Einstein summation convention and δ-functions

Contents

1 Summation convention

1.1 Summation convention basics

Einstein summation convention is a convenient notation when manipulating expressions involving vectors, matrices, or tensors in general. (A tensor is a collection of numbers labeled by indices. The *rank* of a tensor is the number of indices required to specify an entry in the tensor, so a vector is a rank–1 tensor, whereas a matrix is a rank–2 tensor.)
The "rules" of summation convention are:

- Each index can appear at most twice in any term.
- Repeated indices are implicitly summed over.
- Each term must contain identical non-repeated indices.

For example,

$$M_{ij}v_j \equiv \sum_i M_{ij}v_j$$

is a valid expression (it is just left-multiplication of vector $\mathbf{v}$ by matrix M), whereas

$$M_{ij}u_j v_j + w_i$$

is not a valid expression in summation convention, since the index j appears three times in the first term, and

$$T_{ijk}u_k + M_{ip}$$

is invalid since the first term contains the non-repeated index j whereas the second term contains p.

The great advantage of using summation convention is that all quantities in an expression become scalars. And we are free to re-order scalars in whatever way we like (though care must be taken when an expression includes operators, see section 1.3).

1.2 The Kronecker-δ and the ϵ-tensor

When writing vector expressions in summation convention, two tensors appear so often that they are conventionally always given the same symbol: the Kronecker-δ and the ϵ-tensor (also called the Levi-Civita symbol, or the anti-symmetric tensor).

The Kronecker-δ is a rank–2 tensor, defined by:

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

In an expression, it has the effect of replacing one index with another (remember the implicit summation):

$$\delta_{ij} u_j = u_i$$

The Kronecker-δ can be thought of as the identity matrix, e.g. in three dimensions

$$\delta_{ij} = \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right]_{ij}.$$

The ϵ-tensor is rank–3, and is defined by:

$$\epsilon_{ijk} = \begin{cases} 1 & ijk \text{ even} \\ -1 & ijk \text{ odd} \\ 0 & \text{otherwise.} \end{cases}$$

The "even" and "odd" refer to the permutation defined by ijk (e.g. 123 is even whereas 132 is odd), and the "otherwise" includes all combinations with repeated indices. Note that the definition implies that we are always free to cyclically permute indices:

$$\epsilon_{ijk} = \epsilon_{kij} = \epsilon_{jki}.$$

On the other hand, swapping any two indices gives a sign-change:

$$\epsilon_{ijk} = -\epsilon_{ikj},$$

which explains why the ϵ-tensor is sometimes called the completely anti-symmetric tensor. The ϵ-tensor is very useful for writing the vector product (or "cross product") in summation convention:

$$[\mathbf{u} \times \mathbf{v}]_i = \epsilon_{ijk} u_j v_k.$$

An important identity connects the ϵ-tensor with the Kronecker-δ:

$$\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}.$$

1.3 Summation convention and operators

Vector operators are easily handled using summation convention. The vector operator ∇ can be thought of as a vector. E.g. in three dimensions

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}.$$

The grad, div and curl operations in summation convention look like:

$$\begin{aligned} [\mathrm{grad}\,\phi]_i &= [\nabla\phi]_i = \nabla_i\phi, \\ [\mathrm{div}\,\mathbf{A}]_i &= [\nabla\cdot\mathbf{A}]_i = \nabla_i A_i, \\ [\mathrm{curl}\,\mathbf{A}]_i &= [\nabla\times\mathbf{A}]_i = \epsilon_{ijk}\nabla_j A_k. \end{aligned}$$

Note that all operators in summation convention become scalar operators.

The only thing to be careful of when dealing with operators in summation convention is that unlike scalars, scalar *operators* can not necessarily be freely re-ordered, since they must always appear to the left of whatever they operate on. For instance

$$\nabla_i A_i \neq A_i \nabla_i.$$

2 δ-functions

In section 1.2 we met the Kronecker-δ, which is zero unless it's two indices are identical. The indices of the Kronecker δ take discreet values (integers). The δ-function (also called the Dirac δ-function) is the continuous analogue: $\delta(x)$ is zero everywhere apart from the origin $x = 0$. The analogy is clearer if we write the function $\delta(x-y)$, which is zero everywhere except at $x = y$.

Formally, the δ-function is defined by

$$\int_{-\infty}^{+\infty} f(x)\delta(x-a)\mathrm{d}x = f(a) \quad \forall f(x),$$

where $f(x)$ is any function.

It can also be expressed as the limit of certain functions. For example

$$\delta(x) = \lim_{\sigma\to 0} e^{-x^2/\sigma} = \lim_{k\to\infty} \frac{\sin(kx)}{kx}.$$

However, when proving fundamental properties of the δ-function, the definition is more useful than these explicit expressions.

The one-dimensional δ-function defined above can easily be extended to higher dimensions in the obvious way. Higher dimensional δ-functions are just products of one-dimensional ones. E.g. in three dimensions,

$$\delta(\mathbf{r}-\mathbf{r_0}) = \delta(x-x_0)\delta(y-y_0)\delta(z-z_0).$$

As already stated, the δ-function is zero everywhere except at the origin, where it diverges. However, as is obvious from the definition, the area under the δ function is finite:

$$\int_{-\infty}^{+\infty} \delta(x)\mathrm{d}x = 1.$$

It has a number of other interesting (and important) properties. Since it is zero everywhere other than the origin, it is also true that

$$\int_p^q f(x)\delta(x-a)\mathrm{d}x = f(a) \qquad \text{for } p < a < q.$$

Rescaling its argument re-scales the δ-function itself:

$$\delta(ax) = \frac{1}{|a|}\delta(x).$$

The derivative of the δ-function is given by

$$\int_p^q \mathrm{d}x f(x) \frac{\mathrm{d}}{\mathrm{d}x}\delta(x-a) = -\left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=a}.$$

Finally, if $f(x)$ has simple zeros at $x = x_1, x_2, \ldots, x_N$, then

$$\delta(f(x)) = \sum_{i=1}^{N} \frac{\delta(x-x_i)}{\left|\frac{\mathrm{d}f}{\mathrm{d}x}(x_i)\right|}.$$

In mechanics, the δ-function $\delta(t-t_0)$ is used to describe an ideal unit impulse (a "kick") at time $t = t_0$. The response function of a system is defined in terms of an ideal unit impulse. (Since there is an analogy between mechanical systems and electric circuits, it is no surprise that all this applies to circuits too). It is used in electromagnetism to describe idealized point-charges: $q\delta(\mathbf{r}-\mathbf{r}_0)$ describes a point-charge q situated at $\mathbf{r_0}$.