Summary Notes on Tensors

The following notes contain a summary of the material you need to know on tensors. The presentation closely follows that of the set book 'Introducing Einstein's Relativity' by Ray d'Inverno. The equations numbered in bold are the same as the numbers for the corresponding equation in the set book. However these notes are **not a substitute** for reading the book, but simply summarise the key points.

Part A: TENSOR ALGEBRA

§1 Manifolds and Coordinates

We will work with spaces which are *locally* like n-dimensional space. We call such objects *manifolds.*

A point on a manifold is specified by giving n coordinates:

$$(x^1, x^2, \cdots, x^n)$$

or more briefly as

$$x^a, \qquad a = 1, \cdots, n \tag{1}$$

However the choice of coordinates is not unique. Let $x'^1, x'^2, \cdots, x'^n$ be another set of coordinates, then we can write the new coordinates in terms of the old and vice-versa:

$$\begin{aligned}
x'^1 &= f(x^1, x^2, \cdots, x^n) \\
x'^2 &= g(x^1, x^2, \cdots, x^n) \\
\vdots &= \vdots \\
x'^n &= h(x^1, x^2, \cdots, x^n)
\end{aligned}$$

We write this as

$$x'^a = x'^a(x) \tag{\textbf{5.6}}$$

In transforming between coordinate systems it is useful to calculate the $n \times n$ *transformation matrix*

$$\left(\frac{\partial x'^a}{\partial x^b}\right) = \begin{pmatrix} \frac{\partial x'^1}{\partial x^1} & \cdots & \frac{\partial x'^1}{\partial x^n} \\ \vdots & & \vdots \\ \frac{\partial x'^n}{\partial x^1} & \cdots & \frac{\partial x'^n}{\partial x^n} \end{pmatrix} \tag{\textbf{5.7}}$$

The determinant of this matrix is the Jacobian

$$J' = \det\left(\frac{\partial x'^a}{\partial x^b}\right) \tag{2}$$

(Note: we write this as J' rather than J as it is the Jacobian for the transformation to the x' coordinates).

We can also express the x coordinates in terms of the x' and write

$$x^a = x^a(x')$$

Because this transformation is the inverse of (5.6) the corresponding transformation matrix is the inverse of the matrix (5.7). We can write this in components as

$$\sum_{c=1}^{n}\left(\frac{\partial x'^a}{\partial x^c}\right)\left(\frac{\partial x^c}{\partial x'^b}\right) = \delta^a_b \tag{3}$$

where the Kronecker delta δ^a_b is defined to be given by

$$\delta^a_b = \begin{cases} 1, & \text{if } a = b \\ 0, & \text{if } a \neq b \end{cases}$$

and represents the components of the identity matrix.

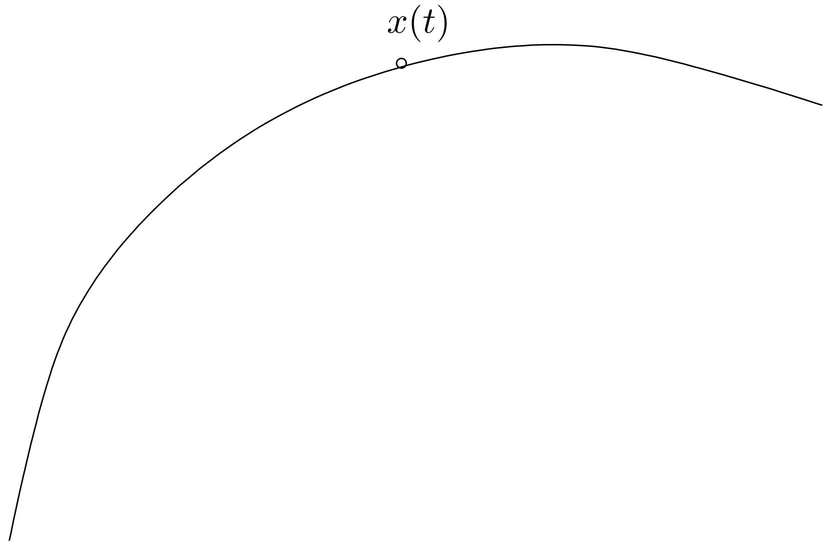
Since a matrix is an inverse for multiplication on either the left or right we also have

$$\sum_{c=1}^{n}\left(\frac{\partial x^a}{\partial x'^c}\right)\left(\frac{\partial x'^c}{\partial x^b}\right) = \delta^a_b \tag{4}$$

