mathbf{V} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ U_1 & U_2 & U_3 \\ V_1 & V_2 & V_3 \end{vmatrix} = (U_1V_2-U_2V_1)\mathbf{e}_3+(U_2V_3-U_3V_2)\mathbf{e}_1+(U_3V_1-U_1V_3)\mathbf{e}_2.$$

Alternatively the cross product can be defined using the Levi-Civita symbol; the three index object ϵ_{ijk} is defined by

$$\epsilon_{123} = 1$$

and the condition that ϵ_{ijk} is totally anti-symmetric. That is it changes sign under any interchange of two indices. This condition forces 21 of the 27 components to be zero, eg. $\epsilon_{113} = 0$. The six non-zero components are

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1, \quad \epsilon_{213} = \epsilon_{321} = \epsilon_{132} = -1.$$

The components of the cross product $\mathbf{U} \times \mathbf{V}$ are given by

$$(\mathbf{U} \times \mathbf{V})_i = \epsilon_{ijk}U_jV_k.$$

It is straightforward to check that this agrees with the standard definition of the cross product (e.g., $(\mathbf{U} \times \mathbf{V})_1 = \epsilon_{1jk}U_jV_k = \epsilon_{123}U_2V_3 + \epsilon_{132}U_3V_2 = U_2V_3 - U_3V_2$).

Vector Calculus

The gradient is defined through

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}$$

$$\nabla = \mathbf{e}_1\frac{\partial}{\partial x_1} + \mathbf{e}_2\frac{\partial}{\partial x_2} + \mathbf{e}_3\frac{\partial}{\partial x_3} = \mathbf{e}_i\frac{\partial}{\partial x_i}$$

where x_1, x_2, x_3 are components of the position vector $\mathbf{r}$ with respect to the orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, that is

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3.$$

[1] This assumes that $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are a right-handed system with $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$ and cyclic permutations.

This can also be written $\nabla = \mathbf{e}_i \partial_i$ using the shorthand

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

In component form $(\nabla\phi)_i = \partial_i \phi$.

The curl $\nabla \times \mathbf{F}$ has components

$$(\nabla \times \mathbf{F})_i = \epsilon_{ijk}\partial_j F_k.$$

In this notation the divergence of a factor field $\mathbf{F}$ is

$$\nabla \cdot \mathbf{F} = \partial_i F_i.$$

The Laplacian is

$$\nabla^2 = \nabla \cdot \nabla = \partial_i \partial_i.$$

Transformation Properties

A vector $\mathbf{V}$ can be written $\mathbf{V} = V_i \mathbf{e}_i$. In the following discussion the assumption that the basis vectors $\mathbf{e}_i$ are orthonormal is dropped. The position vector $\mathbf{r}$ can be written $\mathbf{r} = x_i \mathbf{e}_i$ where the x_i are the coordinates. Consider a linear change of coordinates

$$x_i = R_{ij} x_j.$$

R_{ij} can be viewed as a 3×3 matrix. Under such a transformation the basis vectors also transform

$$\mathbf{e}'_i = S_{ij}\mathbf{e}_j$$

where S_{ij} is another 3×3 matrix. Claim: if $S = (R^T)^{-1}$ then the vector $\mathbf{r}$ is unchanged by the transformation. To see this write $\mathbf{r}$ in two ways:

$$\mathbf{r} = x_i \mathbf{e}_i = x'_i \mathbf{e}'_i = R_{ij} x_j S_{ik} \mathbf{e}_k$$

which hold if $R_{ij} S_{ik} \mathbf{e}_k = \mathbf{e}_j$ which holds if $R_{ij} S_{ik} = \delta_{jk}$. As a matrix equation this is $R^T S = I$ or $S = (R^T)^{-1}$. If R is an orthogonal matrix then $S = R$. Recall that an orthogonal matrix is defined by the property $RR^T = I$. Any rotation matrix is an example of an orthogonal matrix. This has determinant 1. A parity transformation $R = \mathrm{diag}(-1,-1,-1)$ is also orthogonal. However it has determinant -1. Any orthogonal matrix has determinant ± 1 (proof $RR^T = I$, take the determinant, $\det RR^T = \det R \det R^T = (\det R)^2$ so that $\det R = \pm 1$).

