Appendix A

Appendices

A.1 ϵ_{ijk} and cross products

A.1.1 Vector Operations: δ_{ij} and ϵ_{ijk}

These are some notes on the use of the antisymmetric symbol ϵ_{ijk} for expressing cross products. This is an extremely powerful tool for manipulating cross products and their generalizations in higher dimensions, and although many low level courses avoid the use of ϵ, I think this is a mistake and I want you to become proficient with it.

In a cartesian coordinate system a vector $\vec{V}$ has components V_i along each of the three orthonormal basis vectors $\hat{e}_i$, or $\vec{V} = \sum_i V_i\hat{e}_i$. The dot product of two vectors, $\vec{A}\cdot\vec{B}$, is bilinear and can therefore be written as

$$\begin{aligned}
\vec{A}\cdot\vec{B} &= (\sum_i A_i\hat{e}_i)\cdot\sum_j B_j\hat{e}_j && \text{(A.1)}\\
&= \sum_i\sum_j A_iB_j\hat{e}_i\cdot\hat{e}_j && \text{(A.2)}\\
&= \sum_i\sum_j A_iB_j\delta_{ij}, && \text{(A.3)}
\end{aligned}$$

where the Kronecker delta δ_{ij} is defined to be 1 if $i=j$ and 0 otherwise. As the basis vectors $\hat{e}_k$ are orthonormal, *i.e.* orthogonal to each other and of unit length, we have $\hat{e}_i\cdot\hat{e}_j=\delta_{ij}$.

Doing a sum over an index j of an expression involving a δ_{ij} is very simple, because the only term in the sum which contributes is the one with $j=i$. Thus $\sum_j F(i,j)\delta_{ij} = F(i,i)$, which is to say, one just replaces j with i in all

the other factors, and drops the δ_{ij} and the summation over j. So we have $\vec{A}\cdot\vec{B}=\sum_i A_iB_i$, the standard expression for the dot product[1]

We now consider the cross product of two vectors, $\vec{A}\times\vec{B}$, which is also a bilinear expression, so we must have $\vec{A}\times\vec{B} = (\sum_i A_i\hat{e}_i)\times(\sum_j B_j\hat{e}_j) = \sum_i\sum_j A_iB_j(\hat{e}_i\times\hat{e}_j)$. The cross product $\hat{e}_i\times\hat{e}_j$ is a vector, which can therefore be written as $\vec{V}=\sum_k V_k\hat{e}_k$. But the vector result depends also on the two input vectors, so the coefficients V_k really depend on i and j as well. Define them to be ϵ_{ijk}, so

$$\hat{e}_i\times\hat{e}_j = \sum_k \epsilon_{kij}\hat{e}_k.$$

It is easy to evaluate the 27 coefficients ϵ_{kij}, because the cross product of two orthogonal unit vectors is a unit vector orthogonal to both of them. Thus $\hat{e}_1\times\hat{e}_2=\hat{e}_3$, so $\epsilon_{312}=1$ and $\epsilon_{k12}=0$ if $k=1$ or 2. Applying the same argument to $\hat{e}_2\times\hat{e}_3$ and $\hat{e}_3\times\hat{e}_1$, and using the antisymmetry of the cross product, $\vec{A}\times\vec{B}=-\vec{B}\times\vec{A}$, we see that

$$\epsilon_{123}=\epsilon_{231}=\epsilon_{312}=1; \qquad \epsilon_{132}=\epsilon_{213}=\epsilon_{321}=-1,$$

and $\epsilon_{ijk}=0$ for all other values of the indices, *i.e.* $\epsilon_{ijk}=0$ whenever any two of the indices are equal. Note that ϵ changes sign not only when the last two indices are interchanged (a consequence of the antisymmetry of the cross product), but whenever *any* two of its indices are interchanged. Thus ϵ_{ijk} is zero unless $(1,2,3)\to(i,j,k)$ is a permutation, and is equal to the sign of the permutation if it exists.

