Introduction to indicial notation

Theo Hopman*

October 2002

Vectors and their relationships are the same regardless of the coordinate system in which they are written. For example, the vector $(4,4)$ in Cartesian coordinates is the same vector as $\sqrt{8}\angle\frac{\pi}{4}$ written in plane polar coordinates. Similarly, in Cartesian coordinates, if a relationship is true for a vector, it is true for each of its components. This fact is the basis of indicial notation.

Indicial notation allows one to avoid geometrical proofs, which are often intuitively satisfying but inelegant. One no longer has to memorize numerous vector relationships when doing a symbolic proof. Otherwise complex and unwieldy results are reduced to simple algebra.

1 The basics

1.1 Free index

The key concept in indicial notation is that of an *index*. Consider a vector $\vec{v} = (v_x, v_y, v_z)$. The components of the vector are indexed in this case by the coordinate labels x, y and z. We could also write $\vec{v} = (v_1, v_2, v_3)$ where in this case $\vec{v}$ is indexed by the numbers 1, 2 and 3.

When doing arithmetic or algebra on vectors, we have to do the same thing to each component. Rather than going explicitly through the math for each component, we can work with a generic component indexed with a *free index*. If we were working on $\vec{v}$, for example, we could work with v_i, where i is a free index which can take on values 1, 2 or 3, or perhaps x, y or z. We don't specify which value i takes.

The name of a free index must remain the same throughout a calculation; it shouldn't be changed midway through. v_i, for example, must remain v_i throughout a calculation if i is a free index. We can't suddenly say that $v_i = v_j$, since j is another free index which can also take on any of the three coordinate labels. In general, $v_x \neq v_y$.

1.2 The Einstein summation convention

The Einstein summation convention is this: repeated Roman indices are summed over. Consider the dot product:

$$\vec{a} \cdot \vec{b} = \sum_{i=1}^{3} a_i b_i$$

Notice that in the right hand side, $\vec{a}$ and $\vec{b}$ are indexed by i. It's clear here what values of i are allowed $(1 \ldots 3)$, and what we're supposed to do with them. In the Einstein summation convention, we get rid of the (redundant) summation symbol so that

$$\vec{a} \cdot \vec{b} = a_i b_i \tag{1}$$

The repeated Roman index i is summed over. That is, we take all possible values of i, substitute them into $a_i b_i$, then sum the resulting terms.

*Department of Physics, University of Guelph

1.3 Dummy indices

An index which does not appear in an equation after a summation is carried out is called a *dummy index*. Because a dummy index does not appear in the final result, we can change its name to whatever is convenient. Just as $\sum_{i=1}^{3} a_i b_i = \sum_{j=1}^{3} a_j b_j$, so also $a_i b_i = a_j b_j$.

2 Necessary functions

There are two functions which are extremely useful in indicial notation: the Kronecker δ function and the Levi-Civita ϵ function.

2.1 The Kronecker δ

The Kronecker δ (delta) function is defined very simply:

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases} \tag{2}$$

Notice that both i and j are free indices here, and may each take on any value. Only for the three possible cases where $i = j$ is δ_{ij} non-zero. Note that $\delta_{ij} = \delta_{ji}$.

Consider: why does $\delta_{ii} = 3$, not 1?

2.2 The Levi-Civita ϵ

The Levi-Civita ϵ (epsilon) function is rather more complicated, since it is a function of three free indices, meaning there are 27 possible combinations. However, it can be summarized as

$$\epsilon_{ijk} = \begin{cases} 0 & \text{if any of } i, j \text{ or } k \text{ are the same,} \\ +1 & \text{if } ijk \text{ is an even permutation of 123,} \\ -1 & \text{if } ijk \text{ is an odd permuation of 123.} \end{cases} \tag{3}$$

By *even permutation* I mean that ijk is one of 123, 231 or 312. We see that these three combinations require an even number of swaps of indices to return them to 123, and that they are cyclic permutations of each other. For example, if we swap the first two indices in 312, we get 132. If we then swap the last two, we get 123, a total of two swaps, meaning that 312 is an even permutation of 123.

Just as there are three cases where ϵ_{ijk} is $+1$, there are three cases where it is -1. In these cases, ijk is one of 132, 321 or 213. You should be able to convince yourself that each of these combinations can be returned to 123 with and *odd* number of permutations. For example, $321 \to 231 \to 213 \to 123$. These three combinations are cyclic permutations of each other as well.

