4 Index notation

4.1 Einstein Summation Convention

Recall that in n dimensions, the indices i, j etc. labelling the components of vectors run from 1 to n, e.g. we write $\underline{v} = (v_1, v_2, \ldots, v_n)$.

Einstein spotted that in quantities like

$$\underline{u}.\underline{v} = \sum_{i=1}^{n} u_i v_i$$

$$\underline{\nabla}.\underline{u} = \sum_{i=1}^{n} \frac{\partial u_i}{\partial x_i}$$

$$((\underline{a}.\underline{x})\underline{x})_i = \sum_{j=1}^{n} a_j x_j x_i,$$

the index to be summed appears exactly twice in a term or product of terms, while all other indices appear only once (the reason for this is to do with invariance under rotations, or for those of you studying Special Relativity this year, Lorentz transformations). He suggested dropping the summation sign, with the convention that wherever an index is repeated you sum over it.

So, we write

$$\underline{u}.\underline{v} = \sum_{i} u_i v_i \;=\; u_i v_i$$

$$\underline{\nabla}.\underline{u} = \sum_{i} \frac{\partial u_i}{\partial x_i} \;=\; \frac{\partial u_i}{\partial x_i}$$

$$((\underline{a}.\underline{x})\underline{x})_i = a_j x_j x_i.$$

We call the repeated indices dummy indices, and those that are not repeated are called free indices. The dummy indices can be renamed without changing the expression, i.e.

$$a_j x_j x_i = a_k x_k x_i,$$

since clearly

$$\sum_{j=1}^{n} a_j x_j x_i = \sum_{k=1}^{n} a_k x_k x_i = (\underline{a}.\underline{x}) x_i.$$

However, the free indices must match on both sides of an equation.

We must also be careful never to repeat an index more than twice in any single term or product of terms in an expression. If we were to write $a_i x_i x_i$, we can't tell whether this is supposed to be the component form of $(\underline{a}.\underline{x})x_i$, or of $(\underline{x}.\underline{x})a_i$. So if we want to write $(\underline{u}.\underline{v})^2$ in index notation, we should write $u_i v_i u_j v_j$ and not $u_i v_i u_i v_i$.

Although writing an expression like this in index notation can sometimes looks messy, we'll see shortly that it can be incredibly efficient for calculations.

4.2 The Kronecker delta, δ_{ij}

A very useful object for manipulating expressions in terms of their components, is the Kronecker delta. This is an object with two indices, defined by

$$\delta_{ij} \;\equiv\; \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \;.$$

We can think of the Kronecker delta as the components of the $n \times n$ identity matrix I, and in fact this was how you first met δ_{ij}. E.g. for $n = 3$

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \Rightarrow \quad I_{ij} = \delta_{ij},$$

where I_{ij} represents the element of I on the ith row and jth column.

The Kronecker delta appears naturally when we take partial derivatives, as we have

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij}$$

Example 15. If δ_{ij} is the 3-dim Kronecker delta, and $\underline{A} = (A_1, A_2, A_3)$, simplify

1. $A_s \delta_{ts}$
2. $\delta_{rs}\delta_{st}$
3. $\delta_{rs}\delta_{sr}$

Answers:

1.
$$\begin{aligned} A_s \delta_{ts} &= A_1\delta_{t1} + A_2\delta_{t2} + A_3\delta_{t3} \quad (0 \text{ if } t \neq 3 \text{ etc.}) \\ &= A_t \end{aligned}$$

2.
$$\begin{aligned} \delta_{rs}\delta_{st} &= \delta_{r1}\delta_{1t} + \delta_{r2}\delta_{2t} + \delta_{r3}\delta_{3t} \quad (0 \text{ unless } r = 3, t = 3 \text{ etc.}) \\ &= \delta_{rt} \end{aligned}$$

   From these two examples, we can now see the formal rule:

   If a δ has a dummy index, then delete the δ and replace the dummy index in the rest of the expression by the other index on the deleted δ. E.g. $A_s\delta_{ts} = A_t$.

