Einstein's Notation

Kartik Gokhale

April 2021

1 Introduction

The Einstein notation or Einstein summation convention is a notational convention that implies summation over a set of indexed terms in a formula, thus achieving notational brevity. When an index appears twice in a term and is not otherwise defined, it is taken as the sum over all possible values of said index. For example,

$$\Sigma_{i=1}^{3} c_i x^i = c_1 x^1 + c_2 x^2 + c_3 x^3 \tag{1}$$

would just be written as $c_i x^i$. The summation is implicit.
According to this convention:

1. The whole term is summed over any index which appears **twice** in that term
2. Any index appearing only once in the term in not summed over and must be present in both LHS and RHS

2 Two important symbols

2.1 Kronecker Delta

δ_{ij} is the **Kronecker Delta** symbol. It is defined as

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

2.2 Levi-Civita

ϵ_{ijk} is the **Levi-Civita** symbol(in 3 dimensions). It is defined as

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if ijk is a cyclic permutation of xyz, eg xyz, yzx, zxy} \\ -1 & \text{if ijk is an anticyclic permuation of xyz, eg xzy, zyx, yxz} \\ 0 & \text{if in ijk any index is repeating, eg xxy, xyy etc} \end{cases}$$

3 Application to Vectors

3.1 Dot Product

Consider 2 vectors in 3-d space.

$$\vec{a} = a_1\hat{\imath} + a_2\hat{\jmath} + a_3\hat{k}$$
$$\vec{b} = b_1\hat{\imath} + b_2\hat{\jmath} + b_3\hat{k}$$

Hence, their dot product is

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

This can be written quite succinctly using Einstein's notation as

$$\vec{a} \cdot \vec{b} = a_i b_j \delta_{ij} \tag{2}$$

As the summation over the indices is implicit

3.2 Cross Product

Consider 2 vectors again

$$\vec{a} = a_1\hat{\imath} + a_2\hat{\jmath} + a_3\hat{k}$$
$$\vec{b} = b_1\hat{\imath} + b_2\hat{\jmath} + b_3\hat{k}$$

Hence, their cross product is

$$\vec{a} \times \vec{b} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

This can be written quite succinctly using Einstein's notation as

$$\vec{a} \times \vec{b} = a_i b_j \epsilon_{ijk} \hat{e_k} \tag{3}$$

where $\hat{e_k}$ is the unit vector in k direction.

As the summation over the indices is implicit

3.3 Product of Levi-Civitas

We can resolve the product of two Levi-Civitas(in 3-d) as

$$\epsilon_{ijk}\epsilon_{lpq} = \begin{vmatrix} \delta_{il} & \delta_{ip} & \delta_{iq} \\ \delta_{jl} & \delta_{jp} & \delta_{jq} \\ \delta_{kl} & \delta_{kp} & \delta_{kq} \end{vmatrix}$$

4 Problems

4.1 BAC-CAB Rule

Consider 3 vectors, $\vec{A}, \vec{B}, \vec{C}$.Our goal is to simplify

$$\vec{A} \times (\vec{B} \times \vec{C})$$

Solution

Let us define another vector $\vec{D} = \vec{B} \times \vec{C}$

Hence, our expression simplifies to

$$\begin{aligned} \vec{A} \times (\vec{B} \times \vec{C}) &= \vec{A} \times \vec{D} \\ &= a_i d_j \epsilon_{ijk} \hat{e_k} \end{aligned}$$

Since D is a cross product of 2 vectors

$$d_j = b_l c_m \epsilon_{lmj} \hat{e_j}$$

By back-substitution

$$a_i d_j \epsilon_{ijk} \hat{e_k} = a_i b_l c_m \epsilon_{lmj} \epsilon_{ijk} \hat{e_k}$$

