Math Handouts: Index Notation

This handout is an introduction to **index notation**, a powerful and essential tool for any practicing (or aspiring!) physicist. In summary, index notation, or the (Einstein) summation convention, is the practice of *implicitly* summing over terms with *twice-repeated* indices. This is both more and less complicated than it seems.

As far as prerequisites go, you should know the basics of multivariable calculus, specifically partial derivatives, div, grad, curl, and all that. Integration would be nice too, but it's not necessary.

Problems marked with "**Essential**" *must* be thoroughly understood to say that you know the basics of this subject. The point system is just for checking how much you understand; most questions are worth 1 point, but harder ones (which I'll mark as such) will be worth $2-4$ points. Aim to get $\geq 80\%$ of the available points. There are **53** points.

There's also a more advanced section at the end, mainly for the interest of the readers; it covers index notation and some tensor manipulation in more general contexts, with the intention to prepare you for subjects like quantum field theory and general relativity.

(**fix introduction**)

1 The Basics

Idea 1.1 (Vectors and the Summation Convention)

If we choose the standard Cartesian basis $\{\mathbf{e}_i\}_{i=1}^3$, we may decompose a 3-vector $\mathbf{v}$ into

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3 = \sum_{a=1}^{3} v_a\mathbf{e}_a,$$

therefore expressing $\mathbf{v}$ as a sum of its components. In practice, it becomes very useful to adopt a convention where any *twice-repeated index*[a] is *implicitly* summed over. This is the exact same as writing the explicit $\sum$ sign, but we are now dropping it for simplicity's sake. It looks like

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3 = \sum_{a=1}^{3} v_a\mathbf{e}_a =: v_a\mathbf{e}_a.$$

[a] Not a once-repeated nor a thrice-repeated index counts; only **twice**.

Idea 1.2 (Kronecker Delta)

We now define one of the main players of index notation: the **Kronecker Delta**. Consider two orthonormal vectors, $\mathbf{e}_a$ and $\mathbf{e}_b$ (where $a, b = 1, 2, 3$). If we take the dot product of these vectors, the following relation holds,

$$\mathbf{e}_a \cdot \mathbf{e}_b = \begin{cases} 1 & a = b \\ 0 & a \neq b. \end{cases}$$

This is true by definition, but you can convince yourself of it with a few examples. Let $\{\mathbf{e}_i\}_{i=1}^3$

be the Cartesian basis (i.e. the unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$). We would have

$$0 = \mathbf{e}_1 \cdot \mathbf{e}_2 \qquad\qquad 1 = \mathbf{e}_1 \cdot \mathbf{e}_1$$
$$0 = \mathbf{e}_2 \cdot \mathbf{e}_3 \qquad\qquad 1 = \mathbf{e}_3 \cdot \mathbf{e}_3.$$

We now define the *Kronecker Delta* as follows.

$$\mathbf{e}_a \cdot \mathbf{e}_b = \delta_{ab} := \begin{cases} 1 & a = b \\ 0 & a \neq b. \end{cases}$$

Remark 1.1 (Selecting for components)**.** With these definitions, it should be clear that we can *select* for the a^{th} component of some vector $\mathbf{v}$ by taking its dot product with a basis vector $\mathbf{e}_a$. In writing, this looks like,

$$v_a = \mathbf{v} \cdot \mathbf{e}_a.$$

Another thing to take note of is that the Kronecker Delta is symmetric, i.e.

$$\delta_{ab} = \delta_{ba}.$$

You can easily show that this is true from the definition (exercise!).

Remark 1.2 (How do we know when to stop?)**.** With implicit summation, you may have the concern that we won't know where to *stop* our summation; after all, we don't have a convenient upper-index to tell us where to stop counting. This actually isn't a problem in practice, as context makes it clear as to what we're summing over. The standard convention is $a = 1, 2, 3$, and in relativistic contexts, $\mu = 0, 1, 2, 3$ is standard. A word on notation: Latin indices like i, j, k or a, b, c are usually used in non-relativistic contexts, and Greek indices like μ, ν, and ρ are usually used in relativistic contexts. The authors will tell you what the upper bound on their sum is, and if they don't, it'll be abundantly clear from context[a].

[a] For example, if you're talking about special relativity, it should be obvious that $\mu = 0, \ldots, 3$.

Remark 1.3 (Dummy vs. free indices)**.** It is important to distinguish between the two types of indices there are; there are **dummy indices**, and **free indices**. A dummy index is one that appears twice; these are the ones we're familiar with. The reason these are called dummy indices is because they are completely arbitrary; we could have just as easily written the vector $\mathbf{v}$ as

$$\mathbf{v} = v_\mu \mathbf{e}_\mu = v_i \mathbf{e}_i = v_a \mathbf{e}_a = v_\heartsuit \mathbf{e}_\heartsuit.$$

A free index is one that appears only once. An index like this just tells you that your final equation is one that deals with components instead of the entire object at once; an example of this is

$$v_a = \mathbf{v} \cdot \mathbf{e}_a.$$

As you can see, there is one free index in this expression. Thus, this equation is talking about the *components* of $\mathbf{a}$, as opposed to the "full thing". Here are two **vitally** facts to keep in mind:

- Free indices *must* balance over all terms in an equation. For example, we could have $\mathbf{a}+\mathbf{b}=\mathbf{c} \iff a_a+b_a=c_a$. As an non-example, the following expression is illegal: $a_a+b_b=c_a$. This is because the free indices a and b are not shared by every term in the equation.

