EINSTEIN SUMMATION NOTATION

Overview

In class, we began the discussion of how we can write vectors in a more convenient and compact convention. In addition to the advantage of compactness, writing vectors in this way allows us to manipulate vector calculations and prove vector identities in a much more elegant and less laborious manner.

The notation convention we will use, the Einstein summation notation, tells us that whenever we have an expression with a repeated index, we implicitly know to sum over that index from 1 to 3, (or from 1 to N where N is the dimensionality of the space we are investigating).

Vectors in Component Form

You are familiar with writing vectors in component form; the three dimensional vectors **A** and **B** can be expressed as:

$$\mathbf{A} = A_x \hat{x} + A_y \hat{y} + A_z \hat{z}$$
$$\mathbf{B} = B_x \hat{x} + B_y \hat{y} + B_z \hat{z}$$

The dot product of the vectors, **A** and **B**, is:

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z \qquad (1)$$

We see immediately that the result of a dot product is a scalar, and that this resulting scalar is the sum of products. Since the dot product is a sum, we can write this as :

$$\mathbf{A} \cdot \mathbf{B} = \sum_{i=1}^{3} A_i \; B_i \qquad (2)$$

Where i is the arbitrary choice for indexing, and the summation runs from 1 to 3 to capture each of the three components of our vectors.

We can also write the expression in (2) in Einstein summation notation; since we do have a repeated index (in this case the index i), and our expression for a dot product becomes, simply:

$$\mathbf{A} \cdot \mathbf{B} = A_i \; B_i \qquad (3)$$

where summation over i from 1 to 3 is assumed. Learning to write vectors in this notation will make our later work enormously easier.

Our First Vector Proof

Let' s consider our first proof, and we will solve it using standard component notation and then solve it using Einstein summation notation. We will show that :

$$\frac{d}{dt}(\mathbf{A}\cdot\mathbf{B}) = \mathbf{A}\bullet\frac{d\mathbf{B}}{dt} + \mathbf{B}\bullet\frac{d\mathbf{A}}{dt} \tag{4}$$

where the vectors A and B are both functions of time. Using component notation, we write out the dot product of A and B using (1) from above :

$$\mathbf{A}\bullet\mathbf{B} = A_x B_x + A_y B_y + A_z B_z$$

taking the derivative, and using the product rule for differentiation :

$$\frac{d}{dt}(\mathbf{A}\bullet\mathbf{B}) = \frac{d}{dt}\left(A_x B_x + A_y B_y + A_z B_z\right) =$$

$$A_x \frac{dB_x}{dt} + B_x \frac{dA_x}{dt} + A_y \frac{dB_y}{dt} + B_y \frac{dA_y}{dt} + A_z \frac{dB_z}{dt} + B_z \frac{dA_z}{dt}$$

We can group these terms :

$$\left[A_x \frac{dB_x}{dt} + A_y \frac{dB_y}{dt} + A_z \frac{dB_z}{dt}\right] + \left[B_x \frac{dA_x}{dt} + B_y \frac{dA_y}{dt} + B_z \frac{dA_z}{dt}\right] \tag{5}$$

Look carefully at the terms in the brackets and compare with equation (1). Notice that the bracket on the left is just $\boldsymbol{A}\bullet\frac{d\boldsymbol{B}}{dt}$ and that the bracket on the right is $\boldsymbol{B}\bullet\frac{dA}{dt}$, and we have proven our theory using component notation.

In Einstein summation notation, we begin with the definition of the dot product :

$$\mathbf{A}\bullet\mathbf{B} = A_i\, B_i$$

and differentiate using the product rule. Remembering that all of our terms on the right are scalars, we write :

$$\frac{d}{dt}(\mathbf{A}\bullet\mathbf{B}) = \frac{d}{dt}(A_i B_i) = A_i \frac{dB_i}{dt} + B_i \frac{dA_i}{dt} \tag{6}$$

The terms on the right each have a repeated index, so we recognize each term on the right represents a summation over three coordinates. Compare the terms on the right in (6) with the terms in (5), and you will see that the right hand terms in (6) are equivalent to the bracketed terms in (5); in other words, eq. (6) is a one line proof of our identity; all that remains is to equate this to $\frac{d}{dt}(A\cdot B)$.This simple vector proof shows the power of using Einstein summation notation.

