LECTURE NOTES 1

Revised October17, 2019 *

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 we will also show how these expressions are related to the standard expressions in terms of Legendre polynomials and spherical harmonics. 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. Another good source is David J. Griffiths' book on quantum mechanics, **Introduction to Quantum Mechanics**, 2nd Edition (Pearson Prentice Hall, 2004). But our intention is that the lectures, these notes, plus the recitation sections should give you a complete picture, without the need for these other references.

---

 * Added what is now Eqs. (1.39), and fixed a typo in what is now Eq. (1.40), changing the last subscript from $i_{\ell-2}$ to i_ℓ.

1. LAPLACE'S EQUATION IN SPHERICAL COORDINATES:

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

$$
\nabla^2 \varphi(r, \theta, \phi) = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \varphi}{\partial r} \right) + \frac{1}{r^2} \nabla^2_{\text{ang}} \varphi = 0 ,
\tag{1.1}
$$

where the angular part is given by

$$
\nabla^2_{\text{ang}} \varphi \equiv \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \varphi}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 \varphi}{\partial \phi^2} .
\tag{1.2}
$$

It is sometimes useful to rewrite the first term in Eq. (1.1) using

$$
\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \varphi}{\partial r} \right) = \frac{1}{r} \frac{\partial^2}{\partial r^2} (r\varphi) .
\tag{1.3}
$$

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

$$
\varphi(r, \theta, \phi) = R(r) F(\theta, \varphi) .
\tag{1.4}
$$

Then Laplace's equation can be written as

$$
0 = \frac{r^2}{RF} \nabla^2 \varphi = \frac{1}{R} \frac{\mathrm{d}}{\mathrm{d}r} \left(r^2 \frac{\mathrm{d}R}{\mathrm{d}r} \right) + \frac{1}{F} \nabla^2_{\text{ang}} F .
\tag{1.5}
$$

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

$$
\frac{1}{F} \nabla^2_{\text{ang}} F = C_{\text{ang}} ,
\tag{1.6}
$$

$$
\frac{1}{R} \frac{\mathrm{d}}{\mathrm{d}r} \left(r^2 \frac{\mathrm{d}R}{\mathrm{d}r} \right) = -C_{\text{ang}} .
\tag{1.7}
$$

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

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

$$\nabla_{\text{ang}}^2 F = C_{\text{ang}} F , \tag{1.8}$$

which is a rewriting of Eq. (1.6). If such a function F can be found, we say that F is an *eigenfunction* of the operator ∇_{ang}^2, with *eigenvalue C_{ang}*.

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

$$\hat{\boldsymbol{n}} = \sin\theta \cos\phi\, \hat{\boldsymbol{e}}_{\boldsymbol{1}} + \sin\theta \sin\phi\, \hat{\boldsymbol{e}}_{\boldsymbol{2}} + \cos\theta\, \hat{\boldsymbol{e}}_{\boldsymbol{3}} , \tag{1.9}$$

where $\hat{\boldsymbol{e}}_{\boldsymbol{1}}$, $\hat{\boldsymbol{e}}_{\boldsymbol{2}}$, and $\hat{\boldsymbol{e}}_{\boldsymbol{3}}$ are the unit vectors in the x, y, and z directions, respectively, and can also be written as $\hat{\boldsymbol{e}}_{\boldsymbol{x}}$, $\hat{\boldsymbol{e}}_{\boldsymbol{y}}$, and $\hat{\boldsymbol{e}}_{\boldsymbol{z}}$, or as $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$. 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 $\hat{\boldsymbol{n}}$, or more precisely as a power series in the components of $\hat{\boldsymbol{n}}$. 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

$$\boxed{F(\hat{\boldsymbol{n}}) = C^{(0)} + C_i^{(1)}\hat{\boldsymbol{n}}_i + C_{ij}^{(2)}\hat{\boldsymbol{n}}_i\hat{\boldsymbol{n}}_j + \ldots + C_{i_1 i_2 \ldots i_\ell}^{(\ell)}\hat{\boldsymbol{n}}_{i_1}\hat{\boldsymbol{n}}_{i_2}\ldots\hat{\boldsymbol{n}}_{i_\ell} + \ldots ,} \tag{1.10}$$