3.
$$\begin{aligned} \delta_{rs}\delta_{sr} &= \delta_{rr} \qquad (\text{by } (2.)) \\ &= \delta_{11} + \delta_{22} + \delta_{33} = 3 \end{aligned}$$

4.3 The Levi-Cevita symbol, ϵ_{ijk}

If we are working in 3 dimension, there is a another device which is useful for handling expressions involving the vector cross product, such as $\underline{A} \times (\underline{B} \times \underline{C})$ or $\underline{\nabla} \times (\underline{\nabla} \times \underline{v})$.

Consider $\underline{C} = \underline{A} \times \underline{B}$. If $\underline{A} = A_1\underline{e}_1 + A_2\underline{e}_2 + A_3\underline{e}_3$, and $\underline{B} = B_1\underline{e}_1 + B_2\underline{e}_2 + B_3\underline{e}_3$, then

$$\begin{aligned} \underline{C} \;=\; \underline{A} \times \underline{B} &= \begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \\ &= \underline{e}_1(A_2B_3 - A_3B_2) + \underline{e}_2(A_3B_1 - A_1B_3) + \underline{e}_3(A_1B_2 - A_2B_1). \end{aligned}$$

The components of C are then given by

$$\begin{aligned} C_1 &= A_2B_3 - A_3B_2 \\ C_2 &= A_3B_1 - A_1B_3 \\ C_3 &= A_1B_2 - A_2B_1. \end{aligned}$$

We can write these equations as a single equation, by introducing ϵ_{ijk}, a set of numbers labelled by three indices i, j and k, each of which is equal to 1, 2 or 3.

This symbol is know as the Levi-Civita symbol, ϵ_{ijk} and is defined by:

$$\begin{array}{lll} I & \epsilon_{ijk} = -\epsilon_{jik} = -\epsilon_{ikj} & \text{(antisymmetric)} \\ II & \epsilon_{123} = 1 & \end{array}$$

These definitions imply the following properties:

1. $\epsilon_{ijk} = -\epsilon_{kji}$ (i.e. also antisymmetric when swapping 1st and 3rd index).

   Proof:
   $$\epsilon_{ijk} = -\epsilon_{jik} = +\epsilon_{jki} = -\epsilon_{kji}$$

2. $\epsilon_{ijk} = 0$ if any two indices have the same value.

   Proof: e.g.
   $$\epsilon_{112} = -\epsilon_{112} \Rightarrow 2\epsilon_{112} = 0 \Rightarrow \epsilon_{112} = 0$$

3. The only non-zero ϵ_{ijk} therefore have ijk all different (by property 2.), so (ijk) is some permutation of (123).

4. $\epsilon_{ijk} = +1$ if ijk is an even permutation of 123 ("even" = even # swaps)

   $\epsilon_{ijk} = -1$ if ijk is an odd permutation of 123 ("odd" = odd # swaps)

5. $\epsilon_{ijk} = \epsilon_{kij} = \epsilon_{jki}$ (cyclic permutations).

   So e.g. $1 = \epsilon_{123} = \epsilon_{312} = \epsilon_{231}$, and $-1 = \epsilon_{132} = \epsilon_{321} = \epsilon_{213}$

Most importantly, now the vector product $\underline{C} = \underline{A} \times \underline{B}$ can be written as:

$$C_i = \epsilon_{ijk} A_j B_k. \tag{4.1}$$

Check:

$$\begin{aligned} C_1 &= \epsilon_{1jk} A_j B_k = \sum_{j,k=1}^{3} \epsilon_{1jk} A_j B_k \\ &\quad \text{(must have } j,k \neq 1 \text{ and } j \neq k \text{ by property 2.,} \\ &\quad \text{so either } j=2, k=3 \text{ or } j=3, k=2) \\ &= \epsilon_{123} A_2 B_3 + \epsilon_{132} A_3 B_2 \\ &\quad \text{but } \epsilon_{123} = 1 \text{ by } (II) \text{ above, and } \epsilon_{132} = -1 \text{ by both } (I) \text{ and } (II) \\ &= A_2 B_3 - A_3 B_2 \end{aligned}$$

I leave it as an exercise for you to check other two components C_2 and C_3.