§2 Curves and Surfaces

Curves

Let t be some real parameter, then as t varies $x^a(t)$ traces out a curve. For example $x^a(t)$ might be the position of a particle at time t, then as t varies $x^a(t)$ moves along a curve (called the worldline of the particle).



Surfaces

In three dimensions a surface is a (smooth) 2-dimensional subspace. In n-dimensions a surface (or more properly a *hypersurface*) is a $(n-1)$-dimensional subspace. We can define a hypersurface by requiring that the coordinates of the points on a surface satisfy a constraint

$$\phi(x^1, x^2, \cdots, x^n) = 0 \tag{5}$$

For example if (x^1, x^2, x^3) are cartesian coordinates in 3-dimensions and

$$\phi(x^1, x^2, x^3) = (x^1)^2 + (x^2)^2 + (x^3)^2 - a^2$$

then

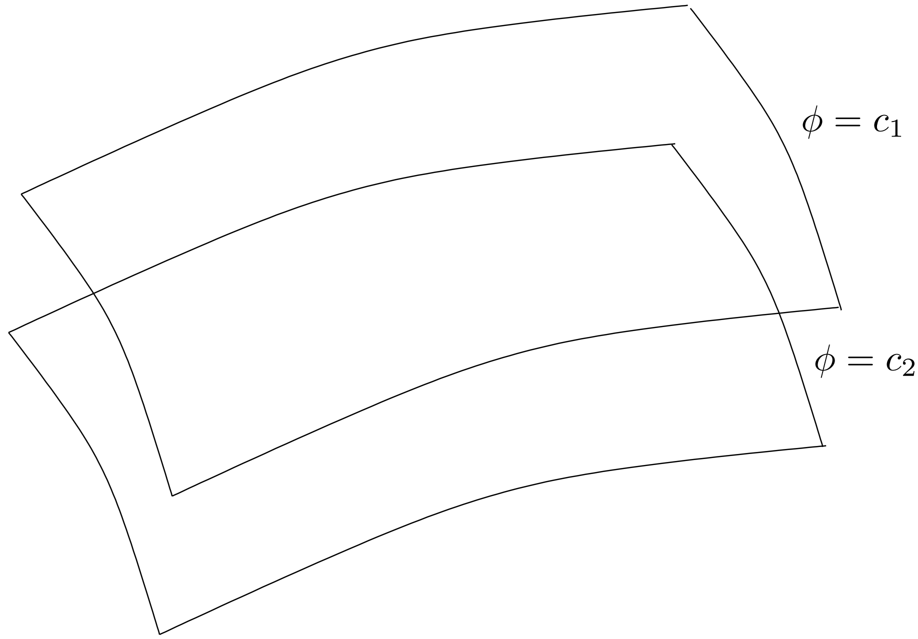
$$\phi(x^1, x^2, x^3) = 0$$

defines the surface of a sphere radius a.

More generally

$$\phi(x^1, x^2, \cdots, x^n) = c \tag{6}$$

defines a hypersurface for every fixed value of c. As c varies then (6) gives a 1-parameter family of surfaces.