An orthogonal transformation with determinant 1 is called a *proper rotation*. An orthogonal transformation with determinant -1 is called an *improper rotation* (this is a combination of a rotation and a parity transformation).

The components of a vector must transform in the same way as the position coordinates

$$V_i' = R_{ij} V_j.$$

The transformation rule is sometimes used as the *definition* of a vector. A vector can be understood to be three numbers V_i $(i = 1, 2, 3)$ with the same transformation properties as the position coordinates under an orthogonal transformation. This approach allows us to work with vectors without using basis vectors.

Tensors

A vector V_i a tensor of rank one which we can understand as 3 numbers which transform in the same way as the position cordinates x_i under the orthogonal transformation $x_i' = R_{ij} x_j$

A cartesian tensor of rank 2 is a two index object T_{ij} (which is 9 numbers) with the transformation property

$$T_{ij}' = R_{ip} R_{jq} T_{pq}.$$

The Kronecker delta δ_{ij} is a tensor of rank two. For it to be a tensor it must have the transformation property

$$\delta_{ij}' = R_{ip} R_{jq} \delta_{pq}$$

. But as δ_{ij} is a symbol it cannot transform (by definition) so for it to be a tensor requires $\delta_{ij}' = \delta_{ij} = R_{ip} R_{jq} \delta_{pq}$ which is satisfied if R is orthogonal. That is the Kronecker delta is a tensor even though it does not transform. This is an *isotropic* tensor of rank 2.

A *cartesian tensor* of rank p is a p index object (3^p numbers) $T_{i_1\ i_2\ \dots\ i_p}$ with the transformation property

$$T_{i_1\ i_2\ \dots\ i_p}' = R_{i_1 j_1} R_{i_2 j_2} \dots R_{i_p j_p} T_{j_1\ j_2\ \dots\ j_p}.$$

A rank 3 tensor is 27 numbers T_{ijk} with the transformation property

$$T_{ijk}' = R_{ip} R_{jq} R_{kr} T_{pqr}.$$

Remark: $\partial_i = \partial / \partial x_i$ is a vector operator under an orthogonal transformation $x_i' = R_{ij} x_j$ and can be used to construct tensors (more later).

Tensor Algebra

Tensors of the same rank may be added to produce a new tensor of the same rank, e.g. if T_{ij} and W_{ij} are tensors of rank two, then

$$W_{ij} = T_{ij} + U_{ij}$$

is also a tensor of rank two. Multiplying a tensor by a scalar gives a tensor of the same rank.

Tensors of any rank can be multiplied; if $T_{i_1 i_2 \dots i_p}$ has rank p and $U_{j_1 j_2 \dots j_q}$ has rank q then

$$W_{i_1 \dots i_p j_1 \dots j_q} = T_{i_1 i_2 \dots i_p} U_{j_1 j_2 \dots j_q}$$

has rank $p + q$. For example the tensor product of two vectors, U_i and V_j, yields a tensor of rank 2:

$$T_{ij} = U_i V_j.$$

Symmetric and Anti-symmetric Tensors

Consider tensors of rank 2 S_{ij} with the property

$$S_{ij} = S_{ji}.$$

This is called a symmetric tensor. While a tensor of rank 2 is nine quantities a symmetric tensor has six independent numbers. A tensor with the property

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

is called an anti-symmetric tensor which represents three independent quantities. Much as a function can be decomposed into even and odd parts a tensor of rank 2 can be decomposed into symmetric and anti-symmetric parts:

$$T_{ij} = \frac{T_{ij} + T_{ji}}{2} + \frac{T_{ij} - T_{ji}}{2}.$$

Here $\frac{1}{2}(T_{ij} + T_{ji})$ is the symmetric part of T_{ij} and $\frac{1}{2}(T_{ij} - T_{ji})$ is the anti-symmetric part.