Now that we have an expression for $\hat{e}_i\times\hat{e}_j$, we can evaluate

$$\vec{A}\times\vec{B} = \sum_i\sum_j A_iB_j(\hat{e}_i\times\hat{e}_j) = \sum_i\sum_j\sum_k \epsilon_{kij}A_iB_j\hat{e}_k. \tag{A.4}$$

Much of the usefulness of expressing cross products in terms of ϵ's comes from the identity

$$\sum_k \epsilon_{kij}\epsilon_{k\ell m} = \delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell}, \tag{A.5}$$

which can be shown as follows. To get a contribution to the sum, k must be different from the unequal indices i and j, and also different from ℓ and m. Thus we get 0 unless the pair (i,j) and the pair (ℓ,m) are the same pair of

[1]Note that this only holds because we have expressed our vectors in terms of *orthonormal* basis vectors.

different indices. There are only two ways that can happen, as given by the two terms, and we only need to verify the coefficients. If $i = \ell$ and $j = m$, the two ϵ's are equal and the square is 1, so the first term has the proper coefficient of 1. The second term differs by one transposition of two indices on one epsilon, so it must have the opposite sign.

We now turn to some applications. Let us first evaluate

$$\vec{A} \cdot (\vec{B} \times \vec{C}) = \sum_i A_i \sum_{jk} \epsilon_{ijk} B_j C_k = \sum_{ijk} \epsilon_{ijk} A_i B_j C_k. \tag{A.6}$$

Note that $\vec{A} \cdot (\vec{B} \times \vec{C})$ is, up to sign, the volume of the parallelopiped formed by the vectors $\vec{A}$, $\vec{B}$, and $\vec{C}$. From the fact that the ϵ changes sign under transpositions of any two indices, we see that the same is true for transposing the vectors, so that

$$\begin{aligned} \vec{A} \cdot (\vec{B} \times \vec{C}) = -\vec{A} \cdot (\vec{C} \times \vec{B}) &= \vec{B} \cdot (\vec{C} \times \vec{A}) = -\vec{B} \cdot (\vec{A} \times \vec{C}) \\ &= \vec{C} \cdot (\vec{A} \times \vec{B}) = -\vec{C} \cdot (\vec{B} \times \vec{A}). \end{aligned}$$

Now consider $\vec{V} = \vec{A} \times (\vec{B} \times \vec{C})$. Using our formulas,

$$\vec{V} = \sum_{ijk} \epsilon_{kij} \hat{e}_k A_i (\vec{B} \times \vec{C})_j = \sum_{ijk} \epsilon_{kij} \hat{e}_k A_i \sum_{lm} \epsilon_{jlm} B_l C_m.$$

Notice that the sum on j involves only the two epsilons, and we can use

$$\sum_j \epsilon_{kij} \epsilon_{jlm} = \sum_j \epsilon_{jki} \epsilon_{jlm} = \delta_{kl}\delta_{im} - \delta_{km}\delta_{il}.$$

Thus

$$\begin{aligned} V_k &= \sum_{ilm} (\sum_j \epsilon_{kij}\epsilon_{jlm}) A_i B_l C_m = \sum_{ilm} (\delta_{kl}\delta_{im} - \delta_{km}\delta_{il}) A_i B_l C_m \\ &= \sum_{ilm} \delta_{kl}\delta_{im} A_i B_l C_m - \sum_{ilm} \delta_{km}\delta_{il} A_i B_l C_m \\ &= \sum_i A_i B_k C_i - \sum_i A_i B_i C_k = \vec{A} \cdot \vec{C}\, B_k - \vec{A} \cdot \vec{B}\, C_k, \end{aligned}$$

so

$$\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}\, \vec{A} \cdot \vec{C} - \vec{C}\, \vec{A} \cdot \vec{B}. \tag{A.7}$$

This is sometimes known as the **bac-cab** formula.

Exercise: Using (A.5) for the manipulation of cross products, show that

$$(\vec{A} \times \vec{B}) \cdot (\vec{C} \times \vec{D}) = \vec{A} \cdot \vec{C} \vec{B} \cdot \vec{D} - \vec{A} \cdot \vec{D} \vec{B} \cdot \vec{C}.$$

The determinant of a matrix can be defined using the ϵ symbol. For a 3×3 matrix A,

$$\det A = \sum_{ijk} \epsilon_{ijk} A_{1i} A_{2j} A_{3k} = \sum_{ijk} \epsilon_{ijk} A_{i1} A_{j2} A_{k3}.$$