- As stated previously, summation is implied over **twice-repeated** indices. You cannot have three identical indices appear in the same term; this is illegal[a]. What qualifies as the "same term"? Roughly, something that multiplies something else, e.g. in a_a+b_a, a_a and b_a are separate terms, but in $a_ab_a+c_bd_b$, a_ab_a and c_bd_b are the separate terms.

In practice, these facts become quite useful for catching errors and sillies. If one has a free index on the left-hand side of an equation that isn't shared with each term on the right-hand side, and vice versa, you know you've made an error. Similarly, if you see a term with a thrice-repeated index in your equation, you also know you have messed up. As a side note, in this handout, the indices have **no relation** to the vectors; an expression like a_a isn't any more special than b_a.

[a] If you think is weird or arbitrary, try to do this, but input the summation symbols explictly. You will see things break quickly.

[6] **Problem 1.** (**Essential**)
Show the following identities are true:

(a) $\mathbf{x}\cdot\mathbf{x}=x_ax_a$.

(b) $a_a=a_b\delta_{ab}$ (this is an example of what's called *index contraction*).

(c) $\mathbf{y}\cdot\mathbf{x}=y_ax_a$ (Hint: use previous part).

(d) $\delta_{ab}\delta_{bc}=\delta_{ac}$ (this is another example of index contraction).

(e) $\partial_ax_b=\delta_{ab}$.

(f) $\delta_{ab}\partial_a=\partial_b$ (ditto last comment).

(g) $\delta_{aa}=3$.

Note: you *must* understand the above examples. These will become like breathing to you, so you must understand why they work. An analogy here is basic operations in solving algebraic equations; in the equation $5+3x=8$, you do not consciously think about subtracting 5 from both sides and then dividing by 3 to get $x=1$. The thing is, though, is that you only *got* to that point by practicing the "slow way" enough times that you could do it without thinking. With some practice, this will occur with index notation as well; to get there, though, you need to understand the basics, i.e. these problems, and some of the other problems later in this handout. The solutions to the essential problems will be more detailed than the other solutions to help aid this understanding.

Solution. (a) $\mathbf{x}\cdot\mathbf{x}=x_a\mathbf{e}_a\cdot x_a\mathbf{e}_a=x_ax_a(\mathbf{e}_a\cdot\mathbf{e}_a)=x_ax_b(1)=x_ax_a$. The reason you can factor out the x_a terms from the dot product is because they're scalars; an example of this is $3\mathbf{x}\cdot 5\mathbf{x}=15(\mathbf{x}\cdot\mathbf{x})$.

(b) We can see that the RHS is going to equate to zero if $b \neq a$, so the only non-zero entries are going to be when $a = b$, so when $a_a = a_a(1)$. The result follows.

(c) $\mathbf{y} \cdot \mathbf{x} = y_a \mathbf{e}_a \cdot x_b \mathbf{e}_b = y_a x_b (\mathbf{e}_a \cdot \mathbf{e}_b) = y_a x_b \delta_{ab} = y_a x_a$ (note that I used the index contraction of the previous step).

(d) You can see this is true, as for when $a \neq b$ or $b \neq c$, the expression $\delta_{ab}\delta_{bc}$ will equate to zero, which leads to a new "effective" expression of δ_{ac} (where the only non-zero components come from when $a = b = c$). If you really wanted to prove this explicitly, feel free to try every combination of a, b, and c, and prove that the expressions are equivalent via truth table.

(e) You can see that this is true by noticing that $\partial x_b / \partial x_a = 0$ unless $a = b$, in which case it equals 1 (convince yourself of this fact by plugging in a few examples). This is a logically equivalent statement to δ_{ab}.

(f) Again, by identical logic as last time, $\delta_{ab}\partial_a = 0$ if $a \neq b$. The result follows.

(g) Summing from $a = 1, \ldots, 3$ gives us this identity,

$$\delta_{11} + \delta_{22} + \delta_{33} = 1 + 1 + 1 = 3.$$

Idea 1.3 (The Gradient and Vector-Valued Functions)

We define the gradient as follows,

$$\nabla \coloneqq \mathbf{e}_a \partial_a,$$

where we've used the shorthand $\partial_a \coloneqq \partial / \partial x_a$. We also define a vector-valued function as

$$\mathbf{g}(\mathbf{x}) \coloneqq \mathbf{e}_a g_a(\mathbf{x}).$$

A scalar function obviously cannot have indices, as it has no components to sum over.

Example 1.1 (Deriving the inside cover of Griffiths, pt. 1)

Let's derive two essential identities from the inside cover of Griffiths, and in doing so, illustrate the general process of deriving such things.

$$\nabla(fg) = (\nabla f)g + f(\nabla g) \qquad\qquad \nabla \cdot (f\mathbf{A}) = \nabla f \cdot \mathbf{A} + f(\nabla \cdot \mathbf{A}).$$

The first thing we would like to do is put things in terms of indices; this just means replacing the terms in our equation with the equivalent index definitions. For the above examples, we have

$$\nabla(fg) = \mathbf{e}_a \partial_a (fg) \qquad\qquad \nabla \cdot (f\mathbf{A}) = \mathbf{e}_a \partial_a \cdot (f \mathbf{e}_b A_b).$$

