Separation of Variables in Spherical Coordinates

13

The laws of physics are almost exclusively written in the form of differential equations (DEs). In (point) particle mechanics there is only one independent variable, leading to ordinary differential equations (ODEs). In other areas of physics in which extended objects such as fields are studied, variations with respect to position are also important. Partial derivatives with respect to coordinate variables show up in the differential equations, which are therefore called partial differential equations (PDEs). We list the most common PDEs of mathematical physics in the following.

13.1 PDEs of Mathematical Physics

In electrostatics, where time-independent scalar fields, such as potentials, and vector fields such as electrostatic fields, are studied, the law is described by **Poisson's equation**,

Poisson's equation

$$\nabla^2 \Phi(\mathbf{r}) = -4\pi\rho(\mathbf{r}). \tag{13.1}$$

In vacuum, where $\rho(\mathbf{r}) = 0$, Eq. (13.1) reduces to **Laplace's equation**,

Laplace's equation

$$\nabla^2 \Phi(\mathbf{r}) = 0. \tag{13.2}$$

Many electrostatic problems involve conductors held at constant potentials and situated in vacuum. In the space between such conducting surfaces, the electrostatic potential obeys Eq. (13.2).

The most simplified version of the **heat equation** is

heat equation

$$\frac{\partial T}{\partial t} = a^2 \nabla^2 T(\mathbf{r}), \tag{13.3}$$

where T is the temperature and a is a constant characterizing the medium in which heat flows.

One of the most frequently recurring PDEs encountered in mathematical physics is the **wave equation**,

wave equation

$$\nabla^2 \Psi - \frac{1}{c^2}\frac{\partial^2 \Psi}{\partial t^2} = 0. \tag{13.4}$$

S. Hassani, *Mathematical Physics*, DOI 10.1007/978-3-319-01195-0_13,

This equation (or its simplification to lower dimensions) is applied to the vibration of strings and drums; the propagation of sound in gases, solids, and liquids; the propagation of disturbances in plasmas; and the propagation of electromagnetic waves.

Schrödinger equation

The **Schrödinger equation**, describing nonrelativistic quantum phenomena, is

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V(\mathbf{r})\Psi = i\hbar\frac{\partial\Psi}{\partial t}, \tag{13.5}$$

where m is the mass of a subatomic particle, $\hbar$ is Planck's constant (divided by 2π), V is the potential energy of the particle, and $|\Psi(\mathbf{r}, t)|^2$ is the probability density of finding the particle at $\mathbf{r}$ at time t.

Klein-Gordon equation

A relativistic generalization of the Schrödinger equation for a free particle of mass m is the **Klein-Gordon equation**, which, in terms of the natural units ($\hbar = 1 = c$), reduces to

$$\nabla^2\phi - m^2\phi = \frac{\partial^2\phi}{\partial t^2}. \tag{13.6}$$

Equations (13.3)–(13.6) have partial derivatives with respect to time. As a first step toward solving these PDEs and as an introduction to similar techniques used in the solution of PDEs not involving time,[1] let us separate the time variable. We will denote the functions in all four equations by the generic symbol $\Psi(\mathbf{r}, t)$. The basic idea is to separate the $\mathbf{r}$ and t dependence into factors: $\Psi(\mathbf{r}, t) \equiv R(\mathbf{r})T(t)$. This factorization permits us to separate the two operations of space differentiation and time differentiation. Let $\mathbf{S}$ stand for all spatial derivative operators and write all the relevant equations either as $\mathbf{S}\Psi = \partial\Psi/\partial t$ or as $\mathbf{S}\Psi = \partial^2\Psi/\partial t^2$. With this notation and the above separation, we have

time is separated from space

$$\mathbf{S}(RT) = T(\mathbf{S}R) = \begin{cases} R\,dT/dt, \\ R\,d^2T/dt^2. \end{cases}$$

Dividing both sides by RT, we obtain

$$\frac{1}{R}\mathbf{S}R = \begin{cases} \frac{1}{T}\frac{dT}{dt}, \\ \frac{1}{T}\frac{d^2T}{dt^2}. \end{cases} \tag{13.7}$$

Now comes the crucial step in the process of separation of variables. The LHS of Eq. (13.7) is a function of *position alone*, and the RHS is a function of *time alone*. Since $\mathbf{r}$ and t are independent variables, the only way that (13.7) can hold is for *both sides to be constant*, say α:

$$\frac{1}{R}\mathbf{S}R = \alpha \quad \Rightarrow \quad \mathbf{S}R = \alpha R$$

[1] See [Hass 08] for a thorough discussion of separation in Cartesian and cylindrical coordinates. Chapter 19 of this book also contains examples of solutions to some second-order linear DEs resulting from such separation.

and[2]

$$\frac{1}{T}\frac{dT}{dt}=\alpha \quad\Rightarrow\quad \frac{dT}{dt}=\alpha T \quad\text{or}\quad \frac{1}{T}\frac{d^2T}{dt^2}=\alpha \quad\Rightarrow\quad \frac{d^2T}{dt^2}=\alpha T.$$

We have reduced the original time-dependent PDE to an ODE,

$$\frac{dT}{dt}=\alpha T \quad\text{or}\quad \frac{d^2T}{dt^2}=\alpha T, \tag{13.8}$$

and a PDE involving only the position variables, $(\mathbf{S}-\alpha)R=0$. The most general form of $\mathbf{S}-\alpha$ arising from Eqs. (13.3) through (13.6) is $\mathbf{S}-\alpha\equiv\nabla^2+f(\mathbf{r})$. Therefore, Eqs. (13.3)–(13.6) are equivalent to (13.8), and $\nabla^2R+f(\mathbf{r})R=0$, which we rewrite as

$$\nabla^2\Psi(\mathbf{r})+f(\mathbf{r})\Psi(\mathbf{r})=0. \tag{13.9}$$

This is called a *homogeneous* PDE because of the zero on the right-hand side. Of all the PDEs outlined above, Poisson's equation is the only inhomogeneous equation. We will restrict ourselves to the homogeneous case in this chapter.

Depending on the geometry of the problem, Eq. (13.9) is further separated into ODEs each involving a single coordinate of a suitable coordinate system. We shall see examples of all major coordinate systems (Cartesian, cylindrical, and spherical) in Chap. 19. For the rest of this chapter, we shall concentrate on some general aspects of the spherical coordinates.

Historical Notes

Jean Le Rond d'Alembert (1717–1783) was the illegitimate son of a famous salon hostess of eighteenth-century Paris and a cavalry officer. Abandoned by his mother, d'Alembert was raised by a foster family and later educated by the arrangement of his father at a nearby church-sponsored school, in which he received instruction in the classics and above-average instruction in mathematics. After studying law and medicine, he finally chose to pursue a career in mathematics. In the 1740s he joined the ranks of the *philosophes*, a growing group of deistic and materialistic thinkers and writers who actively questioned the social and intellectual standards of the day. He traveled little (he left France only once, to visit the court of Frederick the Great), preferring instead the company of his friends in the salons, among whom he was well known for his wit and laughter.

