MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Physics Department

Physics 8.07: Electromagnetism II October 17, 2012

Prof. Alan Guth

LECTURE NOTES 9

TRACELESS SYMMETRIC TENSOR APPROACH TO LEGENDRE POLYNOMIALS AND SPHERICAL HARMONICS

In these notes I will describe the separation of variable technique for solving Laplace's equation, using spherical polar coordinates. The solutions will involve Legendre polynomials for cases with azimuthal symmetry, and more generally they will involve spherical harmonics. I will construct these solutions using traceless symmetric tensors, but in Lecture Notes 8 I describe how the solutions in this form relate to the more standard expressions in terms of Legendre polynomials and spherical harmonics. (Logically Lecture Notes 8 should come after these notes, although they were posted first.) If you are starting from scratch, I think that the traceless symmetric tensor method is the simplest way to understand this mathematical formalism. If you already know spherical harmonics, I think that you will find the traceless symmetric tensor approach to be a useful addition to your arsenal of mathematical methods. The symmetric traceless tensor approach is particularly useful if one needs to extend the formalism beyond what we will be doing — for example, there are analogues of spherical harmonics in higher dimensions, and there are also vector spherical harmonics that are useful for expanding vector functions of angle. Vector spherical harmonics are used in the most general treatments of electromagnetic radiation, although we will not be introducing them in this course.

I don't know a reference for the traceless symmetric tensor method, which is the main reason I am writing these notes. For the standard method, limited to the case of azimuthal symmetry, our textbook by Griffiths should be sufficient. If you would like to see an additional reference on spherical harmonics, which are needed when there is no azimuthal symmetry, then I would recommend J.D. Jackson, **Classical Electrodynamics**, 3rd Edition (John Wiley & Sons, 1999), Sections 3.1, 3.2, 3.5, and 3.6.

1. LAPLACE'S EQUATION IN SPHERICAL COORDINATES:

In spherical coordinates, Laplace's equation can be written as

$$\boxed{\nabla^2\varphi(r,\theta,\phi) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\varphi}{\partial r}\right) + \frac{1}{r^2}\nabla^2_\theta\,\varphi = 0\ ,} \tag{9.1}$$

where the angular part is given by

$$\boxed{\nabla^2_\theta\,\varphi \equiv \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\varphi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\varphi}{\partial\phi^2}\ .} \tag{9.2}$$

It would be more logical to write ∇^2_θ as $\nabla^2_{\theta,\phi}$, but it would be tiresome to write it that way. It is sometimes useful to rewrite the first term in Eq. (9.1) using

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\varphi}{\partial r}\right) = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\varphi) \ . \tag{9.3}$$

To use the method of separation of variables, we can seek a solution of the form

$$\varphi(r,\theta,\phi) = R(r)F(\theta,\varphi) \ . \tag{9.4}$$

Then Laplace's equation can be written as

$$0 = \frac{r^2}{RF}\nabla^2\varphi = \frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) + \frac{1}{F}\nabla^2_\theta F \ . \tag{9.5}$$

Since the first term on the right-hand side depends only on r, and the second term depends only on θ and ϕ, the only way that the equation can be satisfied is if each term is a constant. Thus we can write

$$\frac{1}{F}\nabla^2_\theta F = C_\theta \ , \tag{9.6}$$

$$\frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) = -C_\theta \ . \tag{9.7}$$

2. THE EXPANSION OF $F(\theta,\phi)$:

We now wish to find the most general solution to the equation

$$\nabla^2_\theta F = C_\theta F \ , \tag{9.8}$$

which is a rewriting of Eq. (9.6). If such a function F can be found, we say that F is an *eigenfunction* of the operator ∇^2_θ, with *eigenvalue* C_θ.

A function of angles (θ,ϕ) can equivalently be thought of as a function of the unit vector $\hat{n}$ that points in the direction of θ and ϕ, which can be written explicitly as

$$\hat{n} = \sin\theta\cos\phi\,\hat{e}_1 + \sin\theta\sin\phi\,\hat{e}_2 + \cos\theta\,\hat{e}_3 \ , \tag{9.9}$$

where $\hat{e}_1$, $\hat{e}_2$, and $\hat{e}_3$ can also be written as $\hat{e}_x$, $\hat{e}_y$, and $\hat{e}_z$. I am labeling the unit vectors using the numbers 1, 2, and 3 when I am thinking about summing over the indices, and otherwise I use x, y, and z.

I now claim that the most general function of (θ,ϕ) can be written as a power series in $\hat{n}$, or more precisely as a power series in the components of $\hat{n}$. I will not prove this, but it is true at least for square-integrable piece-wise continuous functions $F(\theta, \phi)$. Such a power series can be written as

$$\boxed{F(\hat{n}) = C^{(0)} + C^{(1)}_i \hat{n}_i + C^{(2)}_{ij} \hat{n}_i \hat{n}_j + \ldots + C^{(\ell)}_{i_1 i_2 \ldots i_\ell} \hat{n}_{i_1} \hat{n}_{i_2} \ldots \hat{n}_{i_\ell} + \ldots \ ,} \tag{9.10}$$

where repeated indices are summed from 1 to 3 (as Cartesian coordinates). Note that $C^{(\ell)}_{i_1 i_2 \ldots i_\ell} \hat{n}_{i_1} \hat{n}_{i_2} \ldots \hat{n}_{i_\ell}$ represents the general term in the series, where the first three terms correspond to $\ell = 0$, $\ell = 1$, and $\ell = 2$. The indices i_1, i_2, $\ldots$, i_ℓ represent ℓ different indices, like i and j; since they are repeated, they are each summed from 1 to 3.

The coefficients $C^{(\ell)}_{i_1 i_2 \ldots i_\ell}$ are called *tensors*, and the number of indices is called the rank of the tensor. Note that $C^{(0)}$, $C^{(1)}_i$, and $C^{(2)}_{ij}$ are special cases of tensors, although they can also be considered a scalar, a vector, and a matrix.