We now carry out whatever operations seem obvious at the moment. For the first term, we simply apply ∂_a to fg using the product rule. For the second term, we first express the dot product in terms of the Kronecker delta, and then apply the product rule to fA_b. Finally, we "un-indexify" our expression by reinstating the relevant vector and operator terms. Applying this process to the terms above gives

$$\nabla(fg) = \mathbf{e}_a\partial_a(fg) = \mathbf{e}_a\{(\partial_a f)g + f(\partial_a g)\} = \boxed{(\nabla f)g + f(\nabla g),}$$
$$\begin{aligned}\nabla\cdot(f\mathbf{A}) &= \mathbf{e}_a\partial_a\cdot(\mathbf{e}_b f A_b) = \delta_{ab}\partial_a(fA_b) = \partial_a(fA_a) = (\partial_a f)A_a + f(\partial_a A_a)\\ &= \boxed{(\nabla f)\cdot\mathbf{A} + f(\nabla\cdot\mathbf{A}).}\end{aligned}$$

[10] **Problem 2.** Verify the following results:

(a) $\nabla\cdot\mathbf{g}(\mathbf{x}) = \partial_a g_a(\mathbf{x})$

(b) $(\mathbf{x}\cdot\nabla)f(\mathbf{x}) = x_b\partial_b f(\mathbf{x})$.

(c) $\nabla\cdot(\nabla f(\mathbf{x})) = \nabla^2 f(\mathbf{x})$ where $\nabla^2 \coloneqq \partial_x^2 + \partial_y^2 + \partial_z^2$.

(d) $\nabla\cdot\mathbf{x} = \partial_a x_a = \delta_{aa} = 3$.

(e) $\nabla(\mathbf{x}\cdot\mathbf{x}) = 2\mathbf{x}$.

(f) $\nabla^2(\mathbf{x}\cdot\mathbf{x}) = 6$.

(g) (2pts) $\nabla|\mathbf{x}| = \mathbf{x}/|\mathbf{x}|$.

(h) (3pts) $\partial_a(x_b/|\mathbf{x}|) = (x^2\delta_{ab} - x_a x_b)/x^3$.

Solution. (a) $\nabla\cdot\mathbf{g}(\mathbf{x}) = \mathbf{e}_a\partial_a\cdot\mathbf{e}_b g_b(\mathbf{x}) = \delta_{ab}\partial_a g_b(\mathbf{x}) = \partial_a g_a(\mathbf{x})$.

(b) $(\mathbf{x}\cdot\nabla)f(\mathbf{x}) = (\mathbf{e}_a x_a\cdot\mathbf{e}_b\partial_b)f(\mathbf{x}) = (\delta_{ab}x_a\partial_b)f(x) = x_b\partial_b f(\mathbf{x})$.

(c) $\nabla\cdot(\nabla f(\mathbf{x})) = \mathbf{e}_a\partial_a\cdot(\mathbf{e}_b\partial_b f(\mathbf{x})) = \delta_{ab}\partial_a\partial_b f(\mathbf{x}) = \partial_a\partial_a f(\mathbf{x}) = \nabla^2 f(\mathbf{x})$.

(d) $\nabla\cdot\mathbf{x} = \mathbf{e}_a\partial_a\cdot\mathbf{e}_b x_b = \delta_{ab}\partial_a x_b = \delta_{ab}\delta_{ab} = \delta_{aa} = 3$.

(e) $\nabla(\mathbf{x}\cdot\mathbf{x}) = \mathbf{e}_a\partial_a(\mathbf{e}_b x_b\cdot\mathbf{e}_b x_b) = \mathbf{e}_a\partial_a(x_b x_b) = \mathbf{e}_a(\partial_a x_b)x_b + x_b(\partial_a x_b) = 2\mathbf{e}_a\delta_{ab}x_b = 2\mathbf{e}_a x_a = 2\mathbf{x}$.

(f) $\nabla^2(\mathbf{x}\cdot\mathbf{x}) = \nabla\cdot(\nabla(\mathbf{x}\cdot\mathbf{x})) = \nabla\cdot(2\mathbf{x}) = \partial_a 2x_b = 2\delta_{ab} = 2\delta_{11} + 2\delta_{22} + 2\delta_{33} = 6$. Note that in the last step, when we summed over the entries of δ_{ab}, we didn't bother to write any of the cases where $a \neq b$, as they are zero.

(g) These last two are slightly trickier; the key to the problem is realizing that $|\mathbf{x}| = \sqrt{x_a x_a}$.

$$\nabla|\mathbf{x}| = \mathbf{e}_a\partial_a\sqrt{x_b x_b} = \mathbf{e}_a\frac{\delta_{ab}x_b}{\sqrt{x_b x_b}} = \frac{\mathbf{e}_b x_b}{\sqrt{x_b x_b}} = \frac{\mathbf{x}}{|\mathbf{x}|}.$$

(h) We have

$$\partial_a\left(\frac{x_b}{|\mathbf{x}|}\right) = \partial_a\left(\frac{x_b}{\sqrt{x_c x_c}}\right) = \frac{\sqrt{x_c x_c}\,\delta_{ab} - x_b\frac{\delta_{ac}x_c}{\sqrt{x_c x_c}}}{x_c x_c} = \frac{x_c x_c\delta_{ab} - x_a x_b}{(x_c x_c)^{3/2}} = \frac{x^2\delta_{ab} - x_a x_b}{x^3}.$$

Note that we multiplied through by $\sqrt{x_c x_c}$ in the fourth term.

[6] **Problem 3.** Verify the following results. All of the questions below are worth 2pts.

(a) $\nabla(\mathbf{x}\cdot\mathbf{g}(\mathbf{x})) = \mathbf{g}(\mathbf{x}) + \mathbf{e}_a x_b \partial_a g_b(\mathbf{x})$.

(b) $\nabla\cdot(\mathbf{x}f(\mathbf{x})) = 3f(\mathbf{x}) + (\mathbf{x}\cdot\nabla)f(\mathbf{x})$.