where repeated indices are summed from 1 to 3 (as Cartesian coordinates). Note that $C_{i_1 i_2 \ldots i_\ell}^{(\ell)}\hat{\boldsymbol{n}}_{i_1}\hat{\boldsymbol{n}}_{i_2}\ldots\hat{\boldsymbol{n}}_{i_\ell}$ 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 $C_{i_1 i_2 \ldots i_\ell}^{(\ell)}$ are called *tensors*, and the number of indices is called the rank of the tensor. Note that $C^{(0)}$, $C_i^{(1)}$, and $C_{ij}^{(2)}$ 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. (1.10) without actually restricting what can appear on the right-hand side of the equations. In particular, we will insist that

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

$$C_{i_1 i_2 \ldots i_\ell}^{(\ell)} = C_{j_1 j_2 \ldots j_\ell}^{(\ell)} , \tag{1.11}$$

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:

$$C^{(\ell)}_{i_1 i_2 \ldots i_{\ell-2} j j} = 0 \ . \tag{1.12}$$

To explain why these restrictions on the $C^{(\ell)}$'s do not impose any restriction on the right-hand side of Eq. (1.10), 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. (1.10), because $C^{(2)}_{ij}$ multiplies the symmetric tensor $\hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_j$. Thus, if $C^{(2)}_{ij}$ had an antisymmetric part, it would not contribute to the right-hand side of Eq. (1.10). The requirement of tracelessness is less obvious, but suppose that $C^{(2)}_{ij}$ were not traceless. Then we could write

$$C^{(2)}_{ii} = \lambda \neq 0 \ . \tag{1.13}$$

We could then define a new quantity,

$$\tilde{C}^{(2)}_{ij} = C^{(2)}_{ij} - \frac{1}{3} \lambda \delta_{ij} \ . \tag{1.14}$$

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

$$\tilde{C}^{(2)}_{ii} = C^{(2)}_{ii} - \frac{1}{3} \lambda \delta_{ii} = \lambda - \frac{1}{3} \lambda \delta_{ii} = 0 \ , \tag{1.15}$$

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

$$C^{(2)}_{ij} \hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_j = \left[\tilde{C}^{(2)}_{ij} + \frac{1}{3} \lambda \delta_{ij} \right] \hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_j = \tilde{C}^{(2)}_{ij} \hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_j + \frac{1}{3} \lambda \ , \tag{1.16}$$

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

$$\tilde{C}^{(0)} = C^{(0)} + \frac{1}{3} \lambda \ . \tag{1.17}$$

Finally, we can write

$$C^{(0)} + C^{(1)}_i \hat{\boldsymbol{n}}_i + C^{(2)}_{ij} \hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_j = \tilde{C}^{(0)} + C^{(1)}_i \hat{\boldsymbol{n}}_i + \tilde{C}^{(2)}_{ij} \hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_j \ , \tag{1.18}$$

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

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

$$F_\ell(\hat{\boldsymbol{n}}) = C^{(\ell)}_{i_1 i_2 \ldots i_\ell} \hat{\boldsymbol{n}}_{i_1} \hat{\boldsymbol{n}}_{i_2} \ldots \hat{\boldsymbol{n}}_{i_\ell} \ . \tag{1.19}$$

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

To evaluate $\nabla^2_{\mathrm{ang}} F_\ell(\hat{n})$, 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

$$\vec{r} = r\hat{n} \ . \tag{1.20}$$

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

$$\vec{r} = x_i \hat{e}_i = x_1 \hat{e}_1 + x_2 \hat{e}_2 + x_3 \hat{e}_3 \ . \tag{1.21}$$

Then, given any $F_\ell(\hat{n})$ of the form given in Eq. (1.19), we can evaluate the same function, but with argument $\vec{r}$:

$$F_\ell(\vec{r}) = C^{(\ell)}_{i_1 i_2 \ldots i_\ell} x_{i_1} x_{i_2} \ldots x_{i_\ell} = r^\ell F_\ell(\hat{n}) \ . \tag{1.22}$$

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 $F_\ell(\vec{r})$ 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. (1.1), which relates the full Laplacian ∇^2 to the angular Laplacian, ∇^2_{ang}. We will find that in this case the full Laplacian and the radial derivative piece of Eq. (1.1) will both be simple, so we will be able to determine the angular Laplacian by evaluating the other terms in Eq. (1.1).