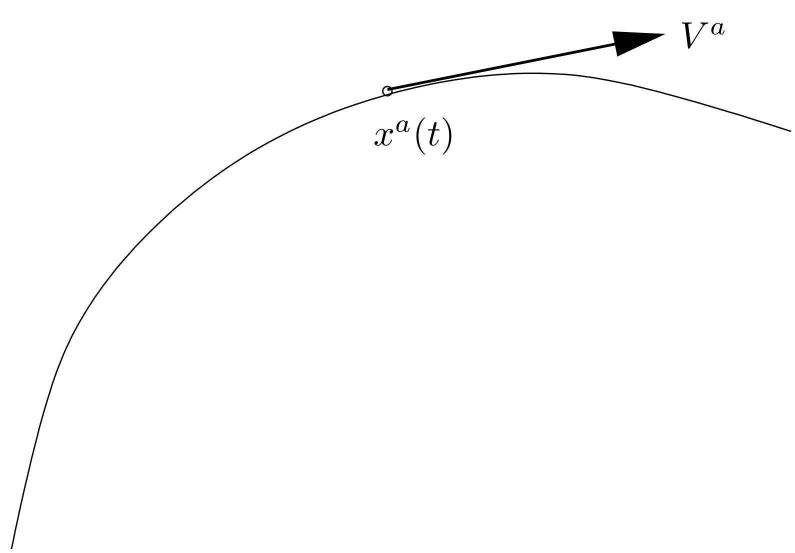


§3 Vectors and co-vectors

Vectors

We start by considering an example. Let a particle have position $x^a(t)$ at time t (as measured in the x coordinate system). Then the *velocity* of the particle V^a is given by

$$V^a = \frac{dx^a}{dt} \tag{7}$$



The velocity is an example of a vector.

What does the velocity look like in a different coordinate system?

$$\text{Let} \quad x'^a = x'^a(x^1, x^2, \cdots, x^n)$$
$$\text{Then} \quad x'^a(t) = x'^a(x^1(t), x^2(t), \cdots, x^n(t))$$

gives the position of the particle in the new coordinates.

So that

$$V'^a = \frac{dx'^a}{dt} \tag{8}$$

is the velocity of the particle in the new coordinates.

Now by the chain rule

$$\begin{aligned}
V'^1 &= \frac{dx'^1}{dt} \\
&= \frac{\partial x'^1}{\partial x^1}\frac{dx^1}{dt} + \frac{\partial x'^1}{\partial x^2}\frac{dx^2}{dt} + \cdots + \frac{\partial x'^1}{\partial x^n}\frac{dx^n}{dt} \\
&= \sum_{b=1}^{n} \frac{\partial x'^1}{\partial x^b}\frac{dx^b}{dt} \\
&= \sum_{b=1}^{n} \frac{\partial x'^1}{\partial x^b} V^b
\end{aligned}$$

And in general

$$V'^a = \sum_{b=1}^{n} \frac{\partial x'^a}{\partial x^b} V^b \tag{9}$$

We now introduce the *Einstein summation convention*: any index which appears twice is summed. (we call such an index a dummy index). Then using the summation convention (9) can be written as

$$V'^a = \frac{\partial x'^a}{\partial x^b} V^b \tag{\textbf{5.16}}$$

Note: the index b appears twice so is summed over.