If rank > 2 the situation is more complicated. A tensor can be symmetric or anti-symmetric in two of the n indices. For example $T_{ijk} = T_{jik}$. T_{ijk} can be be totally symmetric or totally anti-symmetric. Note that if a rank 3 tensor is totally anti-symmetric it is proportional to ϵ_{ijk}. At rank 4 there is no totally anti-symmetric tensor apart from the zero tensor.

Contraction

Given a tensor of rank $p \geq 2$ a new tensor of rank $p - 2$ can be obtained by *contraction*. This is a sum over two of the indices of a tensor (using the summation convention this amounts to setting two induces ‘equal’). For example the rank two tensor T_{ij} can be contracted to yield the scalar (or rank 0 tensor) T_{ii}. For example $\delta_{ii}$is the number three (not one) as $\delta_{ii} = \delta_{11} + \delta_{22} + \delta_{33} = 3$.

Contraction can be viewed as a generalisation of the dot product as

$$\mathbf{U} \cdot \mathbf{V} = U_i V_i$$

which is a contraction of the rank two tensor $U_i V_j$. The Laplacian can be viewed as a contraction of the tensor operator $\partial_i \partial_j$.

If the rank is greater than two there are several ways to contract. For example T_{ijk} has three possible contractions, T_{iik}, T_{iji}, T_{ijj} which are three distinct vectors (tensors of rank one). The cross product can also be obtained via a contraction. Now

$$(\mathbf{U} \times \mathbf{V})_i = \epsilon_{ijk} U_j v_k$$

can be viewed as a double contraction of the rank 5 object $\epsilon_{ijk} U_p V_q$.

Completely contracting a symmetric and anti-symmetric tensor gives zero

$$S_{ij} A_{ij} = 0.$$

This is because $S_{ij} A_{ij}$ has six non-zero contributions which cancel in pairs. What about a partial contraction such as $S_{ij} A_{jk}$?

Is ϵ_{ijk} a rank 3 tensor? Almost! It is a tensor under *proper* rotations (if R is orthogonal with $\det R = 1$). In general

$$R_{ip} R_{jq} R_{kr} \epsilon_{pqr} = \det R \ \epsilon_{ijk}.$$

ϵ_{ijk} is called a pseudo-tensor as it picks up an 'extra' minus sign under an 'improper' rotation. $T_{i_1 i_2 \ldots i_n}$ is rank n pseudo tensor if it has the transformation properly

$$T'_{i_1 i_2 \ldots i_n} = \det \ \ R \ R_{i_1 j_1} \ldots R_{i_n j_n} T_{j_1 j_2 \ldots j_n}.$$

δ_{ij} is an isotropic tensor of rank 2 and ϵ_{ijk} is an isotropic pseudo tensor of rank 3.

Such additional minus signs can also occur when transforming 'vectors'. Under an orthogonal transformation $x'_i = R_{ij} x_j$ a three-vector (also called a polar vector) transforms according to $V'_i = R_{ij} V_k$. The magnetic field has the transformation property

$$V'_i = \det R \ R_{ij} V_j.$$

Such a vector is called or pseudo-vector or axial vector. That is it transforms like a three vector under proper rotations but picks up an extra minus sign under improper rotations. Note that *the cross product of two polar vector is axial*

the cross product of a polar and axial vector is polar

the cross product of two axial vectors is axial

One can see why the magnetic field is an axial vector from the Lorentz force law

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

The momentum $\mathbf{p}$ and electric field $\mathbf{E}$ are polar. Accordingly, $\mathbf{v} \times \mathbf{B}$ must be polar which forces $\mathbf{B}$ to be axial.

Angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ is axial as it is a cross product of two polar vectors. The magnetic field is an axial vector. To see this consider the Lorentz force law

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

The momentum $\mathbf{p}$ and electric field $\mathbf{E}$ are polar. Accordingly, $\mathbf{v} \times \mathbf{B}$ must be polar which forces $\mathbf{B}$ to be axial.