(c) $\nabla e^{i\mathbf{k}\cdot\mathbf{x}} = i\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{x}}$.

Solution. (a) $\nabla(\mathbf{x}\cdot\mathbf{g}(\mathbf{x})) = \mathbf{e}_a\partial_a(x_b g_b(\mathbf{x})) = \mathbf{e}_a(\delta_{ab}g_b(\mathbf{x}) + x_b\partial_a g_b(\mathbf{x})) = \mathbf{g}(\mathbf{x}) + \mathbf{e}_a x_b \partial_a g_b(\mathbf{x})$.

(b) $\nabla\cdot(\mathbf{x}f(\mathbf{x})) = \partial_a(x_a f(\mathbf{x})) = \delta_{aa}f(\mathbf{x}) + x_a\partial_a f(\mathbf{x}) = 3f(\mathbf{x}) + (\mathbf{x}\cdot\nabla)f(\mathbf{x})$.

(c) $\nabla e^{i\mathbf{k}\cdot\mathbf{x}} = \mathbf{e}_a\partial_a e^{ik_b x_b} = i\mathbf{e}_a\delta_{ab}k_b e^{ik_b x_b} = i\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{x}}$. If you're confused about this, note that $\partial_a k_b = 0$.

Example 1.2 (A foray into Fourier analysis)

(fix ts; proof is ass and was written like a year ago right after I first learned it)

Here's an interesting identity:

$$f(\mathbf{x}+\mathbf{a}) = f(\mathbf{x}) + (\mathbf{a}\cdot\nabla)f(\mathbf{x}) + \frac{(\mathbf{a}\cdot\nabla)^2}{2}f(\mathbf{x}) + \cdots = e^{\mathbf{a}\cdot\nabla}f(\mathbf{x}).$$

I distinctly remember being floored in lecture when I saw my professor off-handedly write this on the board. I'll show two ways to see why this is true. The first is a Taylor expansion. Recall that the exponential is *defined as*

$$e^x := \sum_{n=0}^{\infty}\frac{x^n}{n!}.$$

What if we replaced x with an operator? Recall that an operator is just something that takes vectors to other vectors[a]. Replacing x with an operator, which we'll call A, is a perfectly legal thing to do; if this bothers you, just consider e^A to be convenient shorthand for the infinite sum

$$e^A := \sum_{n=0}^{\infty}\frac{(A)^n}{n!},$$

where $A^0 := I$. We know that $\mathbf{a}\cdot\nabla = a_a\partial_a$ is an operator, so if we let $A = \mathbf{a}\cdot\nabla$, we get

$$e^{\mathbf{a}\cdot\nabla} = \sum_{i=0}^{\infty}\frac{(\mathbf{a}\cdot\nabla)^n}{n!},$$

which is a new operator. Ok, now that we have that established, let's go focus on the left-hand side of the equation for a bit. The multivariable Taylor expansion is formally defined as (in components, which you now understand!)

$$f(\mathbf{x}+\mathbf{a}) = f(\mathbf{x}) + a_a\partial_a f(\mathbf{x}) + \frac{a_a a_b\partial_a\partial_b}{2}f(\mathbf{x}) + \cdots. \tag{1}$$

Let's relate this to what we had before. $a_a\partial_a$ is clearly a dot product, $\mathbf{a}\cdot\nabla$, and we can consider $a_a a_b\partial_a\partial_b$ the repeated application of two $\mathbf{a}\cdot\nabla$ terms, so $(\mathbf{a}\cdot\nabla)^2$ is equivalent to the second term in the expression[b]. Ok, so we must have

$$f(\mathbf{x}) + (\mathbf{a}\cdot\nabla)f(\mathbf{x}) + \frac{(\mathbf{a}\cdot\nabla)^2}{2}f(\mathbf{x}) + \cdots = f(\mathbf{x}) + a_a\partial_a f(\mathbf{x}) + \frac{a_a a_b \partial_a \partial_b}{2} f(\mathbf{x}) + \cdots .$$

But we know that (1) is equivalent to $f(\mathbf{x}+\mathbf{a})$, so

$$f(\mathbf{x}+\mathbf{a}) = f(\mathbf{x}) + (\mathbf{a}\cdot\nabla)f(\mathbf{x}) + \frac{(\mathbf{a}\cdot\nabla)^2}{2}f(\mathbf{x}) + \cdots .$$

Writing out $e^{\mathbf{a}\cdot\nabla}f(\mathbf{x})$ from before gives (using the series definition of $e^{\mathbf{a}\cdot\nabla}$)

$$e^{\mathbf{a}\cdot\nabla}f(\mathbf{x}) = f(\mathbf{x}) + (\mathbf{a}\cdot\nabla)f(\mathbf{x}) + \frac{(\mathbf{a}\cdot\nabla)^2}{2}f(\mathbf{x}) + \cdots = f(\mathbf{x}) + a_a\partial_a f(\mathbf{x}) + \frac{a_a a_b \partial_a \partial_b}{2} f(\mathbf{x}) + \cdots .$$