The general game plan in using Einstein notation summation in vector manipulations is:

- Write down your identity in standard vector notation;
- "Translate" the vectors into summation notation; this will allow you to work with the scalar components of the vectors;
- Manipulate the scalar components as needed;
- "Translate" the scalar result back into vector form.

In the proof above, point two is accomplished when you write:

$$\frac{d}{dt}(\mathbf{A}\cdot\mathbf{B}) = \frac{\mathbf{d}}{dt}(A_i B_i),$$

point three is accomplished with the differentiation :

$$A_i \frac{dB_i}{dt} + B_i \frac{dA_i}{dt}$$

and the final result occurs on recognition that is the result we wish to prove.

Cross Products and Einstein Summation Notation

In class, we studied that the vector product between two vectors A and B is called the cross product and written as :

$$\mathbf{C} = \mathbf{A} \times \mathbf{B}$$

The resulting vector C has magnitude equal to | A || B | sinθ, and has a direction mutually perpendicular to the vectors A and B. We showed that in component notation, the cross product is :

$$\mathbf{C} = \left(A_y B_z - A_z B_y\right) \hat{\mathbf{x}} + (A_z B_x - A_x B_z) \hat{\mathbf{y}} + \left(A_x B_y - A_y B_x\right) \hat{\mathbf{z}} \qquad (7)$$

The components of the cross product reveal a number of patterns :

•Each component of the vector C is a difference of products.

•Each component has a permutation of terms, that is, each component consists of terms of the form :

$$A_\alpha B_\beta - A_\beta B_\alpha$$

•The $\hat{x}$ component of **C** has no terms involving A_x or B_x; the same condition applies to the $\hat{y}$ and $\hat{z}$ components of **C**.

Before we can write cross products in summation notation, we need to develop a mathematical formalism that will reproduce the patterns we describe above. Just such a formalism exists, and it is called the **Levi-Civita permutation tensor,** or just the permutation tensor. This permuatation tensor can be written for any number of dimensions, but for the most part we will be dealing with three dimensional space.

The permutation tensor is written as ϵ_{ijk} where i, j, and k are indices corresponding to the three coordinate directions. The permutation tensor is defined to have the following values:

$$\epsilon_{ijk} = \begin{cases} 0 \text{ if any two indices are the same} \\ +1 \text{ if all three indices are different and are cyclic} \\ -1 \text{ if all three indices are different and are anti} - \text{cyclic} \end{cases}$$

Cyclic permutations (or even permutations) are 123, 231, and 312; anti - cyclic (or odd) permutations are 132, 213, and 321.

Properties of the Permutation Tensor

Before moving on to how we write cross products with the permutation tensor, let' s investigate some of its properties.

First, how many different ways can we write the tensor in 3 dimensions? The permuation tensor has three indices, ϵ_{ijk} , and there are three ways of choosing i (it could be 1, 2 or 3), three ways of choosing j, and three ways of choosing k. Therefore, there are 27 possible ways of writing ϵ_{ijk} in 3-space.

While tedious, it is instructive to write all of these possible values:

ϵ_{111} ϵ_{211} ϵ_{311}

ϵ_{112} ϵ_{212} ϵ_{312}

ϵ_{113} ϵ_{213} ϵ_{313}

ϵ_{121} ϵ_{221} ϵ_{321}

ϵ_{122} ϵ_{222} ϵ_{322}

ϵ_{123} ϵ_{223} ϵ_{323}

ϵ_{131} ϵ_{231} ϵ_{331}

ϵ_{132} ϵ_{232} ϵ_{332}

ϵ_{133} ϵ_{233} ϵ_{333}

Of these 27 different values, 21 of them are zero, since they have one or more indices the same (refer to the first of the rules for the value of the permutation tensor above). Of the remaining six possible values, three of them are cyclic permutations (noted in blue) and have value + 1, and six are anti - cyclic permutations of value - 1 (noted in red).