The epsilon symbol can be used to convert an axial vector into a (non-pseudo) tensor of rank 2. The angular momentum tensor can be defined $L_{ij} = \epsilon_{ijk} L_k$ which is a rank two (non-pseudo) tensor. Note that L_{ij} is anti-symmetric. The three numbers L_i are encoded in an anti-symmetric tensor of rank 2. $L_{12} = \epsilon_{12k} L_k = \epsilon_{123} L_3 = L_3$. Similarly $L_{23} = L_1$, $L_{31} = L_2$ and $L_{11} = L_{22} = L_{33} = 0$. For the magnetic field one can define

$$F_{ij} = \epsilon_{ijk} B_k$$

which is a (non-pseudo) anti-symmetric tensor of rank 2. Much as for the angular momentum $F_{12} = -F_{21} = B_3$, $F_{23} = -F_{32} = B_1$, $F_{31} = -F_{13} = B_2$. The Lorentz force law can be written without reference to pseudo tensors:

$$\frac{dp_i}{dt} = q(E_i + F_{ij} v_j).$$

Recall that the magnetic field can be written $\mathbf{B} = \nabla \times \mathbf{A}$. The vector potential $\mathbf{A}$ is polar. In index notation

$$B_i = \epsilon_{ijk} \partial_j A_k.$$

. We have

$$F_{ij} = \epsilon_{ijk} B_k = \epsilon_{ijk} \epsilon_{kpq} \partial_p A_q$$

Now consider the (isotropic non-pseudo) tensor of rank 4

$$\epsilon_{ijk} \epsilon_{kpq} = \delta_{ip} \delta_{jq} - \delta_{iq} \delta_{jp}.$$

Therefore

$$F_{ij} = (\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp})\partial_p A_q = \partial_i A_j - \partial_j A_i$$

(this is an alternative formulation of the curl where the curl of a vector field is a tensor field). Similarly,

$$L_{ij} = x_i p_j - p_i x_j.$$

Maxwell's Equations

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \cdot \mathbf{B} = 0.$$

In index notation the first equation is

$$\partial_i E_i = \frac{\rho}{\epsilon_0}.$$

Try to wire the second equation using F_{ij} instead of $\mathbf{B}$ (see problems).

The preceding discussion assumed that the basis vectors are orthonormal $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$. In order to preserve this property we only considered orthogonal transformations. Now drop this assumption. Define the *metric* g_{ij} as

$$g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j.$$

The position vector $\mathbf{r}$ will be written $\mathbf{r} = x^i \mathbf{e}_i$ using superscripts for the coordinate indices. Use a summation convention with upper and lower indices. Expressions such as T_{ii} are meaningless as both indices are subscripts. The transformation properties are

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

where $R^i{}_j$ is not necessarily orthogonal. The basis vectors transform according to

$$\mathbf{e}'_i = S_i{}^j \mathbf{e}_j.$$

As matrices $S = (R^T)^{-1}$. The gradient

$$\partial_i = \frac{\partial}{\partial x^i}$$

transforms like the basis vectors

$$\partial_i' = \frac{\partial}{\partial x^{i'}} = S_i^{\ j}\partial_j.$$

Define two kinds of vectors. Consider the transformation $x^{i'} = R^i_{\ j}x^j$. If V^i are three numbers transform in the same way as the coordinates, that as

$$V^{i'} = R^i_{\ j}V^j,$$

then V^i is called a contravariant vector. If V_i transforms according to

$$V_i' = S_i^{\ j}V_j$$

then V_i is said to be a covariant vector. For example the gradient of a scalar field is a covariant vector field. Tensors can have both contravariant and covariant indices. A tensor of type (p, q) is 3^{p+q} numbers

$$T^{i_1 i_2 \ldots i_p}_{\qquad\ j_1 j_2 \ldots j_q}$$

with the transformation property

$$T^{i_1 i_2 \ldots i_p}_{\qquad\ j_1 j_2 \ldots j_q}{}' = R^{i_1}_{\ k_1} \ldots R^{i_p}_{\ k_p} S_{j_1}^{\ l_1} \ldots S_{j_q}^{\ l_q} T^{k_1 k_2 \ldots k_p}_{\qquad\ l_1 l_2 \ldots l_q}.$$