And thus $e^{\mathbf{a}\cdot\nabla}f(\mathbf{x}) = f(\mathbf{x}+\mathbf{a})$. Going a bit deeper, we call the operator $e^{\mathbf{a}\cdot\nabla}$ the **generator of translations**[c]; this terminology is borrowed from the mathematics of **Lie groups**, the study of "continuous groups". This rich subject is essential in many areas of modern physics, and it can be seen in little ideas like this. There is another way to show this identity through **Fourier analysis**. Here is a crash course on Fourier analysis. We define the **Fourier transform** of a function $f(\mathbf{x})$ to be (the boxed one; the non-boxed one is called the **inverse Fourier transform**)

$$\boxed{f(\mathbf{x}) = \int đ^3 q\, e^{i\mathbf{q}\cdot\mathbf{x}} \hat{f}(\mathbf{q})} \iff \hat{f}(\mathbf{q}) = \int d^3 x\, e^{-i\mathbf{q}\cdot\mathbf{x}} f(\mathbf{x}),$$

where $đq := dq/2\pi$ (you pronounce $đ$ as "d-bar", similar to hbar in quantum mechanics). (**introduce $đ\mathbf{q} = đ^3 q$ to get them used to good qft notation**) A Fourier transform is just expressing some given function $f(\mathbf{x})$ as the sum of plane waves in "Fourier space", some other space usually consisting of either frequency or wave number. Considering the problem from before gives

$$f(\mathbf{x}+\mathbf{a}) = \int đ^3 q\, e^{i\mathbf{q}\cdot(\mathbf{x}+\mathbf{a})} \hat{f}(\mathbf{q}) = \textbf{(finish)}$$

[a] For the more mathematical people out there, an operator is an **automorphism**, i.e. a morphism $\phi\colon V \to W$ where $V = W$. We call the set of automorphisms on some G-module V, $\mathrm{Aut}(V)$.

[b] Because we *may not* repeat indices more than twice in a term, $a_a a_a \partial_a \partial_a$ is *not allowed* and *makes no sense*.

[c] If you've seen some quantum mechanics, you may think that $e^{\mathbf{a}\cdot\nabla}$ looks awfully familiar to $e^{-iH(t-t_0)/\hbar}$; this isn't a coincidence. The Hamiltonian is the generator of *time* translations in quantum mechanics, so exponentiating it naturally gives finite translations in time. For more about the math behind this, see Georgi's *Lie Algebras in Particle Physics*.

Idea 1.4 (The Levi-Civita Symbol)

We define the Levi-Civita symbol as,

$$\epsilon_{abc} = \begin{cases} \pm 1 & a,\, b,\, c \text{ distinct} \\ 0 & \text{else.} \end{cases}$$

We also define $\epsilon_{123} = 1$, and say that the symbol is completely anti-symmetric; this means that it takes on the opposite sign whenever you flip a pair of indices, e.g. $\epsilon_{123} = -\epsilon_{213}$, where we swapped 1 and 2. Thus you can determine the sign of any configuration of the epsilon symbol by just getting it back to ϵ_{123} and counting the number of swaps it took: an even number means that it's 1 and an odd number means that it's -1. Consider an orthonormal, "right-handed" Cartesian basis. "Right-handed" just means that the vectors obey the cross product, e.g. $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$. In terms of the epsilon symbol, we may more compactly express this relationship,

$$\boxed{\mathbf{e}_a \times \mathbf{e}_b = \mathbf{e}_c \epsilon_{abc}.}$$

Note that both a and b are free indices and c is a repeated index, so summation is only implied over c.

Remark 1.4 (Another way of looking at things)**.** We can define the epsilon symbol in a different way, although one that I think is more confusing. I'll present it here for posterity, and just in case it jives with you a bit more.

$$\epsilon_{abc} = \begin{cases} 1 & \text{cyc. perm. of } (1,2,3) \\ -1 & \text{cyc. perm. of } (3,2,1) \\ 0 & \text{else.} \end{cases}$$

Notice that "else" just means there are repeated indices. Cyc. perm. is an abbreviation of *cyclic permutation*, which is the operator that shifts the numbers left/right one. Here are all of the cyclic permutations of $(1, 2, 3)$ and $(3, 2, 1)$.

$$\text{perm}(1,2,3) = \{(1,2,3),(3,1,2),(2,3,1)\}$$
$$\text{perm}(3,2,1) = \{(3,2,1),(1,3,2),(2,1,3)\}.$$

I personally don't like this as much because it requires you to remember stuff, but to each their own; use whichever one you like more. I will say, though, that when you get into GR or QFT, it will be more difficult to take this approach with $\epsilon^{\mu\nu\rho\sigma}$, the four-dimensional Levi-Civita symbol, so I'd recommend using the approach given in the idea.

Remark 1.5 (What's the difference between "symbol" and "tensor"?)**.** This remark is for more mathematically mature students, so if that's not you, feel free to either ignore this or read it for fun. (**finish**)

Example 1.3 (Deriving the inside cover of Griffiths, pt. 2)

Let's consider the following equation from the inside cover of Griffiths:

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}.$$

This identity is actually quite useful when dealing with wave equations in electrodynamics[a]. Let's derive it. The first step in every index notation calculation is *translation*; turn your expression into an "index-only" expression. Doing this gives

$$\nabla \times (\nabla \times \mathbf{A}) = \mathbf{e}_a\partial_a \times (\mathbf{e}_b\partial_b \times \mathbf{e}_c A_c).$$

We now use our relations above to get rid of the cross products

$$\begin{aligned}\mathbf{e}_a\partial_a \times (\mathbf{e}_b\partial_b \times \mathbf{e}_c A_c) &= \mathbf{e}_a\partial_a \times (\mathbf{e}_d\epsilon_{dbc}\partial_b A_c)\\ &= \epsilon_{ead}\epsilon_{dbc}\partial_a\partial_b A_c.\end{aligned}$$

We now use part (a) of the problem below (it's marked Essential for a reason!) to get rid of the epsilons, and we get our final answer after "un-indexing" the expression,

$$\begin{aligned}\epsilon_{ead}\epsilon_{dbc}\partial_a\partial_b A_c &= \epsilon_{ead}\epsilon_{bcd}\partial_a\partial_b A_c\\ &= (\delta_{eb}\delta_{ac} - \delta_{ec}\delta_{ab})\partial_a\partial_b A_c\\ &= \delta_{eb}\delta_{ac}\partial_a\partial_b A_c - \delta_{ec}\delta_{ab}\partial_a\partial_b A_c\\ &= \partial_e\partial_c A_c - \partial_b\partial_b A_e\\ &= \boxed{\nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A}.}\end{aligned}$$

I'm being quite explicit here, and in all of the calculations I do in this handout, but that's for the sake of pedagogy; in practice, you should be far more succinct in your writing, leaving the calculation of simple identities in your head (analagous to not writing explicit steps when doing algebra or calculus).