Jean Le Rond d'Alembert
1717–1783

D'Alembert turned his mathematical and philosophical talents to many of the outstanding scientific problems of the day, with mixed success. Perhaps his most famous scientific work, entitled *Traité de dynamique*, shows his appreciation that a revolution was taking place in the science of mechanics—the formalization of the principles stated by Newton into a rigorous mathematical framework. The philosophy to which d'Alembert subscribed, however, refused to acknowledge the primacy of a concept as unclear and arbitrary as "force," introducing a certain awkwardness to his treatment and perhaps causing him to overlook the important principle of conservation of energy. Later, d'Alembert produced a treatise on fluid mechanics (the priority of which is still debated by historians), a paper dealing with vibrating strings (in which the wave equation makes its first appearance in physics), and a skillful treatment of celestial mechanics. D'Alembert is also credited with use of the first partial differential equation as well as the first solution

[2]In most cases, α is chosen to be real. In the case of the Schrödinger equation, it is more convenient to choose α to be purely imaginary so that the i in the definition of $\mathbf{S}$ can be compensated. In all cases, the precise nature of α is determined by boundary conditions.

to such an equation using **separation of variables**. (One should be careful interpreting "first": many of d'Alembert's predecessors and contemporaries gave similar, though less satisfactory, treatments of these milestones.) Perhaps his most well-known contribution to mathematics (at least among students) is the ratio test for the convergence of infinite series.

Much of the work for which d'Alembert is remembered occurred outside mathematical physics. He was chosen as the science editor of the *Encyclopédie*, and his lengthy *Discours Préliminaire* in that volume is considered one of the defining documents of the Enlightenment. Other works included writings on law, religion, and music.

Since d'Alembert's final years were not especially happy ones, perhaps this account of his life should end with a glimpse at the humanity his philosophy often gave his work. Like many of his contemporaries, he considered the problem of calculating the relative risk associated with the new practice of smallpox inoculation, which in rare cases caused the disease it was designed to prevent. Although not very successful in the mathematical sense, he was careful to point out that the probability of accidental infection, however slight or elegantly derived, would be small consolation to a father whose child died from the inoculation. It is greatly to his credit that d'Alembert did not believe such considerations irrelevant to the problem.

13.2 Separation of the Angular Part

With Cartesian and cylindrical variables, the boundary conditions are important in determining the nature of the solutions of the ODE obtained from the PDE. In almost all applications, however, the angular part of the spherical variables can be separated and studied very generally. This is because the angular part of the Laplacian in the spherical coordinate system is closely related to the operation of rotation and the angular momentum, which are independent of any particular situation.

The separation of the angular part in spherical coordinates can be done in a fashion exactly analogous to the separation of time by writing Ψ as a product of three functions, each depending on only one of the variables. However, we will follow an approach that is used in quantum mechanical treatments of angular momentum. This approach, which is based on the operator algebra of Chap. 4 and is extremely powerful and elegant, gives solutions for the angular part in closed form.

Define the vector operator $\vec{\mathbf{p}}$ as $\vec{\mathbf{p}} = -i\nabla$ so that its jth Cartesian component is $\mathbf{p}_j = -i\partial/\partial x_j$, for $j = 1, 2, 3$. In quantum mechanics $\vec{\mathbf{p}}$ (multiplied by $\hbar$) is the momentum operator. It is easy to verify that[3]

$$[x_j, \mathbf{p}_k] = i\delta_{jk} \quad \text{and} \quad [x_j, x_k] = 0 = [\mathbf{p}_j, \mathbf{p}_k]. \tag{13.10}$$

angular momentum operator

We can also define the **angular momentum operator** as $\vec{\mathbf{L}} = \vec{\mathbf{r}} \times \vec{\mathbf{p}}$. This is expressed in components as $\mathbf{L}_i = (\vec{\mathbf{r}} \times \vec{\mathbf{p}})_i = \epsilon_{ijk} x_j \mathbf{p}_k$ for $i = 1, 2, 3$, where Einstein's summation convention (summing over repeated indices) is

[3]These operators act on the space of functions possessing enough "nice" properties as to render the space suitable. The operator x_j simply multiplies functions, while $\mathbf{p}_j$ differentiates them.

utilized.[4] Using the commutation relations above, we obtain

commutation relations between components of angular momentum operator

$$[\mathbf{L}_j, \mathbf{L}_k] = i\epsilon_{jkl}\mathbf{L}_l.$$

We will see shortly that $\vec{\mathbf{L}}$ can be written solely in terms of the angles θ and φ. Moreover, there is one factor of $\vec{\mathbf{p}}$ in the definition of $\vec{\mathbf{L}}$, so if we square $\vec{\mathbf{L}}$, we will get two factors of $\vec{\mathbf{p}}$, and a Laplacian may emerge in the expression for $\vec{\mathbf{L}} \cdot \vec{\mathbf{L}}$. In this manner, we may be able to write ∇^2 in terms of $\mathbf{L}^2$, which depends only on angles. Let us try this:

$$\begin{aligned}\mathbf{L}^2 = \vec{\mathbf{L}} \cdot \vec{\mathbf{L}} = \sum_{i=1}^{3} \mathbf{L}_i \mathbf{L}_i &= \epsilon_{ijk} x_j \mathbf{p}_k \epsilon_{imn} x_m \mathbf{p}_n = \epsilon_{ijk}\epsilon_{imn} x_j \mathbf{p}_k x_m \mathbf{p}_n \\ &= (\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}) x_j \mathbf{p}_k x_m \mathbf{p}_n = x_j \mathbf{p}_k x_j \mathbf{p}_k - x_j \mathbf{p}_k x_k \mathbf{p}_j.\end{aligned}$$

We need to write this expression in such a way that factors with the same index are next to each other, to give a dot product. We must also try, when possible, to keep the $\vec{\mathbf{p}}$ factors to the right so that they can operate on functions without intervention from the x factors. We do this by using Eq. (13.10):

$$\begin{aligned}\mathbf{L}^2 &= x_j(x_j \mathbf{p}_k - i\delta_{kj})\mathbf{p}_k - (\mathbf{p}_k x_j + i\delta_{kj}) x_k \mathbf{p}_j \\ &= x_j x_j \mathbf{p}_k \mathbf{p}_k - i x_j \mathbf{p}_j - \mathbf{p}_k x_k x_j \mathbf{p}_j - i x_j \mathbf{p}_j \\ &= x_j x_j \mathbf{p}_k \mathbf{p}_k - 2i x_j \mathbf{p}_j - (x_k \mathbf{p}_k - i\delta_{kk}) x_j \mathbf{p}_j.\end{aligned}$$

