Index notation in 3D

1 Why index notation?

Vectors are objects that have properties that are independent of the coordinate system that they are written in. Vector notation is advantageous since it is elegant and deals with vectors as single objects rather than collections of components. This helps to make the physical and geometric meaning of equations manifest. However vector notation has some difficulties, a major one being that there is a whole heap of vector algebraic and differential identities that are hard to remember and hard to derive using vector notation. For example:

\[ \vec{a} \cdot (\vec{b} \times \vec{c}) = \vec{b} \cdot (\vec{c} \times \vec{a}) = \vec{c} \cdot (\vec{a} \times \vec{b}) \] (1.1)

\[ \vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}(\vec{c} \cdot \vec{a}) - \vec{c}(\vec{a} \cdot \vec{b}) \] (1.2)

\[ (\vec{a} \times \vec{b}) \cdot (\vec{c} \times \vec{d}) = (\vec{a} \cdot \vec{c})(\vec{b} \cdot \vec{d}) - (\vec{a} \cdot \vec{d})(\vec{b} \cdot \vec{c}) \] (1.3)

\[ (\vec{a} \times \vec{b}) \times (\vec{c} \times \vec{d}) = (\vec{a} \cdot \vec{c} \times \vec{d})\vec{b} - (\vec{b} \cdot \vec{c} \times \vec{d})\vec{a} \] (1.4)

\[ \nabla \times \nabla f = 0 \] (1.5)

\[ \nabla \cdot (\nabla \times \vec{a}) = 0 \] (1.6)

\[ \nabla \times (\nabla \times \vec{a}) = \nabla(\nabla \cdot \vec{a}) - \nabla^2 \vec{a} \] (1.7)

\[ \nabla \cdot (\alpha \vec{b}) = \alpha \nabla \cdot \vec{b} + \vec{b} \cdot \nabla \alpha \] (1.8)

\[ \nabla \times (\alpha \vec{b}) = \alpha \nabla \times \vec{b} - \vec{b} \times \nabla \alpha \] (1.9)

\[ \nabla (\vec{a} \cdot \vec{b}) = (\vec{a} \cdot \nabla)\vec{b} + (\vec{b} \cdot \nabla)\vec{a} + \vec{a} \times (\nabla \times \vec{b}) + \vec{b} \times (\nabla \times \vec{a}) \] (1.10)

\[ \nabla \times (\vec{a} \times \vec{b}) = \vec{a}(\nabla \cdot \vec{b}) - \vec{b}(\nabla \cdot \vec{a}) + (\vec{b} \cdot \nabla)\vec{a} - (\vec{a} \cdot \nabla)\vec{b} . \] (1.12)

Most of these identities you should have seen before and all of these identities are used throughout physics, e.g. in classical mechanics, quantum mechanics, electrodynamics etc. (A similar notation is also used in general relativity and other more advanced topics). When using index notation to represent the above identities the proofs become very simple. Thus you don’t even need to remember the identities, you can just rederive whichever result you happen to need as you go.

To move from vector notation to an indexed notation we just need to choose an orthonormal basis in which to expand the vectors. Then any vector can be written as a triple in that basis, for example in Cartesian coordinates a vector can be represented as

\[ \vec{v} = (v_x, v_y, v_z) = v_x \hat{x} + v_y \hat{y} + v_z \hat{z} . \]

In general we choose the basis \( \{ \hat{e}_i, \quad i = 1, 2, 3 \} \) and write

\[ \vec{v} = (v_1, v_2, v_3) = v_1 \hat{e}_1 + v_2 \hat{e}_2 + v_3 \hat{e}_3 . \]

Orthonormal means that

\[ \hat{e}_1 \cdot \hat{e}_1 = \hat{e}_2 \cdot \hat{e}_2 = \hat{e}_3 \cdot \hat{e}_3 = 1 \quad \text{and} \quad \hat{e}_1 \cdot \hat{e}_2 = \hat{e}_2 \cdot \hat{e}_3 = \hat{e}_3 \cdot \hat{e}_1 = 0 . \]
The trick to index notation is to write any vector as an indexed symbol: so the vector \( \vec{v} \in \mathbb{R}^3 \) is represented by \( v_i \in \mathbb{R} \) where \( i \) takes the values 1, 2 or 3. Then all of the operations of vector calculus can be written using indexed objects.