Writing Cross Products with the Permutation Tensor

Let' s start by rewriting eq. (7) from above :

$$\mathbf{C} = \left(A_y\, B_z - A_z\, B_y\right) \hat{\mathbf{x}} + (A_z\, B_x - A_x\, B_z)\, \hat{\mathbf{y}} + \left(A_x\, B_y - A_y\, B_x\right) \hat{\mathbf{z}}$$

We will show that any component of the vector C, produced from the cross product of vectors A and B, can be written in the form :

$$C_i = \epsilon_{ijk}\, A_j\, B_k \qquad (8)$$

This means that the i^{th} component of C (where i can be 1, 2, or 3) can be expressed in terms of the permutation tensor. Let's go into some detail to see how this works. For specificity, let's say we want to find the x component of **C**, in other words, we are going to set i=1 (since i=1 corresponds to the x component, i=2 corresponds to the y component, i=3 to the z component).

If we set i=1, then we know that the values of j and k must be 2 or 3, since the permutation tensor is zero if there are any repeated indices. Looking at (8) above, we can see that both j and k are repeated on the right hand side, so we know we will sum over j and k. (Since we have set i=1 already, I will not show explicitly the summation for j or k =1). Thus, with i=1, our implied summation in (8) is:

$$C_1 = \epsilon_{122} A_2 B_2 + \epsilon_{123} A_2 B_3 + \epsilon_{132} A_3 B_2 + \epsilon_{133} A_3 B_3 \tag{9}$$

The first and last terms on the right hand side of (9) are zero because there are duplicated indices; the terms highlighted in blue and red are the only non zero terms. The term highlighted in blue has a value of + 1 and the value highlighted in red has a value of - 1, so that equation (9) tells us that the x component of C is :

$$C_1 = A_2 B_3 - A_3 B_2 \tag{10}$$

Compare this to the form of eq. (7) and you will see this is exactly equivalent to the expression we obtained from a term by term multiplication of components.

What happens if we set i=2? We know that j and k will be either 1 or 3, and the only non zero terms that survive the summation are:

$$C_2 = \epsilon_{213} A_1 B_3 + \epsilon_{231} A_3 B_1 = A_3 B_1 - A_1 B_3 \tag{11}$$

Similarly, if i = 3, we find the z component of C as :

$$C_3 = \epsilon_{321} A_2 B_1 + \epsilon_{312} A_1 B_2 = A_1 B_2 - A_2 B_1 \tag{12}$$

And we have accurately reproduced each of the individual components of the vector C. A simple final step is to combine all these components to yield the complete vector, so we need to sum all the components as :

$$C_i e_i = \epsilon_{ijk} A_j B_k e_i \tag{13}$$

where we make our first use of "e_i" to mean the i^{th} unit vector; throughout the semester we will make use of $\hat{e}$ to refer to a unit vector in some direction; if we expand the left side of eq. (12) in accordance with our understanding of Einstein summation, we get:

$$C_i e_i = C_1 \hat{e}_1 + C_2 \hat{e}_2 + C_3 \hat{e}_3 = C_1 \hat{x} + C_2 \hat{y} + C_3 \hat{z} \tag{14}$$

As a quick summation, the i^{th} component of the cross product resulting from ***C* = AxB**is:

$$C_i = \epsilon_{ijk} A_j B_k \tag{15}$$

and the complete vector C is :

$$C_i e_i = \epsilon_{ijk} A_j B_k e_i \tag{16}$$

You should not be confused by the slight differences between (15) and (16); (15) tells you the value of any individual component; (16) is the complete vector, and is computed by adding up all the individual components.

A Proof Using Dot and Cross Products

We know that the cross product of two vectors produces a third vector that is perpendicular to each of the original vectors. In other words, if **C=AxB**, we expect that **A•C** and **B•C** should be zero. You can probably imagine how many steps this takes in component form, let's prove this using summation notation. We know the i^{th} component of the vector **C** can be expressed as:

$$C_i = \epsilon_{ijk} A_j B_k \tag{17}$$

Taking the dot product of A and C is equivalent to finding $A_i C_i$, so we can write:

$$A \bullet C = A_i C_i = A_i \epsilon_{ijk} A_j B_k \tag{18}$$

All the terms on the right hand side of (18) are scalars, so we can reorder them in any way we wish since multiplication of scalars is commutative. In particular, we can rewrite (18) as :

$$A_i \epsilon_{ijk} A_j B_k = B_k \epsilon_{ijk} A_i A_j \tag{19}$$

The final term on the right is the k^{th} component of the cross product of **AxA**. Since the cross product of any vector with itself is zero, we have shown that **A•C** is zero. An exactly similar analysis will show that **B•C** is zero.