Recalling that $\delta_{kk} = \sum_{k=1}^{3} \delta_{kk} = 3$ and $x_j x_j = \sum_{j=1}^{3} x_j x_j = \vec{\mathbf{r}} \cdot \vec{\mathbf{r}} = r^2$ etc., we can write $\mathbf{L}^2 = r^2 \vec{\mathbf{p}} \cdot \vec{\mathbf{p}} + i\vec{\mathbf{r}} \cdot \vec{\mathbf{p}} - (\vec{\mathbf{r}} \cdot \vec{\mathbf{p}})(\vec{\mathbf{r}} \cdot \vec{\mathbf{p}})$, which, if we make the substitution $\vec{\mathbf{p}} = -i\boldsymbol{\nabla}$, yields

$$\nabla^2 = -r^{-2}\mathbf{L}^2 + r^{-2}(\mathbf{r} \cdot \boldsymbol{\nabla})(\mathbf{r} \cdot \boldsymbol{\nabla}) + r^{-2}\mathbf{r} \cdot \boldsymbol{\nabla}.$$

Letting both sides act on the function $\Psi(r, \theta, \varphi)$, we get

$$\nabla^2 \Psi = -\frac{1}{r^2}\mathbf{L}^2 \Psi + \frac{1}{r^2}(\mathbf{r} \cdot \boldsymbol{\nabla})(\mathbf{r} \cdot \boldsymbol{\nabla})\Psi + \frac{1}{r^2}\mathbf{r} \cdot \boldsymbol{\nabla}\Psi. \tag{13.11}$$

But we note that $\mathbf{r} \cdot \boldsymbol{\nabla} = r\hat{\mathbf{e}}_r \cdot \boldsymbol{\nabla} = r\partial/\partial r$. We thus get the final form of $\nabla^2 \Psi$ in spherical coordinates:

Laplacian separated into angular and radial parts

$$\nabla^2 \Psi = -\frac{1}{r^2}\mathbf{L}^2 \Psi + \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \Psi}{\partial r}\right) + \frac{1}{r}\frac{\partial \Psi}{\partial r}. \tag{13.12}$$

It is important to note that Eq. (13.11) is a general relation that holds in all coordinate systems. Although all the manipulations leading to it were done in Cartesian coordinates, since it is written in vector notation, there is no indication in the final form that it was derived using specific coordinates.

[4] It is assumed that the reader is familiar with vector algebra using indices and such objects as δ_{ij} and ϵ_{ijk}. For an introductory treatment, sufficient for our present discussion, see [Hass 08]. A more advanced treatment of these objects (tensors) can be found in Part VIII of the present book.

Equation (13.12) is the spherical version of (13.11) and is the version we shall use. We will first make the simplifying assumption that in Eq. (13.9), the master equation, $f(\mathbf{r})$ is a function of r only. Equation (13.9) then becomes

$$-\frac{1}{r^2}\mathbf{L}^2\Psi+\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\Psi}{\partial r}\right)+\frac{1}{r}\frac{\partial\Psi}{\partial r}+f(r)\Psi=0.$$

Assuming, for the time being, that $\mathbf{L}^2$ depends only on θ and φ, and separating Ψ into a product of two functions, $\Psi(r,\theta,\varphi)=R(r)Y(\theta,\varphi)$, we can rewrite this equation as

$$-\frac{1}{r^2}\mathbf{L}^2(RY)+\frac{1}{r}\frac{\partial}{\partial r}\left[r\frac{\partial}{\partial r}(RY)\right]+\frac{1}{r}\frac{\partial}{\partial r}(RY)+f(r)RY=0.$$

Dividing by RY and multiplying by r^2 yields

$$\underbrace{-\frac{1}{Y}\mathbf{L}^2(Y)}_{-\alpha}+\underbrace{\frac{r}{R}\frac{d}{dr}\left(r\frac{dR}{dr}\right)+\frac{r}{R}\frac{dR}{dr}+r^2f(r)}_{+\alpha}=0,$$

or

$$\mathbf{L}^2Y(\theta,\varphi)=\alpha Y(\theta,\varphi) \tag{13.13}$$

and

$$\frac{d^2R}{dr^2}+\frac{2}{r}\frac{dR}{dr}+\left[f(r)-\frac{\alpha}{r^2}\right]R=0. \tag{13.14}$$

We will concentrate on the angular part, Eq. (13.13), leaving the radial part to the general discussion of ODEs. The rest of this section will focus on showing that $\mathbf{L}_1\equiv\mathbf{L}_x$, $\mathbf{L}_2\equiv\mathbf{L}_y$, and $\mathbf{L}_3\equiv\mathbf{L}_z$ are independent of r.

Since $\mathbf{L}_i$ is an operator, we can study its action on an arbitrary function f. Thus,

$$\mathbf{L}_i f=-i\epsilon_{ijk}x_j\nabla_k f\equiv-i\epsilon_{ijk}x_j\partial f/\partial x_k.$$

We can express the Cartesian x_j in terms of r, θ, and φ, and use the chain rule to express $\partial f/\partial x_k$ in terms of spherical coordinates. This will give us $\mathbf{L}_i f$ expressed in terms of r, θ, and φ. It will then emerge that r is absent in the final expression.

Let us start with

$$x=r\sin\theta\cos\varphi,\qquad y=r\sin\theta\sin\varphi,\qquad z=r\cos\theta,$$

and their inverse,

$$r=\left(x^2+y^2+z^2\right)^{1/2},\qquad \cos\theta=\frac{z}{r},\qquad \tan\varphi=\frac{y}{x},$$

and express the Cartesian derivatives in terms of spherical coordinates using the chain rule. The first such derivative is

$$\frac{\partial f}{\partial x}=\frac{\partial f}{\partial r}\frac{\partial r}{\partial x}+\frac{\partial f}{\partial\theta}\frac{\partial\theta}{\partial x}+\frac{\partial f}{\partial\varphi}\frac{\partial\varphi}{\partial x}. \tag{13.15}$$

The derivative of one coordinate system with respect to the other can be easily calculated. For example, $\partial r/\partial x = x/r = \sin\theta\cos\varphi$, and differentiating both sides of the equation $\cos\theta = z/r$, we obtain

$$-\sin\theta\frac{\partial\theta}{\partial x} = -\frac{z\partial r/\partial x}{r^2} = -\frac{zx}{r^3} = -\frac{\cos\theta\sin\theta\cos\varphi}{r}$$
$$\Rightarrow \quad \frac{\partial\theta}{\partial x} = \frac{\cos\theta\cos\varphi}{r}.$$