4.4 The very useful ϵ_{ijk} formula

If we want to write out the vector triple product $\underline{A} \times (\underline{B} \times \underline{C})$, for example, we'll need to be able to put these ϵs together. Luckily there is a neat way to do it:

$$\sum_{k=1}^{3} \epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}. \qquad (\dagger) \tag{4.2}$$

Best just to remember this formula!

Check: Let's first consider the left-hand side (LHS) or Equation (4.2).

$$LHS \text{ of Equation (4.2)} = \epsilon_{ij1}\epsilon_{1lm} + \epsilon_{ij2}\epsilon_{2lm} + \epsilon_{ij3}\epsilon_{3lm} \qquad (*)$$

Now $\epsilon_{ij1}\epsilon_{1lm} = 0$ unless

$$\begin{aligned} (i,j) &= (2,3) \text{ or } (3,2) \\ \text{and} \quad (l,m) &= (2,3) \text{ or } (3,2) \\ \Rightarrow \quad \epsilon_{ij1}\epsilon_{1lm} &= \begin{cases} +1 & \text{if} \quad (l,m) = (i,j) \\ -1 & \text{if} \quad (l,m) = (j,i) \end{cases}, \end{aligned}$$

in which case the second and third terms in (*) will be zero.

A similar argument holds for other terms in (*), $\epsilon_{ij2}\epsilon_{2lm}$ and $\epsilon_{ij3}\epsilon_{3lm}$. Therefore the sum in (*) will be zero, except when

$$\begin{aligned} i \neq j, l \neq m, \quad (i,j) = (l,m) \quad &\Rightarrow \quad (*) \ = \ +1 \\ i \neq j, l \neq m, \quad (i,j) = (m,l) \quad &\Rightarrow \quad (*) \ = \ -1. \end{aligned}$$

Let's now consider the right-hand side (RHS) of Equation (4.2).

$$RHS \text{ of Equation (4.2)} \ = \ \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$$

We have $\delta_{il}\delta_{jm} = 1$ if $i = l$ and $j = m \Leftrightarrow (i,j) = (l,m)$, otherwise this term is zero.

Similarly, $-\delta_{im}\delta_{jl} = -1$ if $i = m$ and $j = l \Leftrightarrow (i,j) = (m,l)$, otherwise this term is zero.

The combination of both terms is zero if either $i = j$ or $l = m$.

This implies that LHS of Equation (4.2) = RHS of Equation (4.2), and hence the formula holds.

Example 16. Show that $\underline{A} \times (\underline{B} \times \underline{C}) \ = \ \underline{B}(\underline{A}.\underline{C}) - \underline{C}(\underline{A}.\underline{B})$

To do this using index notation, we should compute the i^{th} component of $\underline{A} \times (\underline{B} \times \underline{C})$. We can write this as $[\underline{A} \times (\underline{B} \times \underline{C})]_i$. Going slowly, this gives

$$\begin{aligned} [\underline{A} \times (\underline{B} \times \underline{C})]_i &= \epsilon_{ijk}A_j\,(\underline{B} \times \underline{C})_k && j,k: \text{ dummy indices, } i: \text{ free index, 1,2,3} \\ &= \epsilon_{ijk}A_j\epsilon_{klm}B_lC_m && l,m: \text{ more dummy indices} \\ &= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})A_jB_lC_m && \text{by useful formula } (\dagger) \\ &= \delta_{il}\delta_{jm}A_jB_lC_m - \delta_{im}\delta_{jl}A_jB_lC_m \\ &= \delta_{il}A_jB_lC_j - \delta_{im}A_jB_jC_m && \text{by the rule for } \delta \\ &= A_jB_iC_j - A_jB_jC_i && \text{by the rule for } \delta \\ &= B_i(A_jC_j) - C_i(A_jB_j) \\ &= B_i(\underline{A}.\underline{C}) - C_i(\underline{A}.\underline{B}) \\ &= [\underline{B}(\underline{A}.\underline{C}) - \underline{C}(\underline{A}.\underline{B})]_i\,, \end{aligned}$$

i.e. i^{th} component of $\underline{A} \times (\underline{B} \times \underline{C}) = i^{\text{th}}$ component of $\underline{B}(\underline{A}.\underline{C}) - \underline{C}(\underline{A}.\underline{B})$. Since this is true for all $i = 1,2,3$ the result follows.