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

$$\nabla^2 F_0(\vec{r}) = \nabla^2 C_{(0)} = 0 \ , \tag{1.23}$$

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

$$\nabla^2 F_1(\vec{r}) = \nabla^2 C^{(1)}_i x_i = 0 \ , \tag{1.24}$$

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

$$\begin{aligned}
\nabla^2 F_2(\vec{r}) &= \nabla^2 C^{(2)}_{ij} x_i x_j \\
&= C^{(2)}_{ij} \frac{\partial}{\partial x_m} \frac{\partial}{\partial x_m} (x_i x_j) \\
&= C^{(2)}_{ij} \frac{\partial}{\partial x_m} [\delta_{im} x_j + \delta_{jm} x_i] \\
&= 2 C^{(2)}_{ij} [\delta_{im} \delta_{jm}] = 2 C^{(2)}_{ii} = 0 \ ,
\end{aligned} \tag{1.25}$$

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

$$
\begin{aligned}
\nabla^2 F_3(\vec{r}) &= \nabla^2 C_{ijk}^{(3)} x_i x_j x_k \\
&= C_{ijk}^{(3)} \frac{\partial}{\partial x_m} \frac{\partial}{\partial x_m} (x_i x_j x_k) \\
&= C_{ijk}^{(3)} \frac{\partial}{\partial x_m} \Big(\delta_{im} x_j x_k + (\text{terms that symmetrize in } ijk) \Big) \\
&= C_{ijk}^{(3)} \Big(\delta_{im} \delta_{jm} x_k + (\text{terms that symmetrize in } ijk) \Big) \\
&= C_{ijk}^{(3)} \Big(\delta_{ij} x_k + (\text{terms that symmetrize in } ijk) \Big) \\
&= C_{iik}^{(3)} x_k + (\text{terms that symmetrize in } ijk) = 0 ,
\end{aligned}
\tag{1.26}
$$

where again it is the tracelessness of $C_{i_1 i_2 \ldots i_\ell}^{(\ell)}$ 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 $C_{i_1 i_2 \ldots i_\ell}^{(\ell)}$. Since all indices are summed, there are always two indices on $C_{i_1 i_2 \ldots i_\ell}^{(\ell)}$ 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

$$
\nabla^2 F_\ell(\vec{r}) = 0 \qquad \text{for all } \ell.
\tag{1.27}
$$

To see what this says about $\nabla_{\text{ang}}^2 F_\ell(\hat{n})$, recall that $F_\ell(\vec{r}) = r^\ell F_\ell(\hat{n})$. Using Eq. (1.1), we can write

$$
\begin{aligned}
0 = \nabla^2 F_\ell(\vec{r}) &= \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial F_\ell(\vec{r})}{\partial r} \right) + \frac{1}{r^2} \nabla_{\text{ang}}^2 F_\ell(\vec{r}) \\
&= \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dr^\ell}{dr} \right) F_\ell(\hat{n}) + \frac{1}{r^2} r^\ell \nabla_{\text{ang}}^2 F_\ell(\hat{n}) \\
&= r^{\ell-2} \Big[\ell(\ell+1) F_\ell(\hat{n}) + \nabla_{\text{ang}}^2 F_\ell(\hat{n}) \Big] ,
\end{aligned}
$$

and therefore

$$
\boxed{\nabla_{\text{ang}}^2 F_\ell(\hat{n}) = -\ell(\ell+1) F_\ell(\hat{n}) .}
\tag{1.28}
$$