From the second definition, we see that the determinant is the volume of the parallelopiped formed from the images under the linear map A of the three unit vectors $\hat{e}_i$, as

$$(A\hat{e}_1) \cdot ((A\hat{e}_2) \times (A\hat{e}_3)) = \det A.$$

In higher dimensions, the cross product is not a vector, but there is a generalization of ϵ which remains very useful. In an n-dimensional space, $\epsilon_{i_1 i_2 \ldots i_n}$ has n indices and is defined as the sign of the permutation $(1, 2, \ldots, n) \to (i_1 i_2 \ldots i_n)$, if the indices are all unequal, and zero otherwise. The analog of (A.5) has $(n-1)!$ terms from all the permutations of the unsummed indices on the second ϵ. The determinant of an $n \times n$ matrix is defined as

$$\det A = \sum_{i_1, \ldots, i_n} \epsilon_{i_1 i_2 \ldots i_n} \prod_{p=1}^{n} A_{p, i_p}.$$

A.2 The gradient operator

We can define the gradient operator

$$\vec{\nabla} = \sum_i \hat{e}_i \frac{\partial}{\partial x_i}. \tag{A.8}$$

While this looks like an ordinary vector, the coefficients are not numbers V_i but are operators, which do not commute with functions of the coordinates x_i. We can still write out the components straightforwardly, but we must be careful to keep the order of the operators and the fields correct.

The gradient of a scalar field $\Phi(\vec{r})$ is simply evaluated by distributing the gradient operator

$$\vec{\nabla}\Phi = (\sum_i \hat{e}_i \frac{\partial}{\partial x_i}) \Phi(\vec{r}) = \sum_i \hat{e}_i \frac{\partial \Phi}{\partial x_i}. \tag{A.9}$$

Because the individual components obey the Leibnitz rule $\frac{\partial AB}{\partial x_i} = \frac{\partial A}{\partial x_i}B + A\frac{\partial B}{\partial x_i}$, so does the gradient, so if A and B are scalar fields,

$$\vec{\nabla} AB = (\vec{\nabla} A)B + A\vec{\nabla} B. \tag{A.10}$$

The general application of the gradient operator $\vec{\nabla}$ to a *vector* $\vec{A}$ gives an object with coefficients with two indices, a *tensor*. Some parts of this tensor, however, can be simplified. The first (which is the trace of the tensor) is called the *divergence* of the vector, written and defined by

$$\begin{aligned} \vec{\nabla}\cdot\vec{A} &= (\sum_i \hat{e}_i \frac{\partial}{\partial x_i})\cdot(\sum_j \hat{e}_j B_j) = \sum_{ij} \hat{e}_i\cdot\hat{e}_j \frac{\partial B_j}{\partial x_i} = \sum_{ij}\delta_{ij}\frac{\partial B_j}{\partial x_i} \\ &= \sum_i \frac{\partial B_i}{\partial x_i}. \end{aligned} \tag{A.11}$$

In asking about Leibnitz' rule, we must remember to apply the divergence operator only to vectors. One possibility is to apply it to the vector $\vec{V} = \Phi\vec{A}$, with components $V_i = \Phi A_i$. Thus

$$\begin{aligned} \vec{\nabla}\cdot(\Phi\vec{A}) &= \sum_i \frac{\partial(\Phi A_i)}{\partial x_i} = \sum_i \frac{\partial \Phi}{\partial x_i}A_i + \Phi\sum_i \frac{\partial A_i}{\partial x_i} \\ &= (\vec{\nabla}\Phi)\cdot\vec{A} + \Phi\vec{\nabla}\cdot\vec{A}. \end{aligned} \tag{A.12}$$

We could also apply the divergence to the cross product of two vectors,

$$\begin{aligned} \vec{\nabla}\cdot(\vec{A}\times\vec{B}) &= \sum_i \frac{\partial(\vec{A}\times\vec{B})_i}{\partial x_i} = \sum_i \frac{\partial(\sum_{jk}\epsilon_{ijk}A_jB_k)}{\partial x_i} = \sum_{ijk}\epsilon_{ijk}\frac{\partial(A_jB_k)}{\partial x_i} \\ &= \sum_{ijk}\epsilon_{ijk}\frac{\partial A_j}{\partial x_i}B_k + \sum_{ijk}\epsilon_{ijk}A_j\frac{\partial B_k}{\partial x_i}. \end{aligned} \tag{A.13}$$

This is expressible in terms of the *curls* of $\vec{A}$ and $\vec{B}$.