The Kronecker delta is a tensor of type $(1, 1)$. $\delta^i_{\ j}$. One can write

$$\delta^i_{\ j} = \frac{\partial x^i}{\partial x^j} = \partial_j x^i,$$

which clearly has one contravariant and on covariant index, g_{ij} is a tensor of type $(0, 2)$. The inverse metric, written g^{ij}, is a tensor of type $(2, 0)$. The metric can be used to convert a contravariant vector into a covariant one

$$V_i = g_{ij}V^j.$$

Similarly the inverse metric can be used to convert a covariant vector into a contravariant one

$$V^i = g^{ij}V_j.$$

The kinetic energy of a particle is $T = \frac{1}{2}m\mathbf{v}\cdot\mathbf{v}$. Now $\mathbf{r} = x^i\mathbf{e}_i$. Assuming the basis vectors are static $\mathbf{v} = \dot{\mathbf{r}} = \dot{x}^i\mathbf{e}_i$. This gives

$$T = \frac{1}{2}m x^i\mathbf{e}_i \cdot \dot{x}^j\mathbf{e}_j = \frac{1}{2}m g_{ij}\dot{x}^i\dot{x}^j.$$

The formalism is based on linear transformations of coordinates. Now consider a general change of coordinates (in particular non-linear teansformations). We have seen a hint of this in Lagrangian mechanics where the Euler-Lagrange equations have the same form in any coordinate system. However, this is not fully coordinate independent as the Lagrangian changes form under a coordinate transformation. For example a particle moving in three dimensional space has the Lagrangian

$$L = T - V = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - V(x, y, z),$$

where x, y and z are cartesian coordinates. In spherical polar coordinates (r, θ, ϕ) the Lagrangian is

$$L = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta\ \dot{\phi}^2\right) - V(r, \theta, \phi).$$

This change of variables is useful when the potential energy takes a simpler form in spherical polar coordinates.

In this example the form of the Lagrangian changes. The 'new' V is the 'old' V written as a function of the new coordinates. However, the form of the kinetic energy is different. In spherical polars the kinetic term is quadratic in the velocities but picks up position dependent coefficients. For example $\dot{\phi}^2$ is multiplied by $\frac{1}{2}mr^2 \sin^2\theta$. Now try to exploit the fact that the kinetic energy is quadratic in the velocities to write a general form for T. Let x^1, x^2, x^3 be coordinates of a particle (these can be any coordinates, cartesian, polar or otherwise) or

$$x^i \qquad i = 1, 2, 3.$$

Write T as a quadratic form in the velocities

$$T = \frac{m}{2} g_{ij}(x) \frac{dx^i}{dt}\frac{dx^j}{dt}.$$

where the $g_{ij}(x)$ can be viewed as the entries of a 3×3 symmetric matrix. It is 9 quantities (or 6 when taking into account that g_{ij} is symmetric)

$$g_{ij} = g_{ji}.$$

Essentially, the elements of g_{ij} are the coefficients of the quadratic terms in T. g_{ij} is called the *metric tensor.*

Examples Returning to the discussion of T in cartesian and spherical polar coordinates: In cartesian coordinates we have $g_{xx} = g_{yy} = g_{zz} = 1$ and $g_{xy} = g_{yz} = g_{zx} = 0$. In spherical polar coordinates $g_{rr} = 1$, $g_{\theta\theta} = r^2$, $g_{\phi\phi} = r^2 \sin^2\theta$, with all other components zero $g_{r\theta} = g_{\theta\phi} = g_{\phi r} = 0$.

In both examples the metric is diagonal ($g_{ij} = 0$ for $i \neq j$). In general, there will be off-diagonal terms due to 'cross terms' (eg. a term proportional to $\dot{x}^1\dot{x}^2$) in the kinetic energy.

An alternative definition of the metric is (again using the summation convention)

$$(ds)^2 = g_{ij}(x)dx^i dx^j$$

Here ds is the distance between 'neighbouring' points with coordinates

$g_{ij}(x)$ is a tensor of type $(2,0)$. A general contravariant vector V^i transforms like dx^i. A general covariant vector V_i transforms like ∂_i

$$V^{i'} = \frac{\partial x^{i'}}{\partial x^j} V^j$$

$$V_i' = \frac{\partial x^j}{\partial x^{i'}} V_j.$$

That is the matrices R and S become Jacobian matrices.