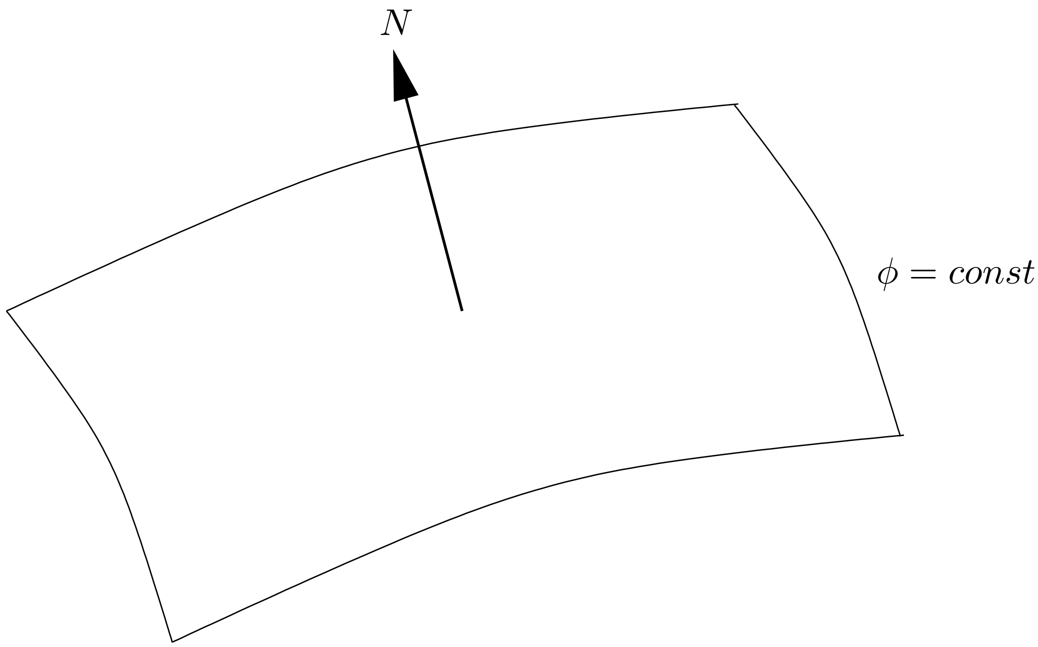
In general a (contravariant) vector is an object whose components transform according to (5.16)

Co-vectors

Another way of constructing vectors in 3-dimensions is using the gradient of a function.

Example

Let ϕ be a function of the cartesian coordinates x, y and z then $\nabla\phi$ is the vector with components $(\partial\phi/\partial x, \partial\phi/\partial y, \partial\phi/\partial z)$. Geometrically $\nabla\phi$ is normal to the surface $\phi = const.$



In n-dimensions we define the gradient of a function $\phi(x^1, x^2, \cdots, x^n)$ (in the x coordinate system) to be

$$\begin{aligned}\nabla\phi &= \left(\frac{\partial\phi}{\partial x^1}, \frac{\partial\phi}{\partial x^2}, \cdots, \frac{\partial\phi}{\partial x^n}\right)\\ &= (\phi_{,1}, \phi_{,2}, \cdots, \phi_{,n})\end{aligned}$$

We now define the co-vector $N = \nabla\phi$ to have components

$$N_a = \phi_{,a} \tag{10}$$

What does this look like in the x' coordinate system? By definition

$$N'_a = \frac{\partial\phi}{\partial x'^a} \tag{11}$$

To relate this to N_a we again use the chain rule:

$$\begin{aligned}
N'_1 &= \frac{\partial \phi}{\partial x'^1} \\
&= \frac{\partial \phi}{\partial x^1}\frac{\partial x^1}{\partial x'^1} + \frac{\partial \phi}{\partial x^2}\frac{\partial x^2}{\partial x'^1} + \cdots + \frac{\partial \phi}{\partial x^n}\frac{\partial x^n}{\partial x'^1} \\
&= \sum_{b=1}^{n} \frac{\partial \phi}{\partial x^b}\frac{\partial x^b}{\partial x'^1} \\
&= \frac{\partial \phi}{\partial x^b}\frac{\partial x^b}{\partial x'^1} \qquad \text{(using the Einstein summation convention)} \\
&= \frac{\partial x^b}{\partial x'^1} N_b
\end{aligned}$$

And in general

$$N'_a = \frac{\partial x^b}{\partial x'^a} N_b \tag{\textbf{5.21}}$$

In general a co-vector (or covariant vector) is an object whose components transform according to (5.21)

Contraction of a vector and a co-vector

Let V^a be a vector and N_a a co-vector, then we can define a new quantity $V^a N_a$ called the *contraction* of V with N. Remembering that we are using the summation convention we see that if we write this out in full we get

$$V^a N_a = V^1 N_1 + V^2 N_2 + \cdots + V^n N_n$$

so that the contraction is a bit like the dot product.