The curl is like a cross product with the first vector replaced by the differential operator, so we may write the i'th component as

$$(\vec{\nabla}\times\vec{A})_i = \sum_{jk}\epsilon_{ijk}\frac{\partial}{\partial x_j}A_k. \tag{A.14}$$

We see that the last expression in (A.13) is

$$\sum_k(\sum_{ij}\epsilon_{kij}\frac{\partial A_j}{\partial x_i})B_k - \sum_j A_j\sum_{ik}\epsilon_{jik}\frac{\partial B_k}{\partial x_i} = (\vec{\nabla}\times\vec{A})\cdot\vec{B} - \vec{A}\cdot(\vec{\nabla}\times\vec{B}). \tag{A.15}$$

where the sign which changed did so due to the transpositions in the indices on the ϵ, which we have done in order to put things in the form of the definition of the curl. Thus

$$\vec{\nabla}\cdot(\vec{A}\times\vec{B}) = (\vec{\nabla}\times\vec{A})\cdot\vec{B} - \vec{A}\cdot(\vec{\nabla}\times\vec{B}). \tag{A.16}$$

Vector algebra identities apply to the curl as to any ordinary vector, except that one must be careful not to change, by reordering, what the differential operators act on. In particular, Eq. A.7 is

$$\vec{A}\times(\vec{\nabla}\times\vec{B}) = \sum_i A_i\vec{\nabla}B_i - \sum_i A_i\frac{\partial\vec{B}}{\partial x_i}. \tag{A.17}$$

A.3 Gradient in Spherical Coordinates

The transformation between Cartesian and spherical coordinates is given by

$$\begin{array}{ll} r = (x^2+y^2+z^2)^{\frac{1}{2}} & x = r\sin\theta\cos\phi \\ \theta = \cos^{-1}(z/r) & y = r\sin\theta\sin\phi \\ \phi = \tan^{-1}(y/x) & z = r\cos\theta \end{array}$$

The basis vectors $\{\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi\}$ at the point (r,θ,ϕ) are given in terms of the cartesian basis vectors by

$$\begin{aligned} \hat{e}_r &= \sin\theta\cos\phi\,\hat{e}_x + \sin\theta\sin\phi\,\hat{e}_y + \cos\theta\,\hat{e}_z \\ \hat{e}_\theta &= \cos\theta\cos\phi\,\hat{e}_x + \cos\theta\sin\phi\,\hat{e}_y - \sin\theta\,\hat{e}_z \\ \hat{e}_\phi &= -\sin\phi\,\hat{e}_x + \cos\phi\,\hat{e}_y. \end{aligned}$$

By the chain rule, if we have two sets of coordinates, say s_i and c_i, and we know the form a function $f(s_i)$ and the dependence of s_i on c_j, we can find $\frac{\partial f}{\partial c_i} = \sum_j \left.\frac{\partial f}{\partial s_j}\right|_s \left.\frac{\partial s_j}{\partial c_i}\right|_c$, where $|_s$ means hold the other s's fixed while varying s_j. In our case, the s_j are the spherical coordinates r,θ,ϕ, while the c_i are x,y,z.

Thus

$$\begin{aligned} \vec{\nabla}f &= \left(\left.\frac{\partial f}{\partial r}\right|_{\theta\phi}\left.\frac{\partial r}{\partial x}\right|_{yz} + \left.\frac{\partial f}{\partial \theta}\right|_{r\phi}\left.\frac{\partial \theta}{\partial x}\right|_{yz} + \left.\frac{\partial f}{\partial \phi}\right|_{r\theta}\left.\frac{\partial \phi}{\partial x}\right|_{yz}\right)\hat{e}_x \\ &+ \left(\left.\frac{\partial f}{\partial r}\right|_{\theta\phi}\left.\frac{\partial r}{\partial y}\right|_{xz} + \left.\frac{\partial f}{\partial \theta}\right|_{r\phi}\left.\frac{\partial \theta}{\partial y}\right|_{xz} + \left.\frac{\partial f}{\partial \phi}\right|_{r\theta}\left.\frac{\partial \phi}{\partial y}\right|_{xz}\right)\hat{e}_y \end{aligned} \tag{A.18}$$

$$+\left(\left.\frac{\partial f}{\partial r}\right|_{\theta\phi}\left.\frac{\partial r}{\partial z}\right|_{xy}+\left.\frac{\partial f}{\partial \theta}\right|_{r\phi}\left.\frac{\partial \theta}{\partial z}\right|_{xy}+\left.\frac{\partial f}{\partial \phi}\right|_{r\theta}\left.\frac{\partial \phi}{\partial z}\right|_{xy}\right)\hat{e}_z$$