## 2 Algebraic relations

### 2.1 Free and dummy indices

Any index that occurs precisely once on each side of an equation is a free index. An equation with a free index represents a vector equation, as it has to be satisfied for each value the index can take. So for example if \( \vec{A} \) and \( \vec{B} \) are vectors then

\[
A_i = B_i \iff \vec{A} = \vec{B} .
\]

For any index that is repeated on one side of the equation, we normally use the Einstein summation convention. That is we sum over all possible values of that index (in 3D that is 1, 2 and 3). For example

\[
A_i B_{ij} = \sum_{i=1}^{3} A_i B_{ij} .
\]

These indices are then called dummy indices because they just stand in for the summation index and can be replaced with any other index without changing the meaning (just like integration variables are dummy variables).

### 2.2 Dot product

To write the dot product in index notation we need the Kronecker \( \delta \) symbol. It is just the identity matrix, 1, written in an indexed form:

\[
\delta_{ij} = \begin{pmatrix}
1 & 0 & \cdots \\
0 & 1 & \cdots \\
\vdots & \ddots & \ddots
\end{pmatrix} .
\]  \hspace{1cm} (2.13)

Then we see that the orthonormality condition on the basis vectors can be written as

\[
\vec{e}_i \cdot \vec{e}_j = \delta_{ij} .
\]  \hspace{1cm} (2.14)

Using the linearity of the dot product we can now see how it works in index notation:

\[
\vec{a} \cdot \vec{b} = (a_i \vec{e}_i) \cdot (b_j \vec{e}_j) = a_i b_j (\vec{e}_i \cdot \vec{e}_j) = a_i \delta_{ij} b_j .
\]  \hspace{1cm} (2.15)

Since \( \delta_{ij} \beta_j = \beta_i \) (which can be seen from the definition above or from the vector equivalent \( 1\vec{b} = \vec{b} \)) we see the dot product take the normal form of the sum of the components:

\[
\vec{a} \cdot \vec{b} = a_i \delta_{ij} b_j = a_i b_i .
\]  \hspace{1cm} (2.16)
The symmetry of the dot product can be seen either from the symmetry of $\delta_{ij}$ combined with the fact that the components $a_i$ and $b_j$ are just numbers and therefore commute.

The magnitude of a vector is just

$$|\vec{a}|^2 = \vec{a} \cdot \vec{a} = a_i a_i.$$  

### 2.3 Cross product

Once again we expand two vectors wrt the basis and use linearity to get

$$\vec{a} \times \vec{b} = a_i b_j (\vec{e}_i \times \vec{e}_j).$$

we know that the cross product of a vector times itself is zero, and using the standard definition of the cross product (right hand rule etc) we see that

$$\vec{e}_1 \times \vec{e}_2 = \vec{e}_3, \quad \vec{e}_2 \times \vec{e}_3 = \vec{e}_1, \quad \vec{e}_3 \times \vec{e}_1 = \vec{e}_2,$$

the others being obtained via antisymmetry of the cross product. This can be summarised by defining a new symbol $\varepsilon_{ijk}$ such that

$$\vec{e}_i \times \vec{e}_j = \varepsilon_{ijk} \vec{e}_k.$$

This symbol is known as the Levi-Civita symbol, it is also called the totally antisymmetric tensor or the permutation symbol. It is defined by:

$$\varepsilon_{ijk} = \begin{cases} 
0 & \text{if any of } i, j \text{ or } k \text{ are the same} \\
1 & \text{if } ijk \text{ is an even permutation of } 123 = \text{sgn } \begin{pmatrix} 1 & 2 & 3 \\ i & j & k \end{pmatrix} \\
-1 & \text{if } ijk \text{ is an odd permutation of } 123.
\end{cases} \quad (2.17)$$

An equivalent definition is

$$\varepsilon_{ijk} = \varepsilon_{jki} = -\varepsilon_{jik}, \quad \varepsilon_{123} = 1. \quad (2.18)$$

We can illustrate this by looking at, for example, the first component. So setting $i = 1$ and expanding the implied summation

$$(\vec{a} \times \vec{b})_1 = \varepsilon_{1jk} a_j b_k = \varepsilon_{123} a_2 b_3 + \varepsilon_{132} a_3 b_2 + \ldots = a_2 b_3 - a_3 b_2,$$

where the dots are terms with repeated indices on the Levi-Civita symbol and thus zero. The other two components can be similarly found. The basic antisymmetry of the cross product can be thought of in terms of the antisymmetry of $\varepsilon_{ijk}$

$$(\vec{a} \times \vec{b})_i = \varepsilon_{ijk} a_j b_k = -\varepsilon_{ikj} a_j b_k = -\varepsilon_{ikj} b_k a_j$$

relabelling the dummy indices to get

$$(\vec{a} \times \vec{b})_i = -\varepsilon_{ijk} b_j a_k = -(\vec{b} \times \vec{a})_i.$$
2.4 Relating $\delta$ and $\varepsilon$