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

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

Now that we know the eigenfunctions of ∇^2_{ang}, we can return to the solution to Laplace's equation by the separation of variables in spherical coordinates. We now know that in Eq. (1.6), the only allowed values of C_{ang} are $-\ell(\ell+1)$, where ℓ is an integer. Thus, Eq. (1.7) becomes

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) = \ell(\ell+1)R \ , \tag{1.29}$$

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

$$p(p+1) = \ell(\ell+1) \ , \tag{1.30}$$

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

$$\boxed{R_\ell(r) = r^\ell \ \text{ or } \ R_\ell(r) = \frac{1}{r^{\ell+1}} \ .} \tag{1.31}$$

Each of the radial functions can be combined with its own angular function $F_\ell(\hat{\boldsymbol{n}})$, so the most general solution to Laplace's equation, in spherical coordinates, can then be written as

$$\boxed{\Phi(\vec{\boldsymbol{r}}) = \sum_{\ell=0}^{\infty}\left(C^{(\ell)}_{i_1 i_2 \ldots i_\ell} r^\ell + \frac{C'^{(\ell)}_{i_1 i_2 \ldots i_\ell}}{r^{\ell+1}}\right)\hat{\boldsymbol{n}}_{i_1}\hat{\boldsymbol{n}}_{i_2}\ldots\hat{\boldsymbol{n}}_{i_\ell} \ ,} \tag{1.32}$$

where $C^{(\ell)}_{i_1 i_2 \ldots i_\ell}$ and $C'^{(\ell)}_{i_1 i_2 \ldots i_\ell}$ are arbitrary traceless symmetric tensors, and $\vec{\boldsymbol{r}} = r\hat{\boldsymbol{n}}$.

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 $N_{\text{sym}}(\ell)$, 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 $S_i^{(1)}$, since the 3 values $S_1^{(1)}$, $S_2^{(1)}$, and $S_3^{(1)}$ can each be specified independently. For $S_{ij}^{(2)}$, however, we see the constraints of symmetry: $S_{ij}^{(2)}$ has to equal $S_{ji}^{(2)}$, so there are fewer than 9 independent values; there are 6, which can be taken to be $S_{11}^{(2)}$, $S_{12}^{(2)}$, $S_{13}^{(2)}$, $S_{22}^{(2)}$, $S_{23}^{(2)}$, $S_{33}^{(2)}$, with $S_{21}^{(2)}$, $S_{31}^{(2)}$, and $S_{32}^{(2)}$ 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 $S_{ij}^{(2)}$) 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 $S_{111223333}^{(9)}$. 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,

$$N_{\text{sym}}(\ell) = \frac{(\ell + 2)!}{2\ell!} = \frac{1}{2}(\ell + 1)(\ell + 2) \ . \tag{1.33}$$

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

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

$$S_{i_1 \ldots i_{\ell-2} jj}^{(\ell)} = 0 \ . \tag{1.34}$$

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 $N_{\text{sym}}(\ell - 2)$.

The number of linearly indendent traceless symmetric tensors is then given by

$$N_{\text{traceless}-\text{sym}}(\ell) = N_{\text{sym}}(\ell) - N_{\text{sym}}(\ell - 2) = \frac{1}{2}(\ell + 1)(\ell + 2) - \frac{1}{2}(\ell - 1)\ell$$

$$= \boxed{2\ell + 1 \ .}$$

(1.35)

For some purposes it will be useful to adopt a standard basis for the traceless symmetric tensors of rank ℓ. We now know that such a basis should have $2\ell+1$ elements. The standard spherical harmonic functions $Y_{\ell m}(\theta, \phi)$ are constructed by choosing a particular basis for the traceless symmetric tensors of rank ℓ, where each basis element is labeled by an integer m, with $-\ell \le m \le \ell$. We will denote these basis tensors by $C_{i_1...i_\ell}^{(\ell,m)}$, and the spherical harmonics will be written as

$$\boxed{Y_{\ell m}(\theta, \phi) = C_{i_1...i_\ell}^{(\ell,m)} \, \hat{\boldsymbol{n}}_{i_1} \ldots \hat{\boldsymbol{n}}_{i_\ell} \ .}$$

(1.36)

Later we will discuss the precise form of the basis tensors $C_{i_1...i_\ell}^{(\ell,m)}$.