It is possible to impose restrictions on the coefficients of Eq. (9.10) without actually restricting what can appear on the right-hand side of the equations. In particular, we will insist that

1) The tensors $C^{(\ell)}_{i_1 i_2 \ldots i_\ell}$ are symmetric under any reordering of the indices:

$$C^{(\ell)}_{i_1 i_2 \ldots i_\ell} = C^{(\ell)}_{j_1 j_2 \ldots j_\ell} \ , \tag{9.11}$$

where $\{j_1, j_2, \ldots, j_\ell\}$ is any permutation of $\{i_1, i_2, \ldots, i_\ell\}$.

2) The tensors $C^{(\ell)}_{i_1 i_2 \ldots i_\ell}$ are traceless, in the sense that if any two indices are set equal to each other and summed, the result is equal to zero. Since the tensors are already assumed to be symmetric, it does not matter which indices are summed, so we can choose the last two:

$$C^{(\ell)}_{i_1 i_2 \ldots i_{\ell-2} jj} = 0 \ . \tag{9.12}$$

To explain why these restrictions on the $C^{(\ell)}$'s do not impose any restriction on the right-hand side of Eq. (9.10), I will use the example of $C^{(2)}_{ij}$, but I think you will be able to see that that the argument applies to all ℓ. The insistence that $C^{(2)}_{ij}$ is symmetric can be seen to make no difference to the right-hand side of Eq. (9.10), because $C^{(2)}_{ij}$ multiplies the symmetric tensor $\hat{n}_i \hat{n}_j$. Thus, if $C^{(2)}_{ij}$ had an antisymmetric part, it would not contribute to the right-hand side of Eq. (9.10). The requirement of tracelessness is less obvious, but suppose that $C^{(2)}_{ij}$ were not traceless. Then we could write

$$C^{(2)}_{ii} = \lambda \neq 0 \ . \tag{9.13}$$

We could then define a new quantity,

$$\tilde{C}^{(2)}_{ij} = C^{(2)}_{ij} - \frac{1}{3}\lambda\delta_{ij} \ . \tag{9.14}$$

It follows that $\tilde{C}^{(2)}_{ij}$ is traceless:

$$\tilde{C}^{(2)}_{ii} = C^{(2)}_{ii} - \frac{1}{3}\lambda\delta_{ii} = \lambda - \frac{1}{3}\lambda\delta_{ii} = 0 \ , \tag{9.15}$$

since $\delta_{ii} = 3$. The original term $C^{(2)}_{ij}\hat{n}_i\hat{n}_j$ can then be expressed in terms of $\tilde{C}^{(2)}_{ij}$:

$$C^{(2)}_{ij}\hat{n}_i\hat{n}_j = \left[\tilde{C}^{(2)}_{ij} + \frac{1}{3}\lambda\delta_{ij}\right]\hat{n}_i\hat{n}_j = \tilde{C}^{(2)}_{ij}\hat{n}_i\hat{n}_j + \frac{1}{3}\lambda \ , \tag{9.16}$$

where we used the fact that $\delta_{ij}\hat{n}_i\hat{n}_j = 1$, since $\hat{n}$ is a unit vector. The extra term, $\frac{1}{3}\lambda$, can then be absorbed into a redefinition of $C^{(0)}$:

$$\tilde{C}^{(0)} = C^{(0)} + \frac{1}{3}\lambda \ . \tag{9.17}$$

Finally, we can write

$$C^{(0)} + C^{(1)}_i\hat{n}_i + C^{(2)}_{ij}\hat{n}_i\hat{n}_j = \tilde{C}^{(0)} + C^{(1)}_i\hat{n}_i + \tilde{C}^{(2)}_{ij}\hat{n}_i\hat{n}_j \ , \tag{9.18}$$

so we can insist that the tensor that multiplies $\hat{n}_i\hat{n}_j$ be traceless with no restriction on what functions can be expressed in this form.

I will call the ℓ'th term of this expansion $F_\ell(\hat{n})$, so

$$F_\ell(\hat{n}) = C^{(\ell)}_{i_1 i_2 \ldots i_\ell}\hat{n}_{i_1}\hat{n}_{i_2}\ldots\hat{n}_{i_\ell} \ . \tag{9.19}$$

3. EVALUATION OF $\nabla^2_\theta F_\ell(\hat{n})$:

To evaluate $\nabla^2_\theta F_\ell(\hat{n})$, we are going to take advantage of a convenient trick. Instead of dealing directly with $F_\ell(\hat{n})$, we will instead introduce a radial variable r, using it to define a coordinate vector

$$\vec{r} = r\hat{n} \ . \tag{9.20}$$

Following the notation of Griffiths (see his Eq. (1.19)), I will denote the coordinates of $\vec{r}$ by x, y, and z, or in index notation I will call them x_i. So

$$\vec{r} = x_i\hat{e}_i = x_1\hat{e}_1 + x_2\hat{e}_2 + x_3\hat{e}_3 \ . \tag{9.21}$$

Then, given any $F_\ell(\hat{n})$ of the form given in Eq. (9.19), we can define a function $\tilde{F}_\ell(\vec{r})$ by

$$\tilde{F}_\ell(\vec{r}) = C^{(\ell)}_{i_1 i_2 \ldots i_\ell} x_{i_1} x_{i_2} \ldots x_{i_\ell} = r^\ell F_\ell(\hat{n}) \ . \tag{9.22}$$

Note that $C^{(\ell)}_{i_1 i_2 \ldots i_\ell}$ is the same rank ℓ traceless symmetric tensor used to define $F_\ell(\hat{n})$, but we are defining $\tilde{F}_\ell(\vec{r})$ by multiplying $C^{(\ell)}_{i_1 i_2 \ldots i_\ell}$ by $x_{i_1} x_{i_2} \ldots x_{i_\ell}$ and then summing over indices, instead of multiplying by $\hat{n}_{i_1} \hat{n}_{i_2} \ldots \hat{n}_{i_\ell}$ and then summing.

Now we can make use of Eq. (9.1), which relates the full Laplacian ∇^2 to the angular Laplacian, ∇^2_θ. We will find that in this case the full Laplacian and the radial derivative piece of Eq. (9.1) will both be simple, so we will be able to determine the angular Laplacian by evaluating the other terms in Eq. (9.1).