A property of ϵ_{ijk} is that

$$\epsilon_{ijk} = -\epsilon_{ikj} \tag{4}$$

since one side of the equation is an even permutation and the other is an odd permutation. (We can't say specifically that the left hand side is an odd permutation because ijk are free indices and may each take any value.)

2.3 Relating δ and ϵ

A very useful identity is

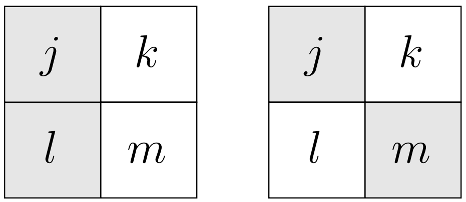
$$\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl} \tag{5}$$

This identity can be proved exhaustively, but we won't do that here since there are $3^4 = 81$ possible combinations of the free indices j, k, l and m, each of which has three terms in the implied sum over i. Instead, let's observe some features of the identity.

First of all, there is a dummy index i on the left hand side; this index is repeated, so it's summed over and doesn't appear on the right hand side. Any time you see two ϵ symbols with the same index, you can use this identity. It may be a little easier to use if you first put the ϵ symbols in the form above using cyclic permutation and (4).

Second, as a mnemonic device, notice that the indices on the deltas in first term on the right hand side are "parallel", while those on the deltas in the second term are "crossed". The diagram below may be helpful in seeing this.

$$\delta_{jl} \quad \delta_{km} \quad - \quad \delta_{jm} \quad \delta_{kl}$$



The first delta uses the first free index from each epsilon; the second uses the second free index from each epsilon. The third delta uses the first free index from the first epsilon and the second from the second, while the fourth uses the converse. Of course, within each term, the delta may be written in either order, and within each delta, the indices may be written in any order.

3 Vector operations

3.1 The dot product

We saw above in §1.2 that the dot product can be represented easily using indicial notation. We can associate the dot product with the Kronecker delta as follows:

$$\vec{a} \cdot \vec{b} = \sum_{i=1}^{3} \sum_{j=1}^{3} a_i b_i \delta_{ij}$$

By applying the properties of the Kronecker delta (zero unless $i = j$), we can replace all the j's we see with i's. When we apply the Einstein summation convention, we end up with (as we saw above)

$$\vec{a} \cdot \vec{b} = a_i b_i$$

3.2 Magnitude

The magnitude of a vector is closely related to the dot product. We can write

$$|\vec{v}| = v = \left(\vec{v} \cdot \vec{v}\right)^{1/2} = v_i v_i^{1/2} \tag{6}$$

3.3 The cross product

Not surprisingly, the cross product is also represented easily with indicial notation; it is associated with the epsilon symbol. The following identity

$$(\vec{a} \times \vec{b})_i = \epsilon_{ijk} a_j b_k \tag{7}$$

gives us the ith component of the cross product $\vec{a} \times \vec{b}$. Let's illustrate this by choosing a specific value for i, say 2, corresponding to the y component of the cross product. With $i = 2$, we'll carefully and slowly evaluate the right hand side of (7). First we explicitly write the (implied) sums, and evaluate one of them.

$$\begin{aligned} \epsilon_{2jk} a_j b_k &= \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{2jk} a_j b_k \\ &= \sum_{j=1}^{3} \left\{ \epsilon_{2j1} a_j b_1 + \epsilon_{2j2} a_j b_2 + \epsilon_{2j3} a_j b_3 \right\} \end{aligned}$$

But the ϵ_{2j2} term is zero since two of its indices are identical. We can sum each term separately.

$$= \sum_{j=1}^{3} \epsilon_{2j1} a_j b_1 + \sum_{j=1}^{3} \epsilon_{2j3} a_j b_3$$

Clearly in the first sum, the only possible value of j which gives a non-zero ϵ_{2j1} is $j = 3$. Similarly, only $j = 1$ in the second sum gives a non-zero term.

$$\begin{aligned} &= \epsilon_{231} a_3 b_1 + \epsilon_{213} a_1 b_3 \\ &= +a_3 b_1 - a_1 b_3 \end{aligned}$$

Therefore the y component of $\vec{a} \times \vec{b}$ is given by

$$(\vec{a} \times \vec{b})_y = a_z b_x - a_x b_z.$$

as we expect. If you wish, you can work through the similar algebra for the x ($i = 1$) and z ($i = 3$) components.