Finally, differentiating both sides of $\tan\varphi = y/x$ with respect to x yields $\partial\varphi/\partial x = -\sin\varphi/(r\sin\theta)$. Using these expressions in Eq. (13.15), we get

$$\frac{\partial f}{\partial x} = \sin\theta\cos\varphi\frac{\partial f}{\partial r} + \frac{\cos\theta\cos\varphi}{r}\frac{\partial f}{\partial\theta} - \frac{\sin\varphi}{r\sin\theta}\frac{\partial f}{\partial\varphi}.$$

In exactly the same way, we obtain

$$\frac{\partial f}{\partial y} = \sin\theta\sin\varphi\frac{\partial f}{\partial r} + \frac{\cos\theta\sin\varphi}{r}\frac{\partial f}{\partial\theta} + \frac{\cos\varphi}{r\sin\theta}\frac{\partial f}{\partial\varphi},$$
$$\frac{\partial f}{\partial z} = \cos\theta\frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial\theta}.$$

We can now calculate $\mathbf{L}_x$ by letting it act on an arbitrary function and expressing all Cartesian coordinates and derivatives in terms of spherical coordinates. The result is

Cartesian components of angular momentum operator expressed in spherical coordinates

$$\mathbf{L}_x f = -iy\frac{\partial f}{\partial z} + iz\frac{\partial f}{\partial y} = i\left(\sin\varphi\frac{\partial}{\partial\theta} + \cot\theta\cos\varphi\frac{\partial}{\partial\varphi}\right)f,$$

or

$$\mathbf{L}_x = i\left(\sin\varphi\frac{\partial}{\partial\theta} + \cot\theta\cos\varphi\frac{\partial}{\partial\varphi}\right). \tag{13.16}$$

Analogous arguments yield

$$\mathbf{L}_y = i\left(-\cos\varphi\frac{\partial}{\partial\theta} + \cot\theta\sin\varphi\frac{\partial}{\partial\varphi}\right), \qquad \mathbf{L}_z = -i\frac{\partial}{\partial\varphi}. \tag{13.17}$$

It is left as a problem for the reader to show that by adding the squares of the components of the angular momentum operator, one obtains

angular momentum squared as differential operator in θ and φ

$$\mathbf{L}^2 = -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) - \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}, \tag{13.18}$$

which is independent of r as promised. Substitution in Eq. (13.12) yields the familiar expression for the Laplacian in spherical coordinates.

13.3 Construction of Eigenvalues of $\mathbf{L}^2$

Now that we have $\mathbf{L}^2$ in terms of θ and φ, we could substitute in Eq. (13.13), separate the θ and φ dependence, and solve the corresponding ODEs. However, there is a much more elegant way of solving this problem algebraically, because Eq. (13.13) is simply an eigenvalue equation for $\mathbf{L}^2$. In this section, we will find the eigenvalues of $\mathbf{L}^2$. The next section will evaluate the eigenvectors of $\mathbf{L}^2$.

Let us consider $\mathbf{L}^2$ as an abstract operator and write (13.13) as

$$\mathbf{L}^2|Y\rangle = \alpha|Y\rangle,$$

where $|Y\rangle$ is an abstract vector whose (θ, φ)th component can be calculated later. Since $\mathbf{L}^2$ is a differential operator, it does not have a (finite-dimensional) matrix representation. Thus, the determinantal procedure for calculating eigenvalues and eigenfunctions will not work here, and we have to find another way.

The equation above specifies an eigenvalue, α, and an eigenvector, $|Y\rangle$. There may be more than one $|Y\rangle$ corresponding to the same α. To distinguish among these so-called *degenerate eigenvectors*, we choose a second operator, say $\mathbf{L}_3 \in \{\mathbf{L}_i\}$ that commutes with $\mathbf{L}^2$. This allows us to select a basis in which both $\mathbf{L}^2$ and $\mathbf{L}_3$ are diagonal, or, equivalently, a basis whose vectors are simultaneous eigenvectors of both $\mathbf{L}^2$ and $\mathbf{L}_3$. This is possible by Theorem 6.4.18 and the fact that both $\mathbf{L}^2$ and $\mathbf{L}_3$ are hermitian operators in the space of square-integrable functions. (The proof is left as a problem.) In general, we would want to continue adding operators until we obtained a maximum set of commuting operators which could label the eigenvectors. In this case, $\mathbf{L}^2$ and $\mathbf{L}_3$ exhaust the set.[5] Using the more common subscripts x, y, and z instead of 1, 2, 3 and attaching labels to the eigenvectors, we have

$$\mathbf{L}^2|Y_{\alpha,\beta}\rangle = \alpha|Y_{\alpha,\beta}\rangle, \qquad \mathbf{L}_z|Y_{\alpha,\beta}\rangle = \beta|Y_{\alpha,\beta}\rangle. \tag{13.19}$$

The hermiticity of $\mathbf{L}^2$ and $\mathbf{L}_z$ implies the reality of α and β. Next we need to determine the possible values for α and β.

Define two new operators $\mathbf{L}_+ \equiv \mathbf{L}_x + i\mathbf{L}_y$ and $\mathbf{L}_- \equiv \mathbf{L}_x - i\mathbf{L}_y$. It is then easily verified that

$$\big[\mathbf{L}^2, \mathbf{L}_\pm\big] = 0, \qquad [\mathbf{L}_z, \mathbf{L}_\pm] = \pm\mathbf{L}_\pm, \qquad [\mathbf{L}_+, \mathbf{L}_-] = 2\mathbf{L}_z. \tag{13.20}$$

The first equation implies that $\mathbf{L}_\pm$ are invariant operators when acting in the subspace corresponding to the eigenvalue α; that is, $\mathbf{L}_\pm|Y_{\alpha,\beta}\rangle$ are eigenvectors of $\mathbf{L}^2$ with the same eigenvalue α:

$$\mathbf{L}^2\big(\mathbf{L}_\pm|Y_{\alpha,\beta}\rangle\big) = \mathbf{L}_\pm\big(\mathbf{L}^2|Y_{\alpha,\beta}\rangle\big) = \mathbf{L}_\pm\big(\alpha|Y_{\alpha,\beta}\rangle\big) = \alpha\mathbf{L}_\pm|Y_{\alpha,\beta}\rangle.$$

The second equation in (13.20) yields

$$\begin{aligned}\mathbf{L}_z\big(\mathbf{L}_+|Y_{\alpha,\beta}\rangle\big) &= (\mathbf{L}_z\mathbf{L}_+)|Y_{\alpha,\beta}\rangle = (\mathbf{L}_+\mathbf{L}_z + \mathbf{L}_+)|Y_{\alpha,\beta}\rangle \\ &= \mathbf{L}_+\mathbf{L}_z|Y_{\alpha,\beta}\rangle + \mathbf{L}_+|Y_{\alpha,\beta}\rangle = \beta\mathbf{L}_+|Y_{\alpha,\beta}\rangle + \mathbf{L}_+|Y_{\alpha,\beta}\rangle \\ &= (\beta+1)\mathbf{L}_+|Y_{\alpha,\beta}\rangle.\end{aligned}$$

[5] We could just as well have chosen $\mathbf{L}^2$ and any other component as our maximal set. However, $\mathbf{L}^2$ and $\mathbf{L}_3$ is the universally accepted choice.