To evaluate $\nabla^2 \tilde{F}_\ell(\vec{r})$, we start with $\ell = 0$. Clearly

$$\nabla^2 \tilde{F}_0(\vec{r}) = \nabla^2 C_{(0)} = 0 \ , \tag{9.23}$$

since the derivative of a constant vanishes. Similarly for $\ell = 1$,

$$\nabla^2 \tilde{F}_1(\vec{r}) = \nabla^2 C^{(1)}_i x_i = 0 \ , \tag{9.24}$$

since the first derivative produces a constant, so the second derivative vanishes. The first nontrivial case is $\ell = 2$:

$$\begin{aligned}
\nabla^2 \tilde{F}_2(\vec{r}) &= \nabla^2 C^{(2)}_{ij} x_i x_j \\
&= C^{(2)}_{ij} \frac{\partial}{\partial x_m} \frac{\partial}{\partial x_m} (x_i x_j) \\
&= C^{(2)}_{ij} \frac{\partial}{\partial x_m} \left[\delta_{im} x_j + \delta_{jm} x_i \right] \\
&= 2C^{(2)}_{ij} \left[\delta_{im} \delta_{jm} \right] = 2C^{(2)}_{ii} = 0 \ ,
\end{aligned} \tag{9.25}$$

where in the last step we used the all-important fact that $C^{(2)}_{ij}$ is traceless. We will look at one more case, and then I hope it will be clear that it generalizes. For $\ell = 3$,

$$\begin{aligned}
\nabla^2 \tilde{F}_3(\vec{r}) &= \nabla^2 C^{(3)}_{ijk} x_i x_j x_k \\
&= C^{(3)}_{ijk} \frac{\partial}{\partial x_m} \frac{\partial}{\partial x_m} (x_i x_j x_k) \\
&= C^{(3)}_{ijk} \frac{\partial}{\partial x_m} \Big(\delta_{im} x_j x_k + (\text{terms that symmetrize in } ijk) \Big) \\
&= C^{(3)}_{ijk} \Big(\delta_{im} \delta_{jm} x_k + (\text{terms that symmetrize in } ijk) \Big) \\
&= C^{(3)}_{ijk} \Big(\delta_{ij} x_k + (\text{terms that symmetrize in } ijk) \Big) \\
&= C^{(3)}_{iik} x_k + (\text{terms that symmetrize in } ijk) = 0 \ ,
\end{aligned} \tag{9.26}$$

where again it is the tracelessness of $C^{(\ell)}_{i_1 i_2 \ldots i_\ell}$ that caused the term to vanish. Thinking about the general term, one can see that after the derivatives are calculated, there are $\ell - 2$ factors of x_i that remain, but there are still ℓ indices on $C^{(\ell)}_{i_1 i_2 \ldots i_\ell}$. Since all indices are summed, there are always two indices on $C^{(\ell)}_{i_1 i_2 \ldots i_\ell}$ which are contracted (i.e., set equal to each other) and summed, which causes the result to vanish by the tracelessness condition. The bottom line, then, is that

$$\nabla^2 \tilde{F}_\ell(\vec{r}) = 0 \qquad \text{for all } \ell. \tag{9.27}$$

To see what this says about $\nabla^2_\theta F_\ell(\hat{n})$, recall that $\tilde{F}_\ell(\vec{r}) = r^\ell F_\ell(\hat{n})$. Using Eq. (9.1), we can write

$$\begin{aligned}
0 = \nabla^2 \tilde{F}_\ell(\vec{r}) &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \tilde{F}_\ell(\vec{r})}{\partial r}\right) + \frac{1}{r^2}\nabla^2_\theta\, \tilde{F}_\ell(\vec{r}) \\
&= \frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2 \frac{\mathrm{d}r^\ell}{\mathrm{d}r}\right) F_\ell(\hat{n}) + \frac{1}{r^2} r^\ell \nabla^2_\theta\, F_\ell(\hat{n}) \\
&= r^{\ell-2}\Big[\ell(\ell+1)F_\ell(\hat{n}) + \nabla^2_\theta\, F_\ell(\hat{n})\Big] \ ,
\end{aligned}$$

and therefore

$$\boxed{\nabla^2_\theta\, F_\ell(\hat{n}) = -\ell(\ell+1)F_\ell(\hat{n}) \ .} \tag{9.28}$$

Thus, we have found the eigenfunctions $\Big(F_\ell(\hat{n}) = C^{(\ell)}_{i_1 i_2 \ldots i_\ell}\hat{n}_{i_1}\hat{n}_{i_2}\ldots\hat{n}_{i_\ell}\Big)$ and eigenvalues $\Big(-\ell(\ell+1)\Big)$ of the differential operator ∇^2_θ. This is a very useful result!

4. GENERAL SOLUTION TO LAPLACE'S EQUATION IN SPHERICAL COORDINATES:

Now that we know the eigenfunctions of ∇^2_θ, we can return to the solution to Laplace's equation by the separation of variables in spherical coordinates. We now know that in Eq. (9.6), the only allowed values of C_θ are $-\ell(\ell+1)$, where ℓ is an integer. Thus, Eq. (9.7) becomes

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) = \ell(\ell+1)R \ , \tag{9.29}$$

and we can look for solutions by trying $R(r) = r^p$. We find consistency provided that

$$p(p+1) = \ell(\ell+1) \ , \tag{9.30}$$

which is a quadratic equation with two roots, $p = \ell$ and $p = -(\ell+1)$. Since we found two solutions to a second order linear differential equation, we know that any solution can be written as a linear sum of these two. Thus we can write

$$\boxed{R_\ell(r) = A_\ell r^\ell + \frac{B_\ell}{r^{\ell+1}}\ ,} \tag{9.31}$$

allowing for different values of A_ℓ and B_ℓ for each ℓ. The most general solution to Laplace's equation, in spherical coordinates, can then be written as

$$\boxed{\Phi(\vec{r}) = \sum_{\ell=0}^{\infty}\left(A_\ell r^\ell + \frac{B_\ell}{r^{\ell+1}}\right) C^{(\ell)}_{i_1 i_2 \ldots i_\ell}\hat{n}_{i_1}\hat{n}_{i_2}\ldots\hat{n}_{i_\ell}\ ,} \tag{9.32}$$

where the A_ℓ's and B_ℓ's are arbitrary constants, and each $C^{(\ell)}_{i_1 i_2 \ldots i_\ell}$ is an arbitrary traceless symmetric tensor.