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 (1.10) 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 do this is to build the traceless symmetric tensor from the vector $\hat{\boldsymbol{z}}$, the unit vector in the z direction. Note that the x-, y-, and z- components of $\hat{\boldsymbol{z}}$ are 0, 0, and 1, respectively, so $\hat{\boldsymbol{z}}_i = \delta_{i3}$. (Note that I have also used $\hat{\boldsymbol{e}}_3$ and $\hat{\boldsymbol{e}}_z$ to indicate the unit vector in the z direction. Since this vector will appear many times in the discussion of azimuthal symmetry, I am here using the simplest notation, $\hat{\boldsymbol{z}}$.)

A rank ℓ tensor can be constructed from $\hat{\boldsymbol{z}}$ simply by taking the product $\hat{\boldsymbol{z}}_{i_1}\hat{\boldsymbol{z}}_{i_2} \ldots \hat{\boldsymbol{z}}_{i_\ell}$. 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. More precisely, we can state a theorem that we will prove later.

Trace Decomposition Theorem: Any symmetric matrix $S_{i_1...i_\ell}$ can be uniquely written in the form

$$S_{i_1...i_\ell} = S_{i_1...i_\ell}^{(\text{TS})} + \underset{i_1...i_\ell}{\text{Sym}} \left[M_{i_1...i_{\ell-2}}\delta_{i_{\ell-1},i_\ell} \right] \ ,$$

(1.37)

where $S_{i_1\ldots i_\ell}^{(\mathrm{TS})}$ is a traceless symmetric tensor, $M_{i_1\ldots i_{\ell-2}}$ is a symmetric tensor, and

$$\underset{i_1\ldots i_\ell}{\mathrm{Sym}}[xxx] \tag{1.38}$$

means to symmetrize the expression xxx in the indices $i_1\ldots i_\ell$. That is, one sums over all $\ell!$ orderings of the indices $i_1\ldots i_\ell$, and then divides by $\ell!$:

$$\underset{i_1\ldots i_\ell}{\mathrm{Sym}}[M_{i_1\ldots i_\ell}] = \frac{1}{\ell!} \sum_{\substack{\text{permutations}\\ \{j_1\ldots j_\ell\}\ \text{of}\ \{i_1\ldots i_\ell\}}} M_{j_1\ldots j_\ell} \ . \tag{1.39}$$

I will use the symbol $\{\ \}_{\mathrm{TS}}$ to denote the traceless symmetric part of any symmetric tensor; i.e., in the language of Eq. (1.37),

$$\boxed{\ \{\, S_{i_1\ldots i_\ell} \,\}_{\mathrm{TS}} \equiv S_{i_1\ldots i_\ell}^{(\mathrm{TS})} = S_{i_1\ldots i_\ell} - \underset{i_1\ldots i_\ell}{\mathrm{Sym}}\left[M_{i_1\ldots i_{\ell-2}}\delta_{i_{\ell-1},i_\ell}\right].\ } \tag{1.40}$$

The general theorem sounds a bit abstract, but we can make it more concrete by applying it to some examples that are relevant to the present discussion. Tensors of rank 0 and 1 (i.e., scalars and vectors) are by definition traceless. Thus,

$$\boxed{\begin{aligned} \{\, 1\, \}_{\mathrm{TS}} &= 1\ , \\ \{\, \hat{\boldsymbol{z}}_i\, \}_{\mathrm{TS}} &= \hat{\boldsymbol{z}}_i\ . \end{aligned}} \tag{1.41}$$

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

$$\boxed{\ \{\, \hat{\boldsymbol{z}}_i\hat{\boldsymbol{z}}_j\, \}_{\mathrm{TS}} = \hat{\boldsymbol{z}}_i\hat{\boldsymbol{z}}_j - \tfrac{1}{3}\delta_{ij}\ .\ } \tag{1.42}$$