The basic identity that relates these two symbols is

$$\varepsilon_{ijk} \varepsilon_{lmn} = \left| \begin{array}{ccc} \delta_{il} & \delta_{jl} & \delta_{kl} \\ \delta_{im} & \delta_{jm} & \delta_{km} \\ \delta_{in} & \delta_{jn} & \delta_{kn} \end{array} \right| ,$$  \hspace{1cm} (2.19)

but this is over-kill. The only identity that is normally needed is obtained by setting $k = l$ above to get

$$\varepsilon_{ijk} \varepsilon_{kmn} = \delta_{im} \delta_{jn} - \delta_{in} \delta_{jm} .$$  \hspace{1cm} (2.20)

This identity can be derived/checked via brute force (try a few cases) or by arguments based on invariant tensors under spatial rotations. It is also good to note that it respects all of the symmetries that it should, i.e. both sides are antisymmetric under exchange of $i \leftrightarrow j$ or equivalently $m \leftrightarrow n$.

2.5 Scalar triple product and determinants

The scalar triple product is easily seen to be

$$\vec{a} \cdot (\vec{b} \times \vec{c}) = a_i (\varepsilon_{ijk} b_j c_k) = \varepsilon_{ijk} a_i b_j c_k$$  \hspace{1cm} (2.21)

and by cycling the indices

$$\varepsilon_{ijk} a_i b_j c_k = \varepsilon_{kij} c_k a_i b_j = \varepsilon_{jki} b_j c_k a_i ;$$

we get the identities

$$\vec{a} \cdot (\vec{b} \times \vec{c}) = \vec{c} \cdot (\vec{a} \times \vec{b}) = \vec{b} \cdot (\vec{c} \times \vec{a}) .$$  \hspace{1cm} (2.22)

The determinant of a $3 \times 3$ matrix can also be written in indicial form and can be shown to be

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \vec{a} \cdot (\vec{b} \times \vec{c}) = \varepsilon_{ijk} a_i b_j c_k .$$  \hspace{1cm} (2.23)

2.6 Vector triple product

The identity we want to prove is $\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})$. Diving in

$$(\vec{a} \times (\vec{b} \times \vec{c}))_i = \varepsilon_{ijk} a_j \varepsilon_{kmn} b_m c_n = (\delta_{im} \delta_{jn} - \delta_{in} \delta_{jm}) a_j b_m c_n$$

$$= b_n a_j c_j - c_n a_j b_j = (\vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b}))_i ;$$  \hspace{1cm} (2.24)

we see that it’s just a simple consequence of the identity \[2.20\].

2.7 More vector products

Products of more vectors can always be reduced down to repeated applications of the two triple products above, but they can still be computed directly (and efficiently) using index notation.
3 Vector Calculus

The basic element of vector calculus is the $\nabla$ operator. It is a vector operator and is sometimes written as $\vec{\nabla}$. Its defining property is that for any unit vector $\hat{s}$ and function $f(x, y, z) = f(\vec{r})$, $\hat{s} \cdot \nabla f$ is the partial derivative of $f$ in the direction of $\hat{s}$.

Its components, with $\vec{r} = (x, y, z)$,

$$\nabla_i = \frac{\partial}{\partial r_i} = \partial_i$$

take partial derivatives in the $r_i$ direction. Since each component of $\nabla_i$ is just a first order partial derivative, then each component must obey the Leibniz (product) rule. This implies that $\vec{\nabla}$ itself obeys the Leibniz rule. For any scalar functions $f = f(\vec{r})$ and $g = g(\vec{r})$,

$$\nabla_i(fg) = (\nabla_i f) g + f \nabla_i g. \quad (3.25)$$

The familiar gradient, divergence and curl are written as

$$\text{grad}(f) = \nabla f = \nabla_i f, \quad (3.26)$$

$$\text{div}(\vec{v}) = \nabla \cdot \vec{v} = \nabla_i v_i \quad \text{and} \quad (3.27)$$

$$\text{curl}(\vec{v}) = \nabla \times \vec{v} = \varepsilon_{ijk} \nabla_j v_k, \quad (3.28)$$

where we introduced the notational convenience $\vec{v}(\vec{r}) = \vec{v} = v_i$. An important (and obvious) result is

$$\nabla_i r_j = \frac{\partial r_j}{\partial r_i} = \delta_{ij}. \quad (3.29)$$

For any constant vector $\vec{b}$ this immediately leads to

$$\nabla(\vec{r} \cdot \vec{b}) = \nabla_i r_j b_j = \delta_{ij} b_j = b_i = \vec{b}, \quad (3.30)$$

it also implies that curl($\vec{r}$) = 0 (prove me!).