3.4 Determinants and the scalar triple product

Determinants and the scalar triple product are identical, and can be written in indicial notation. Suppose we wanted to find the determinant of the 3×3 matrix composed of the vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$. The determinant is

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1(b_2 c_3 - b_3 c_2) + a_2(b_3 c_1 - b_1 c_3) + a_3(b_1 c_2 - b_2 c_1)$$

You will notice that the parenthesized factor in the second term on the right hand side is identical (save for the change of variables $a \to b$ and $b \to c$) to the results for the y component of the cross product. Similarly, the factors in parentheses in the first and third terms can be identified with the x and z components of the cross product. (This shouldn't be surprising, since one way of finding a cross product is via a construction similar to a determinant.) We can therefore substitute these cross product components in.

$$= a_1 \epsilon_{1jk} b_j c_k + a_2 \epsilon_{2jk} b_j c_k + a_3 \epsilon_{3jk} b_j c_k$$

These terms all look very similar, except for the indices 1, 2 and 3. In fact, we can write this as a sum.

$$\begin{aligned} &= \sum_{i=1}^{3} a_i \epsilon_{ijk} b_j c_k \\ &= a_i \epsilon_{ijk} b_j c_k. \end{aligned}$$

This is just the scalar triple product, which is readily seen:

$$\vec{a} \cdot (\vec{b} \times \vec{c}) = a_i \big(\vec{b} \times \vec{c}\big)_i = a_i \epsilon_{ijk} b_j c_k.$$

3.5 The ∇ operator

Indicial notation can also be used with the $\vec{\nabla}$ operator. Since $\vec{\nabla}$ is an operator, not just a number, special care must be taken to keep the order of operations. That is, everything to the right of the operator must stay to the right of the operator.

The notation ∇_i is short for "take the partial derivative of what follows with respect to the ith component of $\vec{r}$". That is,

$$\nabla_i \equiv \frac{\partial}{\partial r_i}$$

where i is a free index which selects differentiation with respect to x, y or z.

The familiar gradient, divergence and curl are written as

$$\vec{\nabla} f = \nabla_i f, \tag{8}$$

$$\vec{\nabla} \cdot \vec{v} = \nabla_i v_i \quad \text{and} \tag{9}$$

$$\vec{\nabla} \times \vec{v} = \epsilon_{ijk} \nabla_j v_k. \tag{10}$$

Let's look a little more closely at, for example, the divergence.

$$\begin{aligned}\vec{\nabla}\cdot\vec{v} &= \nabla_i v_i \\ &= \sum_{i=1}^{3}\nabla_i v_i \\ &= \frac{\partial}{\partial r_1}v_1 + \frac{\partial}{\partial r_2}v_2 + \frac{\partial}{\partial r_3}v_3\end{aligned}$$

But r_1, r_2 and r_3 are just x, y and z respectively.

$$= \frac{\partial}{\partial x}v_1 + \frac{\partial}{\partial y}v_2 + \frac{\partial}{\partial z}v_3.$$

So we see that indicial notation gives us back the familiar form for the divergence.

4 Some worked examples

1. Show the vector triple product identity

$$\big(\vec{a}\times\vec{b}\big)\times\vec{c} = \big(\vec{a}\cdot\vec{c}\big)\vec{b} - \big(\vec{b}\cdot\vec{c}\big)\vec{a}.$$

We need only look at one component, which we do not specify, but identify with the free index i.

$$\begin{aligned}\Big[\big(\vec{a}\times\vec{b}\big)\times\vec{c}\Big]_i &= \epsilon_{ijk}\big(\vec{a}\times\vec{b}\big)_j c_k \\ &= \epsilon_{ijk}\epsilon_{jlm}a_l b_m c_k \\ &= \epsilon_{jki}\epsilon_{jlm}a_l b_m c_k \qquad \text{by cyclic rearrangement} \\ &= \big(\delta_{kl}\delta_{im} - \delta_{km}\delta_{il}\big)a_l b_m c_k \qquad \text{by identity (5)}\end{aligned}$$

We can multiply this out and then apply the Kronecker deltas. Recall that the delta has the effect of forcing indices to be the same (otherwise the term is zero).

$$\begin{aligned}&= a_k b_i c_k - a_i b_k c_k \\ &= \Big[\big(\vec{a}\cdot\vec{c}\big)\vec{b} - \big(\vec{b}\cdot\vec{c}\big)\vec{a}\Big]_i\end{aligned}$$