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

$$\delta_{ii} = 3\ . \tag{1.43}$$

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

$$\boxed{\ \{\, \hat{\boldsymbol{z}}_i\hat{\boldsymbol{z}}_j\hat{\boldsymbol{z}}_k\, \}_{\mathrm{TS}} = \hat{\boldsymbol{z}}_i\hat{\boldsymbol{z}}_j\hat{\boldsymbol{z}}_k - \tfrac{1}{5}\left(\hat{\boldsymbol{z}}_i\delta_{jk} + \hat{\boldsymbol{z}}_j\delta_{ik} + \hat{\boldsymbol{z}}_k\delta_{ij}\right)\ .\ } \tag{1.44}$$

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:

$$
\begin{aligned}
\delta_{ij}\{\,\hat{z}_i\hat{z}_j\hat{z}_k\,\}_{\mathrm{TS}} &= \hat{z}_i\hat{z}_i\hat{z}_k - \tfrac{1}{5}\big(\hat{z}_i\delta_{ik} + \hat{z}_i\delta_{ik} + \hat{z}_k\delta_{ii}\big) \\
&= \hat{z}_k - \tfrac{1}{5}(\hat{z}_k + \hat{z}_k + 3\hat{z}_k) = 0 \ .
\end{aligned}
\tag{1.45}
$$

For rank 4 there is the possibility of subtracting terms with either one or two Kronecker δ-functions, and both are needed to give a traceless result. (In the language of Eq. (1.40), the matrix M itself can include terms that are written in terms of Kronecker delta functions.) We can start with arbitrary coefficients, and see what they have to be:

$$
\begin{aligned}
\{\,\hat{z}_i\hat{z}_j\hat{z}_k\hat{z}_m\,\}_{\mathrm{TS}} &= \hat{z}_i\hat{z}_j\hat{z}_k\hat{z}_m + c_1\big(\hat{z}_i\hat{z}_j\delta_{km} + \hat{z}_i\hat{z}_k\delta_{mj} + \hat{z}_i\hat{z}_m\delta_{jk} + \hat{z}_j\hat{z}_k\delta_{im} \\
&\quad + \hat{z}_j\hat{z}_m\delta_{ik} + \hat{z}_k\hat{z}_m\delta_{ij}\big) + c_2\big(\delta_{ij}\delta_{km} + \delta_{ik}\delta_{jm} + \delta_{im}\delta_{jk}\big) \ .
\end{aligned}
\tag{1.46}
$$

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

$$
\begin{aligned}
\delta_{ij}\{\,\hat{z}_i\hat{z}_j\hat{z}_k\hat{z}_m\,\}_{\mathrm{TS}} &= \hat{z}_k\hat{z}_m + c_1\big(\delta_{km} + \hat{z}_k\hat{z}_m + \hat{z}_k\hat{z}_m + \hat{z}_k\hat{z}_m \\
&\quad + \hat{z}_k\hat{z}_m + 3\hat{z}_k\hat{z}_m\big) + c_2\big(3\delta_{km} + \delta_{km} + \delta_{km}\big) \\
&= \big(1 + 7c_1\big)\hat{z}_k\hat{z}_m + \big(c_1 + 5c_2\big)\delta_{km} \ .
\end{aligned}
\tag{1.47}
$$

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,

$$
\boxed{
\begin{aligned}
\{\,\hat{z}_i\hat{z}_j\hat{z}_k\hat{z}_m\,\}_{\mathrm{TS}} &= \hat{z}_i\hat{z}_j\hat{z}_k\hat{z}_m - \tfrac{1}{7}\big(\hat{z}_i\hat{z}_j\delta_{km} + \hat{z}_i\hat{z}_k\delta_{mj} + \hat{z}_i\hat{z}_m\delta_{jk} + \hat{z}_j\hat{z}_k\delta_{im} \\
&\quad + \hat{z}_j\hat{z}_m\delta_{ik} + \hat{z}_k\hat{z}_m\delta_{ij}\big) + \tfrac{1}{35}\big(\delta_{ij}\delta_{km} + \delta_{ik}\delta_{jm} + \delta_{im}\delta_{jk}\big) \ .
\end{aligned}
}
$$
$$
\tag{1.48}
$$