In Lecture Notes 8 it is shown explicitly that the ℓ'th term here, when compared with the standard expansion in the spherical harmonic functions $Y_{\ell m}(\theta,\phi)$, corresponds to the sum of all terms with the same ℓ, but for all m. The $Y_{\ell m}$'s are defined for integer values of m from $-\ell$ to ℓ, so there are $2\ell+1$ terms for each value of ℓ.

5. COUNTING THE NUMBER LINEARLY INDEPENDENT TRACELESS SYMMETRIC TENSORS:

Using the fact that $C^{(\ell)}_{i_1 i_2 \ldots i_\ell}$ is traceless and symmetric, we can determine how many linearly independent such tensors exist, for any given ℓ. In other words, how many real constants are needed to parameterize the most general traceless symmetric tensor of rank ℓ?

We begin by calculating $N_{\rm sym}(\ell)$, the number of linearly independent symmetric tensors of rank ℓ. Note that we are postponing the consideration of tracelessness. For grounding, we can start by saying that it takes one number to specify $S^{(0)}$, a rank 0 symmetric tensor, because it is just a number. (I am using S for symmetric tensors, while reserving C for traceless symmetric tensors.) It takes 3 numbers to specify $S^{(1)}_i$, since the 3 values $S^{(1)}_1$, $S^{(1)}_2$, and $S^{(1)}_3$ can each be specified independently. For $S^{(2)}_{ij}$, however, we see the constraints of symmetry: $S^{(2)}_{ij}$ has to equal $S^{(2)}_{ji}$, so there are fewer than 9 independent values; there are 6, which can be taken to be $S^{(2)}_{11}$, $S^{(2)}_{12}$, $S^{(2)}_{13}$, $S^{(2)}_{22}$, $S^{(2)}_{23}$, $S^{(2)}_{33}$, with $S^{(2)}_{21}$, $S^{(2)}_{31}$, and $S^{(2)}_{32}$ determined by symmetry. Since the order of the indices does not matter, we can always list the indices in ascending order (as I did for

$S^{(2)}_{ij}$) and then each independent entry will occur once. When the indices are so written in ascending order, then the index values are completely determined if we just specify how many indices are equal to 1, how many are equal to 2, and how many are equal to 3. If it helps, we can imagine three hats, labeled 1, 2, and 3, and ℓ indistinguishable balls, representing the ℓ indices of a rank ℓ tensor. There is then a 1:1 correspondence between independent tensor elements and the different ways that the balls can be put into the hats. For example, if there are 9 balls, with 3 in the first hat, 2 in the second, and 4 in the third, then this arrangement of balls corresponds to $S^{(9)}_{111223333}$. So now we just have to figure out how many different ways we can put ℓ indistinguishable balls into 3 hats.

One way of counting the balls-in-hats problem is to imagine first labeling each ball with a number, so they are no longer indistinguishable. We also introduce 2 dividers, where 2 is one less than the number of hats. Initially we will also assign numbers to the 2 dividers. Thinking of the balls and dividers together, we have $\ell+2$ distinguishable objects. We can imagine listing them in all possible orderings, and with $\ell+2$ distinguishable objects there are $(\ell+2)!$ orderings. For each ordering there is an equivalent balls-in-hats assignment. The balls to the left of the left-most divider are assigned to hat 1, the balls between the two dividers are assigned to hat 2, and the balls to the right of the right-most divider are assigned to hat 3. We have of course overcounted, since many different orderings of our $\ell+2$ objects will lead to the same number of balls in each hat. However, we can see exactly by how much we have overcounted. We can re-order the ℓ balls without changing the number of balls in each hat, and we can interchange the two dividers. So we have overcounted by a factor of $2\ell!$. Thus,

$$N_{\rm sym}(\ell) = \frac{(\ell+2)!}{2\ell!} = \frac{1}{2}(\ell+1)(\ell+2)\ . \tag{9.33}$$

It is easily checked that this gives $N_{\rm sym}(0)=1$, $N_{\rm sym}(1)=3$, and $N_{\rm sym}(2)=6$, consistent with the examples we started with.

To impose tracelessness, we require that our traceless tensors also satisfy

$$S^{(\ell)}_{i_1\ldots i_{\ell-2}jj} = 0\ . \tag{9.34}$$

How many conditions is this? One can see that the tensor on the left is a symmetric tensor of rank $\ell-2$, with free indices $i_1\ldots i_{\ell-2}$. The number of conditions is then $N_{\rm sym}(\ell-2)$. The number of linearly indendent traceless symmetric tensors is then given by

$$\begin{aligned} N_{\rm traceless-sym}(\ell) &= N_{\rm sym}(\ell) - N_{\rm sym}(\ell-2) = \frac{1}{2}(\ell+1)(\ell+2) - \frac{1}{2}(\ell-1)\ell \\ &= \boxed{\ 2\ell+1\ .\ } \end{aligned} \tag{9.35}$$

So the correspondence with the standard $Y_{\ell m}$'s is consistent, as it would have to be. The spherical harmonic expansion is just a rewriting of Eq. (9.32), with a particular choice of basis for the $2\ell+1$ independent traceless symmetric tensors of rank ℓ.

6. SPECIAL CASE: AZIMUTHAL SYMMETRY:

Azimuthal symmetry means symmetry under rotation about an axis, which we will take to be the z-axis. Equivalently, we can say that a problem is azimuthally symmetric if nothing depends on the coordinate ϕ. To specialize the general expansion (9.10) for a function of (θ, ϕ) to the azimuthally symmetric case, we need to construct traceless symmetric tensors which are invariant under rotations about the z-axis. One straightforward way to this is to build the traceless symmetric tensor from the vector $\hat{z}$, the unit vector in the z direction. Note that the x-, y-, and z- components of $\hat{z}$ are 0, 0, and 1, respectively, so $\hat{z}_i = \delta_{i3}$.