Recall product of 2 Levi-Civitas

$$\epsilon_{ijk}\epsilon_{lpq} = \begin{vmatrix} \delta_{il} & \delta_{ip} & \delta_{iq} \\ \delta_{jl} & \delta_{jp} & \delta_{jq} \\ \delta_{kl} & \delta_{kp} & \delta_{kq} \end{vmatrix}$$

Substituting Values and simplifying

$$\begin{aligned} a_i b_l c_m \epsilon_{lmj} \epsilon_{ijk} \hat{e_k} &= a_i b_j c_m (\delta_{lk}\delta_{mi} - \delta_{li}\delta_{mk}) \hat{e_k} \\ &= a_i b_j c_m \delta_{lk}\delta_{mi} \hat{e_k} - a_i b_j c_m \delta_{li}\delta_{mk} \hat{e_k} \end{aligned}$$

Recall these are just dot products

$$\begin{aligned} &= b_k (\vec{A} \cdot \vec{C}) \hat{e_k} - c_k (\vec{A} \cdot \vec{B}) \hat{e_k} \\ &= \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B}) \end{aligned}$$

Hence,we obtain the very popular result

$$\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B}) \tag{4}$$

4.2 More Problems

Question The gradient operator ∇ behaves like a vector in " some sense". For example, divergence of a curl $(\nabla.\nabla \times \vec{A} = 0)$ for any $\vec{A}$, may suggest that it is just like $\vec{A}.\vec{B} \times \vec{C}$ being zero if any two vectors are equal. Prove that $\nabla \times \nabla \times \vec{F} = \nabla(\nabla.\vec{F}) - \nabla^2\vec{F}$. To what extent does this look like the well known expansion of $\vec{A} \times \vec{B} \times \vec{C}$?

Solution

This question is very similar to the previous one.
Now we will represent each component of ∇, $\frac{\partial}{\partial i}$ by ∂_i,

$$\begin{aligned}
\nabla \times \left(\nabla \times \vec{F}\right) &= \epsilon_{ijk}\partial_j \left(\nabla \times \vec{F}\right)_k \hat{e_i} = \epsilon_{ijk}\epsilon_{klm}\partial_j\partial_l F_m\,\hat{e_i} = \epsilon_{kij}\epsilon_{klm}\partial_j\partial_l F_m\,\hat{e_i} \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\partial_j\partial_l F_m\,\hat{e_i} && (\because \epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) \\
&= (\partial_m\partial_i F_m - \partial_l\partial_l F_i)\,\hat{e_i} && (\because \delta_{ij}A_j = A_i) \\
&= \partial_i(\partial_m F_m)\,\hat{e_i} - \partial_l\partial_l(F_i\,\hat{e_i}) && (\because \partial_m\partial_i = \partial_i\partial_m) \\
&= \nabla(\nabla \cdot \vec{F}) - \nabla^2\vec{F} && \blacksquare
\end{aligned}$$

Notice that this proof works because $\partial_m\partial_i = \partial_i\partial_m$ which is just fancy notation for $\frac{\partial^2}{\partial x\partial y} = \frac{\partial^2}{\partial y\partial x}$

We know that $\vec{A} \times \vec{B} \times \vec{C} = \vec{B}(\vec{A} \cdot \vec{C}) - (\vec{A} \cdot \vec{B})\vec{C}$. If we write $\vec{A} = \nabla$, $\vec{B} = \nabla$, $\vec{C} = \vec{F}$, we get $\nabla \times \nabla \times \vec{F} = \nabla(\nabla \cdot \vec{F}) - (\nabla \cdot \nabla)\vec{F} = \nabla(\nabla \cdot \vec{F}) - \nabla^2\vec{F}$. So this serves as a fake proof of the identity

Question As a more involved example, show that the operator $\mathbf{L} = -i\vec{r}\times\nabla$ where $(i = \sqrt{-1})$ satisfies $\mathbf{L}\times\mathbf{L}f = i\mathbf{L}f$ where f is an arbitrary test function. (Notice that the cross product of an operator with itself does not necessarily vanish, can you see why?)