This indicates that $\mathbf{L}_+|Y_{\alpha,\beta}\rangle$ has one more unit of the $\mathbf{L}_z$ eigenvalue than $|Y_{\alpha,\beta}\rangle$ does. In other words, $\mathbf{L}_+$ *raises* the eigenvalue of $\mathbf{L}_z$ by one unit. That is why $\mathbf{L}_+$ is called a **raising operator**. Similarly, $\mathbf{L}_-$ is called a **lowering operator** because $\mathbf{L}_z(\mathbf{L}_-|Y_{\alpha,\beta}\rangle) = (\beta - 1)\mathbf{L}_-|Y_{\alpha,\beta}\rangle$.

angular momentum raising and lowering operators

We can summarize the above discussion as

$$\mathbf{L}_\pm|Y_{\alpha,\beta}\rangle = C_\pm|Y_{\alpha,\beta\pm 1}\rangle,$$

where $C_\pm$ are constants to be determined by a suitable normalization.

There are restrictions on (and relations between) α and β. First note that as $\mathbf{L}^2$ is a sum of squares of hermitian operators, it must be a positive operator; that is, $\langle a|\mathbf{L}^2|a\rangle \geq 0$ for all $|a\rangle$. In particular,

$$0 \leq \langle Y_{\alpha,\beta}|\mathbf{L}^2|Y_{\alpha,\beta}\rangle = \alpha\langle Y_{\alpha,\beta}|Y_{\alpha,\beta}\rangle = \alpha\|Y_{\alpha,\beta}\|^2.$$

Therefore, $\alpha \geq 0$. Next, one can readily show that

$$\mathbf{L}^2 = \mathbf{L}_+\mathbf{L}_- + \mathbf{L}_z^2 - \mathbf{L}_z = \mathbf{L}_-\mathbf{L}_+ + \mathbf{L}_z^2 + \mathbf{L}_z. \tag{13.21}$$

Sandwiching both sides of the first equality between $|Y_{\alpha,\beta}\rangle$ and $\langle Y_{\alpha,\beta}|$ yields

$$\langle Y_{\alpha,\beta}|\mathbf{L}^2|Y_{\alpha,\beta}\rangle = \langle Y_{\alpha,\beta}|\mathbf{L}_+\mathbf{L}_-|Y_{\alpha,\beta}\rangle + \langle Y_{\alpha,\beta}|\mathbf{L}_z^2|Y_{\alpha,\beta}\rangle - \langle Y_{\alpha,\beta}|\mathbf{L}_z|Y_{\alpha,\beta}\rangle,$$

with an analogous expression involving $\mathbf{L}_-\mathbf{L}_+$. Using the fact that $\mathbf{L}_+ = (\mathbf{L}_-)^\dagger$, we get

$$\begin{aligned}\alpha\|Y_{\alpha,\beta}\|^2 &= \langle Y_{\alpha,\beta}|\mathbf{L}_+\mathbf{L}_-|Y_{\alpha,\beta}\rangle + \beta^2\|Y_{\alpha,\beta}\|^2 - \beta\|Y_{\alpha,\beta}\|^2 \\ &= \langle Y_{\alpha,\beta}|\mathbf{L}_-\mathbf{L}_+|Y_{\alpha,\beta}\rangle + \beta^2\|Y_{\alpha,\beta}\|^2 + \beta\|Y_{\alpha,\beta}\|^2 \\ &= \big\|\mathbf{L}_\mp|Y_{\alpha,\beta}\rangle\big\|^2 + \beta^2\|Y_{\alpha,\beta}\|^2 \mp \beta\|Y_{\alpha,\beta}\|^2.\end{aligned} \tag{13.22}$$

Because of the positivity of norms, this yields $\alpha \geq \beta^2 - \beta$ and $\alpha \geq \beta^2 + \beta$. Adding these two inequalities gives $2\alpha \geq 2\beta^2 \Rightarrow -\sqrt{\alpha} \leq \beta \leq \sqrt{\alpha}$. It follows that the values of β are bounded. That is, there exist a maximum β, denoted by β_+, and a minimum β, denoted by β_-, beyond which there are no more values of β. This can happen only if

$$\mathbf{L}_+|Y_{\alpha,\beta_+}\rangle = 0, \qquad \mathbf{L}_-|Y_{\alpha,\beta_-}\rangle = 0,$$

because if $\mathbf{L}_\pm|Y_{\alpha,\beta_\pm}\rangle$ are not zero, then they must have values of β corresponding to $\beta_\pm \pm 1$, which are not allowed.

Using β_+ for β in Eq. (13.22) yields

$$\big(\alpha - \beta_+^2 - \beta_+\big)\|Y_{\alpha,\beta_+}\|^2 = 0.$$

By definition $|Y_{\alpha,\beta_+}\rangle \neq 0$ (otherwise $\beta_+ - 1$ would be the maximum). Thus, we obtain $\alpha = \beta_+^2 + \beta_+$. An analogous procedure using β_- for β yields $\alpha = \beta_-^2 - \beta_-$. We solve these two equations for β_+ and β_-:

$$\beta_+ = \frac{1}{2}(-1 \pm \sqrt{1+4\alpha}), \qquad \beta_- = \frac{1}{2}(1 \pm \sqrt{1+4\alpha}).$$

Since $\beta_+ \geq \beta_-$ and $\sqrt{1+4\alpha} \geq 1$, we must choose

$$\beta_+ = \frac{1}{2}(-1 + \sqrt{1+4\alpha}) = -\beta_-.$$

Starting with $|Y_{\alpha,\beta_+}\rangle$, we can apply $\mathbf{L}_-$ to it repeatedly. In each step we decrease the value of β by one unit. There must be a limit to the number of vectors obtained in this way, because β has a minimum. Therefore, there must exist a nonnegative integer k such that

$$(\mathbf{L}_-)^{k+1}|Y_{\alpha,\beta_+}\rangle = \mathbf{L}_-\big(\mathbf{L}_-^k|Y_{\alpha,\beta_+}\rangle\big) = 0.$$