[You may have noticed an inconsistency in my notation, as earlier (e.g., Eq. (9.9)) I used $\hat{e}_3$ or $\hat{e}_z$ for the unit vector in the z direction. In this case my inconsistency was intentional, with two motivations. First, previously we never needed a notation for the components of a unit basis vector, but here we will. It would be a real pain to write $(\hat{e}_z)_i$. Second, in Lecture Notes 8 I describe a convenient way to construct a basis for the traceless symmetric tensors, which involves the use of a basis for vectors consisting of $\hat{z}$ and two complex vectors $\hat{u}^+$ and $\hat{u}^-$. So, the use of $\hat{z}$ rather than $\hat{e}_z$ will remind us that we are thinking about the $(\hat{u}^+, \hat{u}^-, \hat{z})$ basis, rather than the $(\hat{e}_x, \hat{e}_y, \hat{e}_z)$ basis.]

A rank ℓ tensor can be constructed from $\hat{z}$ simply by taking the product $\hat{z}_{i_1} \hat{z}_{i_2} \ldots \hat{z}_{i_\ell}$. This is symmetric, and can be made traceless by extracting the traceless part. Extracting the traceless part means subtracting terms proportional to one or more Kronecker δ-functions in such a way that the result is traceless. It gets rather complicated to describe how this can be done for a general symmetric tensor of arbitrary rank, so I will just illustrate it by example. Tensors of rank 0 and 1 (i.e., scalars and vectors) are by definition traceless. I will use curly brackets $\{\ldots\}$ to denote the traceless symmetric part of .... Thus,

$$\boxed{\begin{aligned} \{\, 1 \,\} &= 1 \ , \\ \{\, \hat{z}_i \,\} &= \hat{z}_i \ . \end{aligned}} \tag{9.36}$$

But for rank 2, the trace of $\hat{z}_i \hat{z}_j$ is equal to $\hat{z}_i \hat{z}_i = \hat{z} \cdot \hat{z} = 1$. But we can subtract a constant times δ_{ij} so that the result is traceless:

$$\boxed{\{\, \hat{z}_i \hat{z}_j \,\} = \hat{z}_i \hat{z}_j - \tfrac{1}{3} \delta_{ij} \ .} \tag{9.37}$$

The coefficient is 1/3, because the trace of δ_{ij} is

$$\delta_{ii} = 3 \ . \tag{9.38}$$

For rank 3, $\hat{z}_i\hat{z}_j\hat{z}_k$ has trace $\hat{z}_i\hat{z}_i\hat{z}_k = \hat{z}_k$, but we can make it traceless with a subtraction

$$\boxed{\{\, \hat{z}_i\hat{z}_j\hat{z}_k \,\} = \hat{z}_i\hat{z}_j\hat{z}_k - \tfrac{1}{5}\left(\hat{z}_i\delta_{jk} + \hat{z}_j\delta_{ik} + \hat{z}_k\delta_{ij}\right) \ .} \tag{9.39}$$

The subtraction must of course be symmetrized, as shown, since we are trying to construct a traceless symmetric tensor. To verify that 1/5 is the right coefficient to make the expression traceless, we can take its trace. Since it is symmetric we can sum over any pair of indices. I will choose to sum over i and j:

$$\begin{aligned} \delta_{ij}\{\, \hat{z}_i\hat{z}_j\hat{z}_k \,\} &= \hat{z}_i\hat{z}_i\hat{z}_k - \tfrac{1}{5}\left(\hat{z}_i\delta_{ik} + \hat{z}_i\delta_{ik} + \hat{z}_k\delta_{ii}\right) \\ &= \hat{z}_k - \tfrac{1}{5}(\hat{z}_k + \hat{z}_k + 3\hat{z}_k) = 0 \ . \end{aligned} \tag{9.40}$$

For rank 4 there is the option of subtracting terms with either one or two Kronecker δ-functions, and both are needed to give a traceless result. We can start with arbitrary coefficients, and see what they have to be:

$$\begin{aligned} \{\, \hat{z}_i\hat{z}_j\hat{z}_k\hat{z}_m \,\} = \hat{z}_i\hat{z}_j\hat{z}_k\hat{z}_m &+ c_1\big(\hat{z}_i\hat{z}_j\delta_{km} + \hat{z}_i\hat{z}_k\delta_{mj} + \hat{z}_i\hat{z}_m\delta_{jk} + \hat{z}_j\hat{z}_k\delta_{im} \\ &+ \hat{z}_j\hat{z}_m\delta_{ik} + \hat{z}_k\hat{z}_m\delta_{ij}\big) + c_2\big(\delta_{ij}\delta_{km} + \delta_{ik}\delta_{jm} + \delta_{im}\delta_{jk}\big) \ . \end{aligned} \tag{9.41}$$

Within each set of parentheses, the terms are chosen to make the expression symmetric in i, j, k, and m. If we calculate the trace over i and j, we find

$$\begin{aligned} \delta_{ij}\{\, \hat{z}_i\hat{z}_j\hat{z}_k\hat{z}_m \,\} = \hat{z}_k\hat{z}_m &+ c_1\big(\delta_{km} + \hat{z}_k\hat{z}_m + \hat{z}_k\hat{z}_m + \hat{z}_k\hat{z}_m \\ &+ \hat{z}_k\hat{z}_m + 3\hat{z}_k\hat{z}_m\big) + c_2\big(3\delta_{km} + \delta_{km} + \delta_{km}\big) \\ = \big(1 + 7c_1\big)&\hat{z}_k\hat{z}_m + \big(c_1 + 5c_2\big)\delta_{km} \ . \end{aligned} \tag{9.42}$$