Any traceless symmetric tensor of rank ℓ that is invariant under rotations about the z-axis must be constructed from ingredients that are invariant under such rotations, which include $\hat{z}_i$, the Kronecker delta δ_{ij}, and the Levi-Civita tensor ϵ_{ijk}. But ϵ_{ijk} is antisymmetric and so cannot be used to build a symmetric tensor, and any term with a factor of δ_{ij} will become part of the second term of the trace decomposition of Eq. (1.37), and so will not be part of a traceless symmetric tensor. As a result, we expect that any traceless symmetric tensor of rank ℓ that is invariant under rotations about the z-axis must be proportional to $\{\,\hat{z}_{i_1}\ldots\hat{z}_{i_\ell}\,\}_{\mathrm{TS}}$. The argument here is not a real proof, but the conclusion is true. Since Eq. (1.10) 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 $\hat{\boldsymbol{n}}$) can be expanded as

$$
\boxed{
\begin{aligned}
F(\theta) &= c_0 + c_1 \{\, \hat{\boldsymbol{z}}_i \,\}_{\mathrm{TS}} \hat{\boldsymbol{n}}_i + c_2 \{\, \hat{\boldsymbol{z}}_i \hat{\boldsymbol{z}}_j \,\}_{\mathrm{TS}} \hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_j + \dots \\
&\quad + c_\ell \{\, \hat{\boldsymbol{z}}_{i_1} \dots \hat{\boldsymbol{z}}_{i_\ell} \,\}_{\mathrm{TS}} \hat{\boldsymbol{n}}_{i_1} \dots \hat{\boldsymbol{n}}_{i_\ell} + \dots \; ,
\end{aligned}
}
\tag{1.49}
$$

where the c_ℓ's are constants. This corresponds to what is standardly called an expansion in Legendre polynomials. In Lecture Notes 2 I will show exactly how to relate these expressions to the standard conventions for normalizing the Legendre polynomials, but we can see here exactly what these functions are, up to normalization conventions. Using $\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{n}} = \cos\theta$, we have

$$
\boxed{
\begin{aligned}
\{\, 1 \,\}_{\mathrm{TS}} &= 1 \\[4pt]
\{\, \hat{\boldsymbol{z}}_i \,\}_{\mathrm{TS}} \hat{\boldsymbol{n}}_i &= \cos\theta \\[4pt]
\{\, \hat{\boldsymbol{z}}_i \hat{\boldsymbol{z}}_j \,\}_{\mathrm{TS}} \hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_j &= \cos^2\theta - \frac{1}{3} \\[4pt]
\{\, \hat{\boldsymbol{z}}_i \hat{\boldsymbol{z}}_j \hat{\boldsymbol{z}}_k \,\}_{\mathrm{TS}} \hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_j \hat{\boldsymbol{n}}_k &= \cos^3\theta - \frac{3}{5}\cos\theta \\[4pt]
\{\, \hat{\boldsymbol{z}}_i \hat{\boldsymbol{z}}_j \hat{\boldsymbol{z}}_k \hat{\boldsymbol{z}}_m \,\}_{\mathrm{TS}} \hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_j \hat{\boldsymbol{n}}_k \hat{\boldsymbol{n}}_m &= \cos^4\theta - \frac{6}{7}\cos^2\theta + \frac{3}{35} \; .
\end{aligned}
}
\tag{1.50}
$$

The Legendre polynomials are normalized so that $P_\ell(\cos\theta = 1) = 1$, and we will see in the next lecture that this implies

$$
P_\ell(\cos\theta) = \frac{(2\ell)!}{2^\ell (\ell!)^2} \{\, \hat{\boldsymbol{z}}_{i_1} \dots \hat{\boldsymbol{z}}_{i_\ell} \,\}_{\mathrm{TS}} \hat{\boldsymbol{n}}_{i_1} \dots \hat{\boldsymbol{n}}_{i_\ell} \; .
\tag{1.51}
$$