What does the contraction look like in the x' coordinate system?

$$\begin{aligned}
V'^a N'_a &= \left(\frac{\partial x'^a}{\partial x^b} V^b\right)\left(\frac{\partial x^c}{\partial x'^a} N_c\right) \\
&= \frac{\partial x'^a}{\partial x^b}\frac{\partial x^c}{\partial x'^a} V^b N_c \\
&= \delta^c_b V^b N_c \qquad \text{(using equation (3))} \\
&= V^c N_c \\
&= V^a N_a \qquad \text{(changing dummy index from } c \text{ to } a\text{)}
\end{aligned}$$

(Note: in the first line we must use **different** dummy indices for each term. If four dummy indicies had appeared it would have not been clear which terms to sum first. This gives the important rule: no index should appear more than twice.)

We have therefore obtained the important result

$$V^a N_a = V'^a N'_a \tag{12}$$

So that the contraction of the vector V with the co-vector N does not depend upon the coordinate system. This was important to Einstein because it showed that one could use these ideas to write down the laws of physics using tensors and produce experimental predictions that were independent of the coordinates.

§4 General Tensors

In the same way we defined vectors and co-vectors according to the transformation properties of their components we may define general tensors according to the way their components transform. The components of a rank 2 tensor define an $n \times n$ matrix at each point

We define a rank 2 contravariant tensor as an object which transforms according to

$$X'^{ab} = \frac{\partial x'^a}{\partial x^c} \frac{\partial x'^b}{\partial x^d} X^{cd}$$

We define a rank 2 mixed tensor as an object which transforms according to

$$Y'^a{}_b = \frac{\partial x'^a}{\partial x^c} \frac{\partial x^d}{\partial x'^b} Y^c{}_d$$

Finally we define a rank 2 covariant tensor as one whose components transform according to

$$Z'_{ab} = \frac{\partial x^c}{\partial x'^a} \frac{\partial x^d}{\partial x'^b} Z_{cd}$$

We say Z_{ab} is *symmetric* if

$$Z_{ab} = Z_{ba} \tag{13}$$

On the other hand we say Z_{ab} is *antisymmetric* if

$$Z_{ab} = -Z_{ba} \tag{14}$$

A type (p, q) tensor is one with p contravariant and q covariant indices. (Remember: *'co is below'*)

$$T^{a_1 \ldots a_p}{}_{b_1 \ldots b_q}$$

Under a change of coordinates the components of T transform according to

$$T'^{a_1 \ldots a_p}_{b_1 \ldots b_q} = \frac{\partial x'^{a_1}}{\partial x^{c_1}} \cdots \frac{\partial x'^{a_p}}{\partial x^{c_p}} \frac{\partial x^{d_1}}{\partial x'^{b_1}} \cdots \frac{\partial x^{d_q}}{\partial x'^{b_q}} T^{c_1 \ldots c_p}_{d_1 \ldots d_q} \tag{10}$$

The metric

In differential geometry a special role is played by a rank two tensor g_{ab} called the metric. The metric is a symmetric tensor (so that $g_{ab} = g_{ba}$) which has a well defined inverse at each point (so that $\det g \neq 0$).

The metric is used to measure the 'length' of a vector V. We define the length or modulus of a vector $|V|$ by

$$|V|^2 = g_{ab} V^a V^b$$

Because all the quantities transform as tensors and there are no free indices in this expression it turns out that this is independent of the coordinate system.

Example

If we work in 3-dimensional space and use Cartesian coordinates $(x^1, x^2, x^3) = (x, y, z)$ then the usual *Euclidean metric* is given by

$$g_{ab} = \delta_{ab} \qquad (\textbf{in Cartesian coordinates})$$

So that in Euclidean space and Cartesian coordinates

$$\begin{aligned} |V|^2 &= g_{ab} V^a V^b \\ &= \delta_{ab} V^a V^b \\ &= (V^1)^2 + (V^2)^2 + (V^3)^2 \end{aligned}$$

which agrees with the usual expression for the norm of a vector.

We may also use the metric to generalise the notion of dot product between two vectors V and W by defining the quantity $g_{ab} V^a W^b$. (Note this is the same as $g_{ab} W^a V^b$ since g_{ab} is symmetric).