For the expression to vanish for all k and m, the two terms must vanish separately, so $c_1 = -1/7$ and $c_2 = 1/35$. Thus,

$$\boxed{\begin{aligned} \{\, \hat{z}_i\hat{z}_j\hat{z}_k\hat{z}_m \,\} = \hat{z}_i\hat{z}_j\hat{z}_k\hat{z}_m &- \tfrac{1}{7}\big(\hat{z}_i\hat{z}_j\delta_{km} + \hat{z}_i\hat{z}_k\delta_{mj} + \hat{z}_i\hat{z}_m\delta_{jk} + \hat{z}_j\hat{z}_k\delta_{im} \\ &+ \hat{z}_j\hat{z}_m\delta_{ik} + \hat{z}_k\hat{z}_m\delta_{ij}\big) + \tfrac{1}{35}\big(\delta_{ij}\delta_{km} + \delta_{ik}\delta_{jm} + \delta_{im}\delta_{jk}\big) \ . \end{aligned}} \tag{9.43}$$

It can be shown that any traceless symmetric tensor of rank ℓ that is invariant under rotations about the z-axis is proportional to $\{\, \hat{z}_{i_1} \ldots \hat{z}_{i_\ell} \,\}$. (To see this, note that in Lecture Notes 8 we construct a complete $(2\ell+1)$-dimensional basis for the traceless symmetry tensors of rank ℓ. They depend on the azimuthal angle ϕ as $z_m(\phi) \equiv e^{im\phi}$, with m taking integer values from $-\ell$ to ℓ. Since these functions of ϕ are orthogonal in the sense that $\int_0^{2\pi} z^*_{m'}(\phi) z_m(\phi)\,\mathrm{d}\phi = 2\pi\delta_{m'm}$, any traceless symmetric tensor of rank ℓ that is independent of ϕ must proportional to the $m = 0$ basis tensor.) Since Eq. (9.10) tells us how to expand an arbitrary function of θ and ϕ in terms of traceless symmetric tensors, we can now say that functions of θ alone (i.e., azimuthally symmetric functions of $\hat{n}$) can be expanded as

$$\boxed{F(\theta) = c_0 + c_1\{\,\hat{z}_i\,\}\hat{n}_i + c_2\{\,\hat{z}_i\hat{z}_j\,\}\,\hat{n}_i\hat{n}_j + \ldots + c_\ell\{\,\hat{z}_{i_1}\ldots\hat{z}_{i_\ell}\,\}\,\hat{n}_{i_1}\ldots\hat{n}_{i_\ell} + \ldots\ ,} \tag{9.44}$$

where the c_ℓ's are constants. This corresponds to what is standardly called an expansion in Legendre polynomials. In Lecture Notes 8 I show exactly how to relate these terms to the standard conventions for normalizing the Legendre polynomials, but we can see here exactly what these functions are. Using $\hat{z}\cdot\hat{n} = \cos\theta$, we have

$$\boxed{\begin{aligned}
\{\,1\,\} &= 1 \\
\{\,\hat{z}_i\,\}\,\hat{n}_i &= \cos\theta \\
\{\,\hat{z}_i\hat{z}_j\,\}\,\hat{n}_i\hat{n}_j &= \cos^2\theta - \frac{1}{3} \\
\{\,\hat{z}_i\hat{z}_j\hat{z}_k\,\}\,\hat{n}_i\hat{n}_j\hat{n}_k &= \cos^3\theta - \frac{3}{5}\cos\theta \\
\{\,\hat{z}_i\hat{z}_j\hat{z}_k\hat{z}_m\,\}\,\hat{n}_i\hat{n}_j\hat{n}_k\hat{n}_m &= \cos^4\theta - \frac{6}{7}\cos^2\theta + \frac{3}{35}\ .
\end{aligned}} \tag{9.45}$$

Up to a normalization convention described in Lecture Notes 8, these are the Legendre polynomials $P_\ell(\cos\theta)$.

7. THE MULTIPOLE EXPANSION:

The most general solution to Laplace's equation, in spherical coordinates, was given as Eq. (9.32). We now wish to apply that result to a common situation: suppose we have a charged object, and we wish to describe the potential outside of the object. Let's say for definiteness that the charge of the object is entirely contained within a sphere of radius R, centered at the origin. In that case Laplace's equation will hold for all $r > R$, so there should be a solution of the form of Eq. (9.32) that is valid throughout this region. At infinity the potential of a localized charge distribution will always approach a constant, which we can take to be zero, we can see that the A_ℓ coefficients that appear in Eq. (9.32) must all vanish. The B_ℓ factors can be absorbed into the definition of $C^{(\ell)}_{i_1\ldots i_\ell}$, so we can write the expansion as

$$\boxed{\Phi(\vec{r}) = \sum_{\ell=0}^{\infty} \frac{1}{r^{\ell+1}}\, C^{(\ell)}_{i_1\ldots i_\ell}\hat{n}_{i_1}\ldots\hat{n}_{i_\ell}\ .} \tag{9.46}$$

Since each successive term comes with an extra factor of $1/r$, at large distances the sum is dominated by the first term or maybe the first few terms. All the information about the charge distribution of the object is contained in the $C^{(\ell)}_{i_1\ldots i_\ell}$, so knowledge of the first few $C^{(\ell)}_{i_1\ldots i_\ell}$ is enough to describe the field at large distances, no matter how complicated the object.

The first few terms of this series have special names: the $\ell = 0$ term is the monopole term, the $\ell = 1$ term is the dipole, the $\ell = 2$ term is the quadrupole, and the $\ell = 3$ term is the octupole.

If we want to calculate the $C^{(\ell)}_{i_1\ldots i_\ell}$ in terms of the charge distribution, we can start with the general equation for the potential of an arbitrary charge distribution:

$$V(\vec{r}) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|}\,\mathrm{d}^3x \ . \tag{9.47}$$

The multipole expansion can then be derived by expanding $1/\left|\vec{r} - \vec{r}\,'\right|$ in a power series in $\vec{r}\,'$.