Thus, $\mathbf{L}_-^k|Y_{\alpha,\beta_+}\rangle$ must be proportional to $|Y_{\alpha,\beta_-}\rangle$. In particular, since $\mathbf{L}_-^k|Y_{\alpha,\beta_+}\rangle$ has a β value equal to $\beta_+ - k$, we have $\beta_- = \beta_+ - k$. Now, using $\beta_- = -\beta_+$ (derived above) yields the important result

$$\beta_+ = \frac{k}{2} \equiv j \quad \text{for } k \in \mathbb{N},$$

or $\alpha = j(j+1)$, since $\alpha = \beta_+^2 + \beta_+$. This result is important enough to be stated as a theorem.

eigenvalues of $\mathbf{L}^2$ and $\mathbf{L}_z$ given

Theorem 13.3.1 *The eigenvectors of* $\mathbf{L}^2$, *denoted by* $|Y_{jm}\rangle$, *satisfy the eigenvalue relations*

$$\mathbf{L}^2|Y_{jm}\rangle = j(j+1)|Y_{jm}\rangle, \qquad \mathbf{L}_z|Y_{jm}\rangle = m|Y_{jm}\rangle,$$

where j *is a positive integer or half-integer, and* m *can take a value in the set* $\{-j, -j+1, \dots, j-1, j\}$ *of* $2j+1$ *numbers.*

Let us briefly consider the normalization of the eigenvectors. We already know that the $|Y_{jm}\rangle$, being eigenvectors of the hermitian operators $\mathbf{L}^2$ and $\mathbf{L}_z$, are orthogonal. We also demand that they be of unit norm; that is,

$$\langle Y_{jm}|Y_{j'm'}\rangle = \delta_{jj'}\delta_{mm'}. \tag{13.23}$$

This will determine the constants $C_\pm$, introduced earlier. Let us consider C_+ first, which is defined by $\mathbf{L}_+|Y_{jm}\rangle = C_+|Y_{j,m+1}\rangle$. The hermitian conjugate of this equation is $\langle Y_{jm}|\mathbf{L}_- = C_+^*\langle Y_{j,m+1}|$. We contract these two equations to get

$$\langle Y_{jm}|\mathbf{L}_-\mathbf{L}_+|Y_{jm}\rangle = |C_+|^2\langle Y_{j,m+1}|Y_{j,m+1}\rangle.$$

Then we use the second relation in Eq. (13.21), Theorem 13.3.1, and (13.23) to obtain

$$j(j+1) - m(m+1) = |C_+|^2 \quad \Rightarrow \quad |C_+| = \sqrt{j(j+1) - m(m+1)}.$$

Adopting the convention that the argument (phase) of the complex number C_+ is zero (and therefore that C_+ is real), we get

$$C_+ = \sqrt{j(j+1) - m(m+1)}$$

Similarly, $C_- = \sqrt{j(j+1) - m(m-1)}$.

Box 13.3.2 *The raising and lowering operators act on* $|Y_{jm}\rangle$ *as follows*:

$$\begin{aligned} \mathbf{L}_+|Y_{jm}\rangle &= \sqrt{j(j+1) - m(m+1)}|Y_{j,m+1}\rangle, \\ \mathbf{L}_-|Y_{jm}\rangle &= \sqrt{j(j+1) - m(m-1)}|Y_{j,m-1}\rangle. \end{aligned} \tag{13.24}$$

Example 13.3.3 Assume that $j = l$, a positive integer. Let us find an expression for $|Y_{lm}\rangle$ by repeatedly applying $\mathbf{L}_-$ to $|Y_{ll}\rangle$. The action for $\mathbf{L}_-$ is completely described by Eq. (13.24). For the first power of $\mathbf{L}_-$, we obtain

$$\mathbf{L}_-|Y_{ll}\rangle = \sqrt{l(l+1) - l(l-1)}|Y_{l,l-1}\rangle = \sqrt{2l}|Y_{l,l-1}\rangle.$$

We apply $\mathbf{L}_-$ once more:

$$\begin{aligned} (\mathbf{L}_-)^2|Y_{ll}\rangle &= \sqrt{2l}\mathbf{L}_-|Y_{l,l-1}\rangle = \sqrt{2l}\sqrt{l(l+1) - (l-1)(l-2)}|Y_{l,l-2}\rangle \\ &= \sqrt{2l}\sqrt{2(2l-1)}|Y_{l,l-2}\rangle = \sqrt{2(2l)(2l-1)}|Y_{l,l-2}\rangle. \end{aligned}$$

Applying $\mathbf{L}_-$ a third time yields

$$\begin{aligned} (\mathbf{L}_-)^3|Y_{ll}\rangle &= \sqrt{2(2l)(2l-1)}\mathbf{L}_-|Y_{l,l-2}\rangle = \sqrt{2(2l)(2l-1)}\sqrt{6(l-1)}|Y_{l,l-3}\rangle \\ &= \sqrt{3!(2l)(2l-1)(2l-2)}|Y_{l,l-3}\rangle. \end{aligned}$$

The pattern suggests the following formula for a general power k:

$$\mathbf{L}_-^k|Y_{ll}\rangle = \sqrt{k!(2l)(2l-1)\cdots(2l-k+1)}|Y_{l,l-k}\rangle,$$

or $\mathbf{L}_-^k|Y_{ll}\rangle = \sqrt{k!(2l)!/(2l-k)!}|Y_{l,l-k}\rangle$. If we set $l - k = m$ and solve for $|Y_{l,m}\rangle$, we get

$$|Y_{l,m}\rangle = \sqrt{\frac{(l+m)!}{(l-m)!(2l)!}}\mathbf{L}_-^{l-m}|Y_{ll}\rangle.$$

The discussion in this section is the standard treatment of angular momentum in quantum mechanics. In the context of quantum mechanics, Theorem 13.3.1 states the far-reaching physical result that particles can have integer or half-integer spin. Such a conclusion is tied to the rotation group in three dimensions, which, in turn, is an example of a Lie group, or a continuous group of transformations. We shall come back to a study of groups later. It is worth noting that it was the study of differential equations that led the Norwegian mathematician Sophus Lie to the investigation of their symmetries and the development of the beautiful branch of mathematics and theoretical physics that bears his name. Thus, the existence of a connection between group theory (rotation, angular momentum) and the differential equation we are trying to solve should not come as a surprise.

13.4 Eigenvectors of $\mathbf{L}^2$: Spherical Harmonics

The treatment in the preceding section took place in an abstract vector space. Let us go back to the function space and represent the operators and vectors in terms of θ and φ.