If we return to the example of Euclidean space in Cartesian coordinates we find

$$\begin{aligned} g_{ab}V^aW^b &= \delta_{ab}V^aW^b \\ &= (V^1)(W^1) + (V^2)(W^2) + (V^3)(W^3) \\ &= V.W \end{aligned}$$

So that in Euclidean space and Cartesian coordinates we get precisely the usual dot product between vectors.

The contravariant metric

Because g_{ab} is invertible we may define its inverse. We denote the inverse of g_{ab} by g^{ab}. Taking the inverse of the equation

$$g'_{ab} = \frac{\partial x^c}{\partial x'^a}\frac{\partial x^d}{\partial x'^b} g_{cd} \tag{16}$$

We get

$$g'^{ab} = \frac{\partial x'^a}{\partial x^c}\frac{\partial x'^b}{\partial x^d} g^{cd} \tag{17}$$

so that g^{ab} is a (symmetric and invertible) contravariant tensor.

Note that because g^{ab} is the inverse of g_{ab} we have the two important equations:

$$g^{ac}g_{cb} = \delta^a_b \tag{18}$$

and

$$g_{ac}g^{cb} = \delta^a_b \tag{19}$$

Raising and lowering indices

Let X^a be a contravariant vector then we may define a new object $g_{ab}X^b$ by contracting X^a with the metric g_{ab}. Since both X^a and g_{ab} are tensors the new object is also a tensor. Since it has one downstairs index it must represent a *covariant* vector. Since it obtained from the (contravariant) vector X we call it the covariant form of X and write

$$X_a = g_{ab}X^b \tag{20}$$

We call this operation *lowering the index.*

In the same way if we are given a covariant vector Y_a we may define its contravariant form by contracting with g^{ab} so that

$$Y^a = g^{ab} Y_b \tag{21}$$

We call this operation *raising the index.*

However suppose we start with a vector X^a, then lower the index to obtain X_a we need to be sure that when we raise it again we get back to the *same* X^a that we started with. We now show that this is true:

$$X_b = g_{bc} X^c$$

So that

$$\begin{aligned} g^{ab} X_b &= g^{ab} g_{bc} X^c \\ &= \delta^a_c X^x \\ &= X^a \end{aligned}$$

It is possible to raise and lower *any* tensor index using the metric. For example if we start with the type $(0, 2)$ tensor T_{ab} we may define the type $(1, 1)$ tensor

$$T^a{}_b = g^{ac} T_{cb}$$

by raising the first index, and a (in general) different type $(1, 1)$ tensor

$$T_a{}^b = T_{ac} g^{cb}$$

by raising the second index. We can also define a type $(2, 0)$ tensor by raising both indices

$$T^{ab} = g^{ac} g^{bd} T_{cd}$$

The line element

An alternative notation is also sometimes used to represent a metric. Given the metric g_{ab} we (formally) define the *line element* ds^2 according to

$$ds^2 = g_{ab} dx^a dx^b \tag{22}$$

When we write this equation we view it as ds measuring the infinitesimal distance between points with coordinates x^a and $x^a + dx^a$.

The length of a curve

The above interpretation of the line element enables us to define the length of a curve.

Let $x^a(t)$ for $p < t < q$ define some curve γ. Then we may define the *length* of the curve by

$$\ell(\gamma) = \int_\gamma ds \tag{23}$$

where the integral is performed along the curve γ. In order to compute this integral we divide equation (22) by dt^2 and take the square root to obtain

$$\frac{ds}{dt} = \sqrt{g_{ab}\frac{dx^a}{dt}\frac{dx^b}{dt}} \tag{24}$$

Using this we see that the length of the curve γ is given by

$$\ell(\gamma) = \int_{t=p}^{q} \sqrt{g_{ab}\frac{dx^a}{dt}\frac{dx^b}{dt}}\,dt \tag{25}$$

This equation will prove important when we come to calculate the equation of a geodesic – a curve which minimizes the distance between two points – and generalizes the concept of straight line to a curved space(time).