I'll begin by doing it as Griffiths does, which gives the simplest — but not the most useful — form of the multipole expansion. Griffiths rewrote the denominator as

$$\frac{1}{|\vec{r} - \vec{r}\,'|} = \frac{1}{\sqrt{|\vec{r}|^2 + |\vec{r}\,'|^2 - 2\vec{r}\cdot\vec{r}\,'}} = \frac{1}{\sqrt{r^2 + r'^2 - 2rr'\cos\theta'}} \ , \tag{9.48}$$

where r and r' are the lengths of the vectors $\vec{r}$ and $\vec{r}\,'$, respectively, and θ is the angle between these vectors. Next he used the fact that the Legendre polynomials can be defined by the generating function

$$g(x,\lambda) = \frac{1}{\sqrt{1 + \lambda^2 - 2\lambda x}} \ , \tag{9.49}$$

which means that the Legendre polynomials $P_\ell(x)$ can be obtained by expanding $g(x,\lambda)$ in a power series in λ:

$$\boxed{g(x,\lambda) = \frac{1}{\sqrt{1 + \lambda^2 - 2\lambda x}} = \sum_{\ell=0}^{\infty} \lambda^\ell P_\ell(x) \ .} \tag{9.50}$$

Eq. (9.50) is sometimes taken as the definition of the Legendre polynomials, and sometimes it is derived from another definition. In any case, if we accept Eq. (9.50) as valid, then

$$\frac{1}{|\vec{r} - \vec{r}\,'|} = \frac{1}{r\sqrt{1 + \left(\frac{r'}{r}\right)^2 - 2\frac{r'}{r}\cos\theta'}} = \frac{1}{r}\sum_{\ell=0}^{\infty}\left(\frac{r'}{r}\right)^{\ell} P_\ell(\cos\theta) \ . \tag{9.51}$$

Inserting this relation into Eq. (9.47), we find

$$\boxed{V(\vec{r}) = \frac{1}{4\pi\epsilon_0}\sum_{\ell=0}^{\infty}\frac{1}{r^{\ell+1}}\int {r'}^{\ell}\rho(\vec{r}\,')P_\ell(\cos\theta')\,\mathrm{d}^3x \ .} \tag{9.52}$$

This is the easiest way that I know to show that there is an expansion of $V(\vec{r})$ in powers of $1/r$, but the complication is that $\cos\theta'$ appears inside the integral. If we could implement Eq. (9.46), we would be able to calculate (or maybe measure) a small number of the quantities $C^{(\ell)}_{i_1\ldots i_\ell}$, and then we would be able to evaluate $V(\vec{r})$ at large distances in any direction. To use Eq. (9.52) directly, however, one would have to repeat the integration for every direction of $\vec{r}$. Griffiths works around this problem by massaging the formula to extract the monopole and dipole terms, and in Problem 3 of Problem Set 5 you had the opportunity to carry this out for the quadrupole and octupole terms.

The standard method of "improving" Eq. (9.52) is to use spherical harmonics, but here I will derive the equivalent relations using the traceless symmetric tensor approach.

Instead of expanding $1/|\vec{r}-\vec{r}\,'|$ in powers of r', we will think of it as a function of three variables — the components $\vec{r}\,'_i$ of $\vec{r}\,'$, and we will expand it as a Taylor series in 3 variables. To make the formalism clear, I will define the function

$$f(\vec{r}\,') \equiv \frac{1}{|\vec{r}-\vec{r}\,'|} \ . \tag{9.53}$$

The function can then be expanded in a power series using the standard multi-variable Taylor expansion:

$$f(\vec{r}\,') = f(\vec{0}) + \left.\frac{\partial f}{\partial x'_i}\right|_{\vec{r}\,'=\vec{0}} x'_i + \frac{1}{2!}\left.\frac{\partial^2 f}{\partial x'_i\partial x'_j}\right|_{\vec{r}\,'=\vec{0}} x'_i x'_j + \ldots \ , \tag{9.54}$$

where the repeated indices are summed. To separate the angular behavior, we write

$$x'_i = r'\hat{n}'_i \ , \tag{9.55}$$

so Eq. (9.54) becomes

$$f(\vec{r}\,') = f(\vec{0}) + r'\left.\frac{\partial f}{\partial x'_i}\right|_{\vec{r}\,'=\vec{0}} \hat{n}'_i + \frac{r'^2}{2!}\left.\frac{\partial^2 f}{\partial x'_i\partial x'_j}\right|_{\vec{r}\,'=\vec{0}} \hat{n}'_i\hat{n}'_j + \ldots \ . \tag{9.56}$$

The notation can now be simplified by noting that since f is a function of $\vec{r}-\vec{r}\,'$, the derivatives with respect to x'_i can be replaced by derivatives with respect to x_i with a change of sign:

$$\frac{\partial f}{\partial x'_i} = \left.\frac{\partial}{\partial x'_i}\left(\frac{1}{|\vec{r}-\vec{r}\,'|}\right)\right|_{\vec{r}\,'=\vec{0}} = -\left.\frac{\partial}{\partial x_i}\left(\frac{1}{|\vec{r}-\vec{r}\,'|}\right)\right|_{\vec{r}\,'=\vec{0}} = -\frac{\partial}{\partial x_i}\left(\frac{1}{|\vec{r}|}\right) \ . \tag{9.57}$$

This allows us to write the derivatives in the expansion (9.56) much more simply. The ℓ'th derivative is found by repeating the above operation ℓ times:

$$\left.\frac{\partial^\ell f}{\partial x'_{i_1}\ldots\partial x'_{i_\ell}}\right|_{\vec{r}\,'=\vec{0}} = (-1)^\ell \frac{\partial^\ell}{\partial x_{i_1}\ldots\partial x_{i_\ell}}\frac{1}{|\vec{r}|} \ . \tag{9.58}$$

Combining Eqs. (9.58) with (9.56), we can write

$$\frac{1}{|\vec{r}-\vec{r}\,'|} = \sum_{\ell=0}^{\infty}\frac{(-1)^\ell r'^\ell}{\ell!}\left(\frac{\partial^\ell}{\partial x_{i_1}\ldots\partial x_{i_\ell}}\frac{1}{|\vec{r}|}\right)\hat{n}'_{i_1}\ldots\hat{n}'_{i_\ell} \ . \tag{9.59}$$

Note that the quantity in parentheses in the equation above is traceless, because

$$\begin{aligned}\left(\frac{\partial^\ell}{\partial x_i\partial x_i\partial x_{i+2}\ldots\partial x_{i_\ell}}\frac{1}{|\vec{r}|}\right) &= \nabla^2\frac{\partial^\ell}{\partial x_{i+2}\ldots\partial x_{i_\ell}}\frac{1}{|\vec{r}|}\\ &= \frac{\partial^\ell}{\partial x_{i+2}\ldots\partial x_{i_\ell}}\nabla^2\frac{1}{|\vec{r}|} = 0 \ ,\end{aligned} \tag{9.60}$$

because $\nabla^2(1/|\vec{r}|) = 0$ except at $\vec{r} = 0$. So we can see the traceless symmetric tensor formalism emerging.