4.5 Applications

We've already seen in example 16 that index notation can be used to prove the vector triple product identity, $\underline{A} \times (\underline{B} \times \underline{C}) \ = \ \underline{B}(\underline{A}.\underline{C}) - \underline{C}(\underline{A}.\underline{B})$.

For many vector calculus calculations we need $\underline{\nabla} f$, $\underline{\nabla}.\underline{v}$ and $\underline{\nabla} \times \underline{v}$ in index notation

- The ith component of $\underline{\nabla} f$ is simply

$$(\underline{\nabla} f)_i \ = \ \frac{\partial f}{\partial x_i}. \tag{4.3}$$

A common notation used to simplify this further is to write $\frac{\partial}{\partial x_i} \equiv \partial_i$, so then we can write

$$(\underline{\nabla} f)_i \ = \ \partial_i f. \tag{4.4}$$

- $\underline{\nabla}.\underline{v}$ can be thought of simply as $\underline{A}.\underline{v}$ where $\underline{A}$ has "components" $\frac{\partial}{\partial x_i}$:

$$\underline{\nabla}.\underline{v} \;=\; \frac{\partial v_i}{\partial x_i}, \tag{4.5}$$

and using the notation from above this can also be written as

$$\underline{\nabla}.\underline{v} \;=\; \partial_i v_i. \tag{4.6}$$

- Similarly, $\underline{\nabla} \times \underline{v}$ is like $\underline{A} \times \underline{v}$. So if $\underline{u} = \underline{\nabla} \times \underline{v}$, then its ith component is:

$$u_i \;=\; (\underline{\nabla} \times \underline{v})_i \;=\; \epsilon_{ijk}\frac{\partial}{\partial x_j}v_k \;=\; \epsilon_{ijk}\frac{\partial v_k}{\partial x_j}, \tag{4.7}$$

and using the same notation as above this can also be written as

$$u_i \;=\; (\underline{\nabla} \times \underline{v})_i \;=\; \epsilon_{ijk}\partial_j v_k. \tag{4.8}$$

Here are some more examples that show how useful index notation can be for proving identities in vector calculus.

Example 17. Find the gradient of $f(\underline{x}) = |\underline{x}|^2$ in $\mathbb{R}^n$.

$$\begin{aligned}
& f(\underline{x}) = |\underline{x}|^2 \;=\; \underline{x}.\underline{x} \;=\; x_i x_i \;=\; x_j x_j \\
\Rightarrow\quad & (\underline{\nabla} f)_i = \frac{\partial}{\partial x_i}(x_j x_j) \qquad \text{important not to reuse the dummy index } j \\
& = \left(\frac{\partial x_j}{\partial x_i}\right) x_j + x_j\left(\frac{\partial x_j}{\partial x_i}\right) \qquad \text{product rule} \\
& = 2\delta_{ij}x_j \;=\; 2x_i \;=\; (2\underline{x})_i \\
\Rightarrow\quad & \underline{\nabla} f = 2\underline{x}
\end{aligned}$$

Example 18. Find the divergence of $\underline{v}(\underline{x}) = \underline{x}$ and $\underline{u}(\underline{x}) = (\underline{a}.\underline{x})\underline{x}$ in $\mathbb{R}^3$.