We will need all the partial derivatives $\frac{\partial s_j}{\partial c_i}$. From $r^2 = x^2 + y^2 + z^2$ we see that

$$\left.\frac{\partial r}{\partial x}\right|_{yz} = \frac{x}{r} \qquad \left.\frac{\partial r}{\partial y}\right|_{xz} = \frac{y}{r} \qquad \left.\frac{\partial r}{\partial z}\right|_{xy} = \frac{z}{r}.$$

From $\cos\theta = z/r = z/\sqrt{x^2+y^2+z^2}$,

$$-\sin\theta\left.\frac{\partial \theta}{\partial x}\right|_{yz} = \frac{-zx}{(x^2+y^2+z^2)^{3/2}} = \frac{-r^2\cos\theta\sin\theta\cos\phi}{r^3}$$

so

$$\left.\frac{\partial \theta}{\partial x}\right|_{yz} = \frac{\cos\theta\cos\phi}{r}.$$

Similarly,

$$\left.\frac{\partial \theta}{\partial y}\right|_{xz} = \frac{\cos\theta\sin\phi}{r}.$$

There is an extra term when differentiating w.r.t. z, from the numerator, so

$$-\sin\theta\left.\frac{\partial \theta}{\partial z}\right|_{xy} = \frac{1}{r} - \frac{z^2}{r^3} = \frac{1-\cos^2\theta}{r} = r^{-1}\sin^2\theta,$$

so

$$\left.\frac{\partial \theta}{\partial z}\right|_{xy} = -r^{-1}\sin\theta.$$

Finally, the derivatives of ϕ can easily be found from differentiating $\tan\phi = y/x$. Using differentials,

$$\sec^2\phi d\phi = \frac{dy}{x} - \frac{ydx}{x^2} = \frac{dy}{r\sin\theta\cos\phi} - \frac{dx\sin\theta\sin\phi}{r\sin^2\theta\cos^2\phi}$$

so

$$\left.\frac{\partial \phi}{\partial x}\right|_{yz} = -\frac{1}{r}\frac{\sin\phi}{\sin\theta} \qquad \left.\frac{\partial \phi}{\partial y}\right|_{xz} = \frac{1}{r}\frac{\cos\phi}{\sin\theta} \qquad \left.\frac{\partial \phi}{\partial z}\right|_{xy} = 0.$$

Now we are ready to plug this all into (A.18). Grouping together the terms involving each of the three partial derivatives, we find

$$\begin{aligned}
\vec{\nabla} f &= \left.\frac{\partial f}{\partial r}\right|_{\theta\phi}\left(\frac{x}{r}\hat{e}_x + \frac{y}{r}\hat{e}_y + \frac{z}{r}\hat{e}_z\right) \\
&\quad + \left.\frac{\partial f}{\partial \theta}\right|_{r\phi}\left(\frac{\cos\theta\cos\phi}{r}\hat{e}_x + \frac{\cos\theta\sin\phi}{r}\hat{e}_y - \frac{\sin\theta}{r}\hat{e}_z\right) \\
&\quad + \left.\frac{\partial f}{\partial \phi}\right|_{r\theta}\left(-\frac{1}{r}\frac{\sin\phi}{\sin\theta}\hat{e}_x + \frac{1}{r}\frac{\cos\phi}{\sin\theta}\hat{e}_y\right) \\
&= \left.\frac{\partial f}{\partial r}\right|_{\theta\phi}\hat{e}_r + \frac{1}{r}\left.\frac{\partial f}{\partial \theta}\right|_{r\phi}\hat{e}_\theta + \frac{1}{r\sin\theta}\left.\frac{\partial f}{\partial \phi}\right|_{r\theta}\hat{e}_\phi
\end{aligned}$$

Thus we have derived the form for the gradient in spherical coordinates.