To evaluate this quantity, we will work out the first several terms until we recognize the pattern. We write

$$\vec{r} \equiv r\hat{n} \ , \tag{9.61}$$

and adopt the abbreviation

$$\partial_i \equiv \frac{\partial}{x_i} \ . \tag{9.62}$$

It is useful to start by evaluating the derivatives of the basic quantities r and $\hat{n}_i$:

$$\begin{aligned}\partial_i r &= \partial_i(x_j x_j)^{1/2} = \frac{1}{2}(x_k x_k)^{-1/2}\partial_i(x_j x_j) = \frac{1}{2r}2x_j\delta_{ij} = \frac{x_i}{r} = \hat{n}_i \ ,\\ \partial_i\hat{n}_j &= \partial_i\left(\frac{x_j}{r}\right) = \frac{\delta_{ij}}{r} - \frac{1}{r^2}x_j\partial_i r = \frac{1}{r}\left(\delta_{ij} - \hat{n}_i\hat{n}_j\right) \ .\end{aligned} \tag{9.63}$$

It is then straightforward to show that

$$\begin{aligned}\partial_i\left(\frac{1}{r}\right) &= -\frac{1}{r^2}\hat{n}_i \ ,\\ \partial_i\partial_j\left(\frac{1}{r}\right) &= \frac{3}{r^3}\{\,\hat{n}_i\hat{n}_j\,\} \ ,\\ \partial_i\partial_j\partial_k\left(\frac{1}{r}\right) &= -\frac{5\cdot 3}{r^4}\{\,\hat{n}_i\hat{n}_j\hat{n}_k\,\} \ ,\end{aligned} \tag{9.64}$$

where { } denotes the traceless symmetric part, and the relevant cases are shown explicitly in Eqs. (9.36) – (9.39). It becomes clear that the general formula, which can be proven by induction, is

$$\frac{\partial^\ell}{\partial x_{i_1} \ldots \partial x_{i_\ell}} \frac{1}{|\vec{r}|} = \frac{(-1)^\ell (2\ell-1)!!}{r^{\ell+1}} \{ \hat{n}_{i_i} \ldots \hat{n}_{i_\ell} \} , \tag{9.65}$$

where

$$(2\ell-1)!! \equiv (2\ell-1)(2\ell-3)(2\ell-5)\ldots 1 = \frac{(2\ell)!}{2^\ell \ell!} , \text{ with } (-1)!! \equiv 1 . \tag{9.66}$$

Inserting this result into Eq. (9.59), we find

$$\frac{1}{|\vec{r}-\vec{r}\,'|} = \sum_{\ell=0}^{\infty} \frac{(2\ell-1)!!}{\ell!} \frac{r'^\ell}{r^{\ell+1}} \{ \hat{n}_{i_1} \ldots \hat{n}_{i_\ell} \} \hat{n}'_{i_1} \ldots \hat{n}'_{i_\ell} . \tag{9.67}$$

One can write this more symmetrically by writing

$$\boxed{\frac{1}{|\vec{r}-\vec{r}\,'|} = \sum_{\ell=0}^{\infty} \frac{(2\ell-1)!!}{\ell!} \frac{r'^\ell}{r^{\ell+1}} \{ \hat{n}_{i_1} \ldots \hat{n}_{i_\ell} \} \{ \hat{n}'_{i_1} \ldots \hat{n}'_{i_\ell} \} ,} \tag{9.67}$$

since $\{ \hat{n}'_{i_1} \ldots \hat{n}'_{i_\ell} \}$ differs from $\hat{n}'_{i_1} \ldots \hat{n}'_{i_\ell}$ by terms proportional to Kronecker δ-functions, which vanish when summed with the traceless tensor $\{ \hat{n}_{i_1} \ldots \hat{n}_{i_\ell} \}$. Starting with Eq. (9.67), one can if one wishes drop the curly brackets around either factor (but not both!).

Inserting this expression for $1/|\vec{r}-\vec{r}\,'|$ into Eq. (9.47), we have the final result

$$\boxed{V(\vec{r}) = \frac{1}{4\pi\epsilon_0} \sum_{\ell=0}^{\infty} \frac{1}{r^{\ell+1}} C^{(\ell)}_{i_1 \ldots i_\ell} \hat{n}_{i_1} \ldots \hat{n}_{i_\ell} ,} \tag{9.68}$$

where

$$C^{(\ell)}_{i_1 \ldots i_\ell} = \frac{(2\ell-1)!!}{\ell!} \int \rho(\vec{r}\,') \{ \vec{r}\,'_{i_1} \ldots \vec{r}\,'_{i_\ell} \} \, \mathrm{d}^3 x' . \tag{9.69}$$

Note that the coefficient in the above expression can also be written as

$$\frac{(2\ell-1)!!}{\ell!} = \frac{(2\ell)!}{2^\ell (\ell!)^2} . \tag{9.70}$$

For purposes of illustration, I will write out the first two terms — the monopole and dipole terms — in a bit more detail. The monopole term can be written as

$$V_{\rm mono}(\vec{r}) = \frac{1}{4\pi\epsilon_0}\frac{Q}{r} \ , \tag{9.71}$$

where

$$Q = C^{(0)} = \int \rho(\vec{r}\,') \, {\rm d}^3x' \ . \tag{9.72}$$

The dipole term is

$$V_{\rm dip}(\vec{r}) = \frac{1}{4\pi\epsilon_0}\frac{\vec{p}\cdot\hat{n}}{r^2} \ , \tag{9.73}$$

where

$$p_i = C_i^{(1)} = \int \rho(\vec{r}\,')\vec{r}_i \, {\rm d}^3x \ . \tag{9.74}$$