Since we haven't said what component i is, this equation must be true for *all* components. Therefore,

$$\big(\vec{a}\times\vec{b}\big)\times\vec{c} = \big(\vec{a}\cdot\vec{c}\big)\vec{b} - \big(\vec{b}\cdot\vec{c}\big)\vec{a} \qquad \text{Q.E.D.}$$

2. Show that

$$\big(\vec{a}\times\vec{b}\big)\times\big(\vec{c}\times\vec{d}\big) = \big(\vec{a}\vec{c}\vec{d}\big)\vec{b} - \big(\vec{b}\vec{c}\vec{d}\big)\vec{a}$$

where $\big(\vec{a}\vec{b}\vec{c}\big) \equiv \vec{a}\cdot\vec{b}\times\vec{c}$.

Recall that $\vec{a}\cdot\vec{b}\times\vec{c} - a_i\epsilon_{ijk}b_j c_k$. We can therefore write the right-hand side in indicial notation:

$$\begin{aligned}\text{RHS} &= b_i a_j \epsilon_{jkl}c_k d_l - a_i b_j \epsilon_{jkl}c_k d_l \\ &= \big(b_i a_j - a_i b_j\big)\epsilon_{jkl}c_k d_l\end{aligned}$$

We can't go much further here, so let's work with the left-hand side.

$$\begin{aligned}\text{LHS} &= \epsilon_{ijk}\big(\vec{a}\times\vec{b}\big)_j\big(\vec{c}\times\vec{d}\big)_k \\ &= \epsilon_{ijk}\epsilon_{jlm}a_l b_m \epsilon_{knp}c_n d_p \\ &= -\epsilon_{jik}\epsilon_{jlm}a_l b_m \epsilon_{knp}c_n d_p \\ &= -\big(\delta_{il}\delta_{km} - \delta_{im}\delta_{kl}\big)a_l b_m \epsilon_{knp}c_n d_p \\ &= -\big(a_i b_k - b_i a_k\big)\epsilon_{knp}c_n d_p \\ &= \big(b_i a_k - a_i b_k\big)\epsilon_{knp}c_n d_p\end{aligned}$$

We are in fact now done, even though this doesn't look exactly like the right-hand side. The differences, however, are merely dummy variables, whose names we are free to change. If we make the change of dummy variables $k \to j$, $n \to k$ and $p \to l$, we get

$$\begin{aligned}&= \big(b_i a_j - a_i b_j\big)\epsilon_{jkl}c_k d_l \\ &= \text{RHS}\end{aligned}$$

and the proof is complete.

3. Show that if $\vec{a}+\vec{b}+\vec{c}=0$, $\vec{a}\times\vec{b}=\vec{b}\times\vec{c}=\vec{c}\times\vec{a}$.

   We need to show two equalities for a complete proof, but since they are done in almost exactly the same way, we'll do only one here.

$$\begin{aligned}\vec{a}\times\vec{b} &= \epsilon_{ijk}a_jb_k\\ &= \epsilon_{ijk}a_j\big(-c_k-a_k\big)\\ &= -\epsilon_{ijk}a_jc_k-\epsilon_{ijk}a_ja_k\end{aligned}$$

   The second term in zero, since it corresponds to $\vec{a}\times\vec{a}$.

$$\begin{aligned}&= \epsilon_{ikj}c_ka_j\\ &= \vec{c}\times\vec{a}\end{aligned}$$

4. Show that $\vec{\nabla}\cdot\big(\phi\vec{A}\big)=\vec{\nabla}\phi\cdot\vec{A}+\phi\big(\vec{\nabla}\cdot\vec{A}\big)$.

   This is trivial in indicial notation. Note that $\nabla_i\phi=(\vec{\nabla}\phi)_i$.

$$\begin{aligned}\vec{\nabla}\cdot\big(\phi\vec{A}\big) &= \nabla_i\big(\phi A\big)_i \quad \text{in indicial notation}\\ &= \big(\nabla_i\phi\big)A_i+\phi\nabla_iA_i \quad \text{by the product rule}\\ &= \big(\vec{\nabla}\phi\big)\cdot\vec{A}+\phi\big(\vec{\nabla}\cdot\vec{A}\big)\end{aligned}$$

5. Evaluate $\big(\hat{r}\cdot\vec{\nabla}\big)\hat{r}$.