3.1 Five Leibniz type identities

All of these proofs can be done using only the definitions and properties of $\delta_{ij}$ and $\varepsilon_{ijk}$ and the identities $\quad (3.25)$ and $\quad (2.20)$. The first three are trivial:

$$\nabla \cdot (a \vec{b}) = a \nabla \cdot \vec{b} + \vec{b} \cdot \nabla a \quad (3.31)$$

$$\nabla \cdot (\vec{a} \times \vec{b}) = \vec{b} \cdot \nabla \times \vec{a} - \vec{a} \cdot \nabla \times \vec{b} \quad (3.32)$$

$$\nabla \times (a \vec{b}) = a \nabla \times \vec{b} - \vec{b} \times \nabla a. \quad (3.33)$$

To prove

$$\nabla \times (\vec{a} \times \vec{b}) = (\vec{b} \cdot \nabla) \vec{a} - (\vec{a} \cdot \nabla) \vec{b} - \vec{b} (\nabla \cdot \vec{a}) \quad , \quad (3.34)$$

we write

$$LHS = \varepsilon_{ijk} \nabla_j (\varepsilon_{klm} a_l b_m) = (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) (b_m \nabla_j a_l + a_l \nabla_j b_m)$$

$$= b_j \nabla_j a_i + a_i \nabla_j b_j - a_j \nabla_j b_i - b_i \nabla_j a_j = RHS.$$
The proof of the last identity is left as an exercise:

$$\nabla (\vec{a} \cdot \vec{b}) = \vec{a} \times (\nabla \times \vec{b}) + (\vec{a} \cdot \nabla) \vec{b} + \vec{b} \times (\nabla \times \vec{a}) + (\vec{b} \cdot \nabla) \vec{a}$$  \hspace{1cm} (3.35)

Note that this identity has \(3.30\) as a special case.

### 3.2 Repeated derivatives

Assuming we are working in the space of smooth functions, all partial derivatives commute: \(\nabla_i \nabla_j f = \nabla_j \nabla_i f\). Using this, the standard identities are easy to prove:

- \(\text{curl grad } f = \nabla \times (\nabla f) = 0\), \hspace{1cm} (3.36)
- \(\text{div curl } \vec{v} = \nabla \cdot (\nabla \times \vec{v}) = 0\), \hspace{1cm} (3.37)
- \(\text{curl curl } \vec{v} = \nabla \times \nabla \times \vec{v} = \nabla (\nabla \cdot \vec{v}) - \nabla^2 \vec{v} = \text{grad div } \vec{v} - \Delta \vec{v}\). \hspace{1cm} (3.38)

where the Laplacian is defined as \(\Delta = \nabla^2 = \nabla \cdot \nabla\).

### 4 Exercises

Anything that is stated but not derived in these notes!

Also you could try:

- \(\text{div}(\vec{r}) = \nabla \cdot \vec{r} = \ldots\) \hspace{1cm} (4.39)
- \(\text{grad}(|\vec{r}|) = \nabla |\vec{r}| = \ldots\) \hspace{1cm} (4.40)
- \(\nabla f(|\vec{r}|) = \ldots\) \hspace{1cm} (4.41)
- \(\nabla \times \vec{a}(|\vec{r}|) = \ldots\) etc \hspace{1cm} (4.42)

In the Kepler problem of classical mechanics there is one conserved scalar, the energy, and two conserved vectors, the angular momentum \((\hat{L} = \vec{r} \times \vec{p})\) and the Laplace-Runge-Lenz vector. An important component of the LRL vector is \(\vec{p} \times \hat{L}\). Show that it is equal to \(|\vec{p}|^2 \vec{r} - (\vec{p} \cdot \vec{r}) \vec{p}\).

In quantum mechanics the momentum operator is \(\hat{p} = -i \nabla\) so the angular momentum is \(\hat{L} = \vec{r} \times \hat{p} = -i \vec{r} \times \nabla\). Prove that

$$\nabla (\vec{r} \cdot \vec{a}) = \vec{a} + \vec{r} (\nabla \cdot \vec{a}) + i(\hat{L} \times \vec{a})$$.