Solution

$$\begin{aligned}
\mathbf{L}\times\mathbf{L}f &= \epsilon_{ijk}L_jL_kf\,\hat{e}_i \\
&= -i\epsilon_{ijk}\epsilon_{jlm}r_l\partial_mL_kf\,\hat{e}_i \\
&= -i\epsilon_{jki}\epsilon_{jlm}r_l\partial_mL_kf\,\hat{e}_i \\
&= -i(\delta_{kl}\delta_{im}-\delta_{km}\delta_{il})r_l\partial_mL_kf\,\hat{e}_i \\
&= -i(r_l\partial_iL_kf - r_i\partial_kL_kf)\,\hat{e}_i \\
&= -i(A_i - B_i)\,\hat{e}_i
\end{aligned}$$

$$A_i = r_k\partial_iL_kf = -i\epsilon_{kpq}r_k\partial_ir_p\partial_qf$$

$$B_i = r_i\partial_kL_kf = -i\epsilon_{kpq}r_i\partial_kr_p\partial_qf$$

Now we can't directly switch the order of ∂_i & ∂_q,

we have to use the product rule: $(\partial_a(r_b\partial_c))f = (r_b\partial_a\partial_c)f + ((\partial_ar_b)\partial_c)f$

$$\implies A_i = -i\cancelto{0}{\epsilon_{kpq}r_kr_p\partial_i\partial_qf} - i\epsilon_{kpq}r_k(\partial_ir_p)\partial_qf = -i\epsilon_{kpq}r_k(\delta_{ip})\partial_qf = -i\epsilon_{kiq}r_k\partial_qf$$

$$(\because \epsilon_{kpq}r_kr_p\partial_i\partial_qf = -\epsilon_{pkq}r_kr_p\partial_i\partial_qf = -\epsilon_{kpq}r_pr_k\partial_i\partial_qf = -\epsilon_{kpq}r_kr_p\partial_i\partial_qf$$

$$\implies \epsilon_{kpq}r_kr_p\partial_i\partial_qf = 0)$$

The first term cancels because in the summation 1 term will have k,q = x,y and another term k,q = y,z. Both differ only in their sign, which is due to the ϵ_{kpq}. So they cancel out

$$\implies B_i = -i\cancelto{0}{\epsilon_{kpq}r_ir_p\partial_k\partial_qf} - i\epsilon_{kpq}r_i(\partial_kr_p)\partial_qf = -i\epsilon_{kpq}r_k(\delta_{kp})\partial_qf = -i\epsilon_{ppq}r_k\partial_qf = 0$$

$$(\because \epsilon_{kpq}r_ir_p\partial_k\partial_qf = -\epsilon_{qpk}r_ir_p\partial_k\partial_qf = -\epsilon_{kpq}r_ir_k\partial_q\partial_kf = -\epsilon_{kpq}r_ir_p\partial_k\partial_qf$$

$$\implies \epsilon_{kpq}r_ir_p\partial_k\partial_qf = 0)$$

Similar reasoning as above is used here

$$\implies \mathbf{L}\times\mathbf{L}f = -\epsilon_{kiq}r_k\partial_qf\,\hat{e}_i = \epsilon_{ikq}r_k\partial_qf\,\hat{e}_i = i\mathbf{L}f \qquad \blacksquare$$

Notice that we have used $\partial_ar_b = \delta_{ab}$. This is because r_a is simply x if $a = x$, y if $a = y$ and so on. And if we differentiate that w.r.t $b = x$ or y or z, then it is 1 if $a = b$, and 0 otherwise, because $\frac{\partial}{\partial x}x = 1$, $\frac{\partial}{\partial x}y = 0$

The cross product of the operator with itself doesn't vanish here because the operators act on different objects. One $\mathbf{L}$ acts on $\mathbf{L}f$, and the other on f.