[a] Specifically, one can input $\mathbf{B} \coloneqq \nabla \times \mathbf{A}$, with $\mathbf{A}$ being the vector potential, into Maxwell's equations to derive the wave equation for the vector potential (after choosing the Lorenz gauge).

Remark 1.6 (What does the Laplacian of a vector even mean?). Here's a question: why can we write $\nabla^2\mathbf{A}$? After all, we tend to think of $\nabla^2 \coloneqq \nabla\cdot\nabla$, and $\nabla\mathbf{A}$ on the face of it *makes no sense*. Here are two answers to this question:

1. There actually is no problem here; more generally, the Laplacian of any tensor A is defined as $\nabla^2 A = \nabla\cdot(\nabla A)$. The operations $\nabla\cdot$ and ∇ are defined through covariant differentiation and the like (connections), which have no trouble acting on vectors (which are just tensors). We only *think* this is an issue because we don't know enough.

2. We can simply *define* the vector Laplacian *using* the curl-of-a-curl identity, i.e.

$$\nabla^2 \coloneqq \nabla(\nabla\cdot) - \nabla\times(\nabla\times),$$

therefore completely sidestepping the troublesome business of taking $\nabla\mathbf{A}$.

3. In index notation, $\nabla\mathbf{A}$ is really a perfectly reasonable thing to write; it's just the nine-component object $\nabla\mathbf{A} \to \partial_a A_b$. This is a nine-component object, as it has two free indices a and b, so inputting $a, b = 1, 2, 3$ gives its components.

[10] **Problem 4.** (**Essential**)
Derive the following results

(a) $\epsilon_{abc}\epsilon_{dec} = \delta_{ad}\delta_{be} - \delta_{ae}\delta_{bd}$

(b) $\epsilon_{abc}\epsilon_{dbc} = 2\delta_{ad}$

(c) $\epsilon_{abc}\epsilon_{abc} = 6$

(d) $\mathbf{A} \times \mathbf{B} = A_a B_b \epsilon_{abc} \mathbf{e}_c$. Note that this implies that $(\mathbf{A} \times \mathbf{B})_c = A_a B_b \epsilon_{abc}$.

(e) $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = A_a B_b C_c \epsilon_{abc}$

(f) $(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} = (\mathbf{B} \times \mathbf{C}) \cdot \mathbf{A}$. This shows that the triple product is *symmetric under cyclic permutations*, which means that its value is the same if you permute the vectors cyclically[1].

(g) $\mathbf{A} \times \mathbf{A} = \mathbf{0}$

(h) $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$. This is known as the "BAC-CAB" rule, for obvious reasons.

Solution. (a) To confirm this, just check indices, i.e. plug in the non-zero values of the deltas on the RHS and check that it equals the LHS.

(b)

Idea 1.5 (Scalar and Vector Fields)

We can apply the idea of the Levi-Civita symbol to scalar and vector fields as well, just like we did for the Kronecker delta. We define the curl operator $\nabla\times$ acting on some vector field $\mathbf{g}(\mathbf{x}) = \mathbf{e}_a g_a(\mathbf{x})$ to be

$$\nabla \times \mathbf{g}(\mathbf{x}) \coloneqq \mathbf{e}_a \partial_a \times \mathbf{e}_b g_b(\mathbf{x}) = \mathbf{e}_c \epsilon_{abc} \partial_a g_b(\mathbf{x}).$$

Which is exactly what we would expect from our previous work.

[7] **Problem 5.** Using the two ideas above, verify the following results:

(a) $\nabla \times \mathbf{x} = \mathbf{0}$

(b) $\nabla \times (\mathbf{H} \times \mathbf{x}) = 2\mathbf{H}$, for some constant $\mathbf{H}$.

(c) $\nabla \cdot (\nabla \times \mathbf{g}(\mathbf{x})) = 0$

(d) $\nabla \times (\nabla f(\mathbf{x})) = \mathbf{0}$

(e) $\nabla \times (f(\mathbf{x})\mathbf{g}(\mathbf{x})) = f(\nabla \times \mathbf{g}) + (\nabla f) \times \mathbf{g}$

(f) (2pts) $\nabla \times (\mathbf{g}(\mathbf{x}) \times \mathbf{h}(\mathbf{x})) = \mathbf{g}(\nabla \cdot \mathbf{h}) - \mathbf{h}(\nabla \cdot \mathbf{g}) + (\mathbf{h} \cdot \nabla)\mathbf{g} - (\mathbf{g} \cdot \nabla)\mathbf{h}$

Solution. (a) $\nabla \times \mathbf{x} = \mathbf{e}_a \partial_a \times \mathbf{e}_b x_b = \mathbf{e}_c \epsilon_{aac} = 0$.