First, let us consider $\mathbf{L}_z$ in the form of a differential operator, as given in Eq. (13.17). The eigenvalue equation for $\mathbf{L}_z$ becomes

$$-i\frac{\partial}{\partial\varphi}Y_{jm}(\theta,\varphi)=mY_{jm}(\theta,\varphi).$$

We write $Y_{jm}(\theta,\varphi)=P_{jm}(\theta)Q_{jm}(\varphi)$ and substitute in the above equation to obtain the ODE for φ, $dQ_{jm}/d\varphi=imQ_{jm}$, which has a solution of the form $Q_{jm}(\varphi)=C_{jm}e^{im\varphi}$, where C_{jm} is a constant. Absorbing this constant into P_{jm}, we can write

$$Y_{jm}(\theta,\varphi)=P_{jm}(\theta)e^{im\varphi}.$$

In classical physics the value of functions must be the same at φ as at $\varphi+2\pi$. This condition restricts the values of m to integers. In quantum mechanics, on the other hand, it is the absolute values of functions that are physically measurable quantities, and therefore m can also be a half-integer.

Box 13.4.1 *From now on, we shall assume that m is an integer and denote the eigenvectors of $\mathbf{L}^2$ by $Y_{lm}(\theta,\varphi)$, in which l is a nonnegative integer.*

Our task is to find an analytic expression for $Y_{lm}(\theta,\varphi)$. We need differential expressions for $\mathbf{L}_\pm$. These can easily be obtained from the expressions for $\mathbf{L}_x$ and $\mathbf{L}_y$ given in Eqs. (13.16) and (13.17). (The straightforward manipulations are left as a problem.) We thus have

$$\mathbf{L}_\pm=e^{\pm i\varphi}\left(\pm\frac{\partial}{\partial\theta}+i\cot\theta\frac{\partial}{\partial\varphi}\right). \tag{13.25}$$

Since l is the highest value of m, when $\mathbf{L}_+$ acts on $Y_{ll}(\theta,\varphi)=P_{ll}(\theta)e^{il\varphi}$ the result must be zero. This leads to the differential equation

$$\left(\frac{\partial}{\partial\theta}+i\cot\theta\frac{\partial}{\partial\varphi}\right)\left[P_{ll}(\theta)e^{il\varphi}\right]=0\quad\Rightarrow\quad\left(\frac{d}{d\theta}-l\cot\theta\right)P_{ll}(\theta)=0.$$

The solution to this differential equation is readily found to be

$$P_{ll}(\theta)=C_l(\sin\theta)^l. \tag{13.26}$$

The constant is subscripted because each P_{ll} may lead to a different constant of integration. We can now write

$$Y_{ll}(\theta,\varphi)=C_l(\sin\theta)^le^{il\varphi}.$$

With $Y_{ll}(\theta,\varphi)$ at our disposal, we can obtain any $Y_{lm}(\theta,\varphi)$ by repeated application of $\mathbf{L}_-$. In principle, the result of Example 13.3.3 gives all the (abstract) eigenvectors. In practice, however, it is helpful to have a closed form (in terms of derivatives) for just the θ part of $Y_{lm}(\theta,\varphi)$. So, let us apply $\mathbf{L}_-$, as given in Eq. (13.25) to $Y_{ll}(\theta,\varphi)$:

$$\begin{aligned}\mathbf{L}_- Y_{ll} &= e^{-i\varphi}\left(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}\right)\left[P_{ll}(\theta)e^{il\varphi}\right]\\ &= e^{-i\varphi}\left[-\frac{\partial}{\partial\theta} + i\cot\theta(il)\right]\left[P_{ll}(\theta)e^{il\varphi}\right]\\ &= (-1)e^{i(l-1)\varphi}\left(\frac{d}{d\theta} + l\cot\theta\right)P_{ll}(\theta).\end{aligned}$$

It can be shown that for a positive integer,

$$\left(\frac{d}{d\theta} + n\cot\theta\right)f(\theta) = \frac{1}{\sin^n\theta}\frac{d}{d\theta}\left[\sin^n\theta f(\theta)\right]. \tag{13.27}$$

Using this result and (13.26) yields

$$\begin{aligned}\mathbf{L}_- Y_{ll} &= (-1)e^{i(l-1)\varphi}\frac{1}{\sin^l\theta}\frac{d}{d\theta}\left[\sin^l\theta\left(C_l\sin^l\theta\right)\right]\\ &= (-1)C_l\frac{e^{i(l-1)\varphi}}{\sin^l\theta}\frac{d}{d\theta}\left(\sin^{2l}\theta\right).\end{aligned} \tag{13.28}$$

We apply $\mathbf{L}_-$ to (13.28), and use Eq. (13.27) with $n = l-1$ to obtain

$$\begin{aligned}\mathbf{L}_-^2 Y_{ll} &= (-1)^2 C_l e^{i(l-2)\varphi}\frac{1}{\sin^{l-1}\theta}\frac{d}{d\theta}\left[\sin^{l-1}\theta\frac{1}{\sin^l\theta}\frac{d}{d\theta}\left(\sin^{2l}\theta\right)\right]\\ &= (-1)^2 C_l\frac{e^{i(l-2)\varphi}}{\sin^{l-1}\theta}\frac{d}{d\theta}\left[\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin^{2l}\theta\right)\right].\end{aligned}$$

Making the substitution $u = \cos\theta$ yields

$$\mathbf{L}_-^2 Y_{ll} = C_l\frac{e^{i(l-2)\varphi}}{(1-u^2)^{l/2-1}}\frac{d^2}{du^2}\left[\left(1-u^2\right)^l\right].$$

With a little more effort one can detect a pattern and obtain

$$\mathbf{L}_-^k Y_{ll} = C_l\frac{e^{i(l-k)\varphi}}{(1-u^2)^{(l-k)/2}}\frac{d^k}{du^k}\left[\left(1-u^2\right)^l\right].$$

If we let $k = l-m$ and make use of the result obtained in Example 13.3.3, we obtain

$$Y_{lm}(\theta,\varphi) = \sqrt{\frac{(l+m)!}{(l-m)!(2l)!}}C_l\frac{e^{im\varphi}}{(1-u^2)^{m/2}}\frac{d^{l-m}}{du^{l-m}}\left[\left(1-u^2\right)^l\right].$$

To specify $Y_{lm}(\theta, \varphi)$ completely, we need to evaluate C_l. Since C_l does not depend on m, we set $m = 0$ in the above expression, obtaining

$$Y_{l0}(u, \varphi) = \frac{1}{\sqrt{(2l)!}} C_l \frac{d^l}{du^l}\left[\left(1 - u^2\right)^l\right].$$

The RHS looks very much like the Legendre polynomial of Chap. 8. In fact,

$$Y_{l0}(u, \varphi) = \frac{C_l}{\sqrt{(2l)!}}(-1)^l 2^l l! P_l(u) \equiv A_l P_l(u). \tag{13.29}$$

Therefore, the normalization of Y_{l0} and the Legendre polynomials P_l determines C_l.