$$\begin{aligned}
& v_i = x_i \\
\Rightarrow\quad & \underline{\nabla}.\underline{v} = \frac{\partial x_i}{\partial x_i} \;=\; \delta_{ii} \;=\; 3 \\
\text{While}\quad & u_i = (\underline{a}.\underline{x})x_i \;=\; a_j x_j x_i \\
\Rightarrow\quad & \underline{\nabla}.\underline{u} = \frac{\partial}{\partial x_i}(a_j x_j x_i) \\
& = a_j\left(\frac{\partial x_j}{\partial x_i}x_i + x_j\frac{\partial x_i}{\partial x_i}\right) \\
& = a_j(\delta_{ij}x_i + x_j\delta_{ii}) \\
& = a_j(x_j + 3x_j) \;=\; 4a_j x_j \\
& = 4\,\underline{a}.\underline{x}
\end{aligned}$$

Example 19. Find the curl of $\underline{v}(\underline{x}) = \underline{x}$ and $\underline{u}(\underline{x}) = (\underline{a}.\underline{x})\underline{x}$ in $\mathbb{R}^3$.

$$
\begin{aligned}
&& v_i &= x_i \\
\Rightarrow && (\underline{\nabla} \times \underline{v})_i &= \epsilon_{ijk} \frac{\partial}{\partial x_j} x_k \\
&& &= \epsilon_{ijk} \delta_{jk} \\
&& &= \epsilon_{ijj} \\
&& &= \epsilon_{i11} + \epsilon_{i22} + \epsilon_{i33} \\
&& &= 0 \\
&\text{While} & u_k &= a_j x_j x_k \; = \; a_l x_l x_k \\
\Rightarrow && (\underline{\nabla} \times \underline{u})_i &= \epsilon_{ijk} \frac{\partial}{\partial x_j} u_k \\
&& &= \epsilon_{ijk} \frac{\partial}{\partial x_j} (a_l x_l x_k) \\
&& &= a_l \epsilon_{ijk} \left(\frac{\partial x_l}{\partial x_j} x_k + x_l \frac{\partial x_k}{\partial x_j} \right) \\
&& &= a_l \epsilon_{ijk} \left(\delta_{lj} x_k + x_l \delta_{kj} \right) \\
&& &= a_l \epsilon_{ijk} \delta_{lj} x_k + a_l \epsilon_{ijk} x_l \delta_{kj} \\
&& &= a_j \epsilon_{ijk} x_k + a_l \epsilon_{ijj} x_l \\
&& &= \epsilon_{ijk} a_j x_k + 0 \\
&& &= (\underline{a} \times \underline{x})_i \\
\Rightarrow && \underline{\nabla} \times \underline{u} &= \underline{a} \times \underline{x}
\end{aligned}
$$

Example 20. Show that $\underline{\nabla}.(\underline{\nabla} \times \underline{v}) = 0$, (if v_k has continuous 2nd partial derivatives).

$$
\begin{aligned}
\underline{\nabla}.(\underline{\nabla} \times \underline{v}) &= \frac{\partial}{\partial x_i} (\underline{\nabla} \times \underline{v})_i \\
&= \frac{\partial}{\partial x_i} (\epsilon_{ijk} \frac{\partial}{\partial x_j} v_k) \\
&= \epsilon_{ijk} \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} v_k
\end{aligned}
$$

Now ϵ_{ijk} is <u>anti</u>-symmetric when i and j swap, whereas $\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} v_k$ is symmetric, so answer is zero(!).

In more detail:

$$
\begin{aligned}
&& \underline{\nabla}.(\underline{\nabla} \times \underline{v}) &= \epsilon_{ijk} \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} v_k \\
&& &= \epsilon_{jik} \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} v_k && \text{swapping labelling of dummy indices } i \text{ and } j \\
&& &= -\epsilon_{ijk} \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} v_k && \text{anti-symmetric } \epsilon_{ijk} \\
&& &= -\epsilon_{ijk} \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} v_k && \text{symmetric derivatives, due to continuity} \\
&& &= -\underline{\nabla}.(\underline{\nabla} \times \underline{v}) \\
\Rightarrow && \underline{\nabla}.(\underline{\nabla} \times \underline{v}) &= 0
\end{aligned}
$$

Much more elegant (and quicker!) then writing it all out.