(b) $\nabla \times (\mathbf{H} \times \mathbf{x}) = \mathbf{e}_a \partial_a \times (\mathbf{e}_d \epsilon_{bcd} H_b x_c) = \mathbf{e}_f \epsilon_{fad} \epsilon_{bcd} \partial_a H_b x_c = \mathbf{e}_f (\delta_{fb}\delta_{ac} - \delta_{fc}\delta_{ab})\delta_{ac} H_b = \mathbf{e}_f (3H_f - H_f) = 2\mathbf{H}$.

[1] This identity does actually comes in handy sometimes, e.g. when you're showing that angular momentum is associated with rotational symmetry in classical mechanics.

(c) $\nabla \cdot (\nabla \times \mathbf{g}) = \epsilon_{aac}\partial_b g_c = 0$.

(d) $\nabla \times (\nabla f) = \mathbf{e}_a\partial_a \times (\mathbf{e}_b\partial_b f) = \mathbf{e}_c\epsilon_{abc}\partial_a\partial_b f = 0$. The reason why this is zero is because the epsilon symbol is anti-symmetric and the partial derivatives are symmetric. This means that when we permute a and b, we will just end up subtracting something from itself, so we'll get zero, i.e. $\partial_a\partial_b f - \partial_b\partial_a f = 0$ for all c. This idea is quite useful, so keep it in mind for the future.

(e) $\nabla \times (f\mathbf{g}) = \mathbf{e}_a\partial_a \times (f\mathbf{e}_b g_b) = \mathbf{e}_c\epsilon_{abc}\partial_a f g_b = \mathbf{e}_c\epsilon_{abc}\{(\partial_a f)g_b + f\partial_a g_b\} = \{\mathbf{e}_c\epsilon_{abc}(\partial_a f)g_b + f(\mathbf{e}_c\epsilon_{abc}\partial_a g_b)\} = f(\nabla \times \mathbf{g}) + (\nabla f) \times \mathbf{g}$

(f) We have

$$\begin{aligned}
\nabla \times (\mathbf{g} \times \mathbf{h}) &= \mathbf{e}_a\partial_a \times (\mathbf{e}_d\epsilon_{bcd}g_b h_c)\\
&= \mathbf{e}_f\epsilon_{fad}\epsilon_{bcd}\partial_a g_b h_c\\
&= \mathbf{e}_f\{(\delta_{fb}\delta_{ac} - \delta_{fc}\delta_{ab})\partial_a g_b h_c\}\\
&= \mathbf{e}_f\{\partial_c h_c g_f - \partial_a g_a h_f\}\\
&= \mathbf{e}_f\{(\partial_c h_c)g_f + h_c(\partial_c g_f) - (\partial_a g_a)h_f - g_a(\partial_a h_f)\}\\
&= \mathbf{g}(\nabla \cdot \mathbf{h}) - \mathbf{h}(\nabla \cdot \mathbf{g}) + (\mathbf{h} \cdot \nabla)\mathbf{g} - (\mathbf{g} \cdot \nabla)\mathbf{h}.
\end{aligned}$$

[7] **Problem 6.** Verify the following results.

(a) (4pts) For constant $\mathbf{h}$,

$$\oint_\Gamma d\mathbf{x} \cdot \left(\frac{1}{2}\mathbf{h} \times \mathbf{n}\right) = \pi\mathbf{h} \cdot \mathbf{n},$$

where Γ is a circle of unit radius, and $\mathbf{n}$ is the normal vector that specifies the orientation of the circle. **Hint:** use Stokes' theorem.

(b) (3pts) $\nabla^2(1/|\mathbf{x}|) = -4\pi\delta(\mathbf{x})$. **Hint:** use Gauss' theorem.

Solution. (a) Let $\partial S = \Gamma$. We have

$$\begin{aligned}
\oint d\mathbf{x} \cdot \left(\frac{1}{2}\mathbf{h} \times \mathbf{n}\right) &= \frac{1}{2}\oint_S d\mathbf{S} \cdot (\nabla \times (\mathbf{h} \times \mathbf{n}))\\
&= \frac{1}{2}\oint_S dS_a\mathbf{e}_a \cdot (\mathbf{e}_b\partial_b \times (\mathbf{e}_c\epsilon_{cde}h_d n_e))\\
&= \frac{1}{2}\oint_S dS_a\mathbf{e}_a \cdot (\mathbf{e}_f\epsilon_{fbc}\epsilon_{cde}\partial_b h_d n_e)\\
&= \frac{1}{2}\oint_S dS_a\epsilon_{abc}\epsilon_{dec}\partial_b h_d n_e\\
&= \frac{1}{2}\oint_S dS_a\epsilon_{abc}\epsilon_{dec}h_d\delta_{be}\\
&= \frac{1}{2}\oint_S dS_a 2h_d\delta_{ad}\\
&= \oint_S dS_a h_a = \pi\mathbf{h} \cdot \mathbf{n},
\end{aligned}$$

where the last line comes from the fact that $d\mathbf{S} = dS\,\mathbf{n}$ and the fact that $\oint dS = \pi$ by assumption. Let me comment on some of the things done above; the first is Stokes' theorem in the first equality. The next is the fact that $\partial_b h_d = 0$, as $\mathbf{h}$ is constant. Finally, we used the fact that $\epsilon_{aec}\epsilon_{dec} = 2\delta_{ad}$. As mentioned previously, n practice, you will just write out these things without saying what you're doing.

(b) **(to-do)**

Example 1.4

(include interesting physics problem that uses index notation)

(stuff below here is unfinished and very well may be completely deleted)

2 Advanced

This section is for people that want to go beyond the basics, towards more advanced subjects like quantum field theory or general relativity. It requires slightly more maturity than the previous section; if you don't know whether or not you "qualify", just try it and see what shakes out.

This section of the handout is primarily based on Kevin Zhou's notes on tensors, as well as Robert Littlejohn's notes on tensor analysis.

Idea 2.1 (Relativistic index notation)

We now adopt the following convention from here on out: Latin indices are implicitly summed from $i = 1, \ldots, 3$, while Greek indices are implicitly summed from $\mu = 0, 1, \ldots, 3$. This is to more elegantly account for the structure of *spacetime* instead of just *space*. Heuristically, the "time" or "scalar" part of some four-tensor is the zeroth component, and the next three components are the "space", or "vector" parts.

Remark 2.1 (Why use four-vectors?). While in the non-relativistic regieme, (speeds $v \ll c$), normal three-vectors work just fine; specifying a location in space, which we'll call $\mathbf{x}$, and having some background "lab clock" specifying the time was enough. In the relativistic regieme, this is not so. The key observation that necessitates this shift of thinking is that *times and lengths are now relative to your frame of reference*, and thus *a normal three-vector is now an incomplete way to specify an event*. A reasonable shift to make to accomodate this change is to specify a so-called **four-vector**, a vector with components of time *and* space.

Remark 2.2 (A shift in notation). In the relativistic regieme, we now must make an initially confusing move: we label coordinates via *superscripts* instead of *subscripts*. This means that if we have some vector, which we'll call $\mathbf{x}$, instead of labeling the components as $\mathbf{x} = (x_1, x_2, x_3)$, we label them as $\mathbf{x} = (x^1, x^2, x^3)$. Why do we do this? This is because the objects x_μ and x^μ are now *fundamentally different* in a relativistic context. The distinction between the two is that upper-index objects like x^μ are **contravariant** under Lorentz transformations, while lower-index objects like x_μ are **covariant** under Lorentz transformations. We'll explain what this means in a minute.

Idea 2.2 (Four-vectors)

A **four-vector** is a vector is defined as a vector with components

$$x^\mu = (ct, x^1, x^2, x^3).$$

The factor of c input into the zeroth component is done so on dimensional grounds. Letting $c = 1$ gives

$$x^\mu = (t, \mathbf{x}),$$

where we collated $(x^1, x^2, x^3) \to \mathbf{x}$.

Idea 2.3 (More four-vectors)

We may build other useful four vectors. The ones we're going to cover are *momentum*, *potential*, and *current*. They are given as (with $c = 1$)

$$p^\mu = (E, \mathbf{p}) \qquad A^\mu = (\phi, \mathbf{A}) \qquad J^\mu = (\rho, \mathbf{J}).$$

If you don't have even a passing familiarity with these, I'd recommend reading chapter 12 of Griffiths' *Introduction to Electrodynamics*.

Idea 2.4 (The metric)

We define the **metric** to be the matrix

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

and its upper-index counterpart, whose values are the exact same, (i.e. $\eta^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$). You use the metric to raise and lower indices, e.g. $p_\mu = \eta_{\mu\nu} p^\nu$. Functionally this just means that all of the space components of your vector flip sign, so

$$p_0 = p^0, \qquad p_1 = -p^1, \qquad p_2 = -p^2, \qquad p_3 = -p^3.$$

The number of negative signs in the metric dictate what's called its **signature**; we've used the **mostly minus** metric here, but you can also define things to be "mostly plus", i.e. $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$.

Remark 2.3 (Upper vs. lower indices)**.** (**make remark on the difference between these; contracting one with another is only thing allowed unless you're matthew schwartz**)

[3] **Problem 7.** Solve the following problems:

(a) Show that $\eta^{\mu\nu}\eta_{\nu\rho} = \delta^\mu_\rho$.

(b) Find the numeric value of $\eta_{\mu\nu}\eta^{\mu\nu}$.

(c) Here's another example of a covector: the **four-derivative**. We define it to be $\partial_\mu = (\partial_t, \nabla)$. Explicitly write out the components of $\partial \cdot A = \partial_\mu A^\mu$ and $p \cdot A = p_\mu A^\mu$.

Idea 2.5 (Relativistic Levi-Civita)

We define the rank-four Levi-Civita symbol to be

$$\epsilon^{\mu\nu\rho\sigma} = \begin{cases} \pm 1 & \mu, \nu, \rho, \sigma \text{ distinct} \\ 0 & \text{else.} \end{cases}$$

We also define $\epsilon^{0123} = 1$ and note that the symbol is completely anti-symmetric, so it picks up a negative sign upon exchange of any of its indices. Notice how this is exactly the same definition as earlier, just with another component and number.

[1] **Problem 8.** Show that $\eta_{\mu\nu}\epsilon^{\mu\nu\rho\sigma}$ for any ρ and σ. (**what**)

[3] **Problem 9.** We have that[2]

$$\begin{aligned} \epsilon_{\mu\nu\rho\sigma}\epsilon^{\mu\nu\gamma\delta} &= c_1(\delta^\delta_\rho\delta^\gamma_\sigma - \delta^\gamma_\rho\delta^\delta_\sigma) \\ \epsilon_{\mu\nu\rho\sigma}\epsilon^{\mu\nu\rho\delta} &= c_2\delta^\delta_\sigma \\ \epsilon_{\mu\nu\rho\sigma}\epsilon^{\mu\nu\rho\sigma} &= c_3. \end{aligned}$$

Find the coefficients c_i using any method you can think of.

(**finish**)

[2] You can take this on faith, or feel free to prove it yourself.