We now use Eq. (7.25) to obtain the integral form of the orthonormality relation for Y_{lm}:

$$\begin{aligned}\delta_{ll'}\delta_{mm'} &= \langle Y_{l'm'}|Y_{lm}\rangle = \langle Y_{l'm'}|\left(\int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta\, d\theta |\theta, \varphi\rangle\langle\theta, \varphi|\right)|Y_{lm}\rangle \\ &= \int_0^{2\pi} d\varphi \int_0^{\pi} Y^*_{l'm'}(\theta, \varphi) Y_{lm}(\theta, \varphi) \sin\theta\, d\theta,\end{aligned} \tag{13.30}$$

which in terms of $u = \cos\theta$ becomes

$$\int_0^{2\pi} d\varphi \int_{-1}^{1} Y^*_{l'm'}(u, \varphi) Y_{lm}(u, \varphi) du = \delta_{ll'}\delta_{mm'}. \tag{13.31}$$

Problem 13.15 shows that using (13.30) one gets $A_l = \sqrt{(2l+1)/(4\pi)}$. Therefore, Eq. (13.29) yields not only the value of C_l, but also the useful relation

$$Y_{l0}(u, \varphi) = \sqrt{\frac{2l+1}{4\pi}} P_l(u). \tag{13.32}$$

Substituting the value of C_l thus obtained, we finally get

$$\begin{aligned}Y_{lm}(\theta, \varphi) &= (-1)^l \sqrt{\frac{2l+1}{4\pi}} \frac{e^{im\varphi}}{2^l l!} \sqrt{\frac{(l+m)!}{(l-m)!}} \left(1 - u^2\right)^{-m/2} \\ &\quad \times \frac{d^{l-m}}{du^{l-m}}\left[\left(1 - u^2\right)^l\right],\end{aligned} \tag{13.33}$$

where $u = \cos\theta$. These functions, the eigenfunctions of $\mathbf{L}^2$ and $\mathbf{L}_z$, are called **spherical harmonics**. They occur frequently in those physical applications for which the Laplacian is expressed in terms of spherical coordinates.

spherical harmonics

One can immediately read off the θ part of the spherical harmonics:

$$P_{lm}(u) = (-1)^l \sqrt{\frac{2l+1}{4\pi}} \frac{1}{2^l l!} \sqrt{\frac{(l+m)!}{(l-m)!}} \left(1 - u^2\right)^{-m/2} \frac{d^{l-m}}{du^{l-m}}\left[\left(1 - u^2\right)^l\right].$$

associated Legendre functions

However, this is not the version used in the literature. For historical reasons the **associated Legendre functions** $P_l^m(u)$ are used. These are defined by

$$\begin{aligned} P_l^m(u) &= (-1)^m \sqrt{\frac{(l+m)!}{(l-m)!}} \sqrt{\frac{4\pi}{2l+1}} P_{lm}(u) \\ &= (-1)^{l+m} \frac{(l+m)!}{(l-m)!} \frac{(1-u^2)^{-m/2}}{2^l l!} \frac{d^{l-m}}{du^{l-m}} \left[\left(1-u^2\right)^l \right]. \end{aligned}$$

Thus,

Box 13.4.2 *The solutions of the angular part of the Laplacian are*

$$Y_{lm}(\theta, \varphi) = (-1)^m \left[\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!} \right]^{1/2} P_l^m(\cos\theta) e^{im\varphi}, \quad (13.34)$$

where, with $u = \cos\theta$,

$$P_l^m(u) = (-1)^{l+m} \frac{(l+m)!}{(l-m)!} \frac{(1-u^2)^{-m/2}}{2^l l!} \frac{d^{l-m}}{du^{l-m}} \left[\left(1-u^2\right)^l \right]. \quad (13.35)$$

We generated the spherical harmonics starting with $Y_{ll}(\theta, \varphi)$ and applying the lowering operator $\mathbf{L}_-$. We could have started with $Y_{l,-l}(\theta, \varphi)$ instead, and applied the raising operator $\mathbf{L}_+$. The latter procedure is identical to the former; nevertheless, we outline it below because of some important relations that emerge along the way. We first note that

$$|Y_{l,-m}\rangle = \sqrt{\frac{(l+m)!}{(l-m)!(2l)!}} \mathbf{L}_+^{l-m} |Y_{l,-l}\rangle. \quad (13.36)$$

(This can be obtained following the steps of Example 13.3.3.) Next, we use $\mathbf{L}_-|Y_{l,-l}\rangle = 0$ in differential form to obtain

$$\left(\frac{d}{d\theta} - l\cot\theta \right) P_{l,-l}(\theta) = 0,$$

which has the same form as the differential equation for P_{ll}. Thus, the solution is $P_{l,-l}(\theta) = C_l'(\sin\theta)^l$, and

$$Y_{l,-l}(\theta, \varphi) = P_{l,-l}(\theta) e^{-il\varphi} = C_l'(\sin\theta)^l e^{-il\varphi}.$$

Applying $\mathbf{L}_+$ repeatedly yields

$$\mathbf{L}_+^k Y_{l,-l}(u, \varphi) = C_l' \frac{(-1)^k e^{-i(l-k)\varphi}}{(1-u^2)^{(l-k)/2}} \frac{d^k}{du^k} \left[\left(1-u^2\right)^l \right],$$

where $u = \cos\theta$. Substituting $k = l - m$ and using Eq. (13.36) gives

$$Y_{l,-m}(u, \varphi) = \sqrt{\frac{(l+m)!}{(l-m)!(2l)!}} C_l' \frac{(-1)^{l-m} e^{-im\varphi}}{(1-u^2)^{m/2}} \frac{d^{l-m}}{du^{l-m}} \left[\left(1-u^2\right)^l \right].$$

The constant C'_l can be determined as before. In fact, for $m = 0$ we get exactly the same result as before, so we expect C'_l to be identical to C_l. Thus,

$$Y_{l,-m}(u,\varphi) = (-1)^{l+m}\sqrt{\frac{2l+1}{4\pi}}\frac{e^{-im\varphi}}{2^l l!}\sqrt{\frac{(l+m)!}{(l-m)!}} \times \left(1-u^2\right)^{-m/2}\frac{d^{l-m}}{du^{l-m}}\left[\left(1-u^2\right)^l\right].$$

Comparison with Eq. (13.33) yields

$$Y_{l,-m}(\theta,\varphi) = (-1)^m Y^*_{l,m}(\theta,\varphi), \tag{13.37}$$

and using the definition $Y_{l,-m}(\theta,\varphi) = P_{l,-m}(\theta)e^{-im\varphi}$ and the first part of Eq. (13.35), we obtain

$$P_l^{-m}(\theta) = (-1)^m\frac{(l-m)!}{(l+m)!}P_l^m(\theta). \tag{13.38}$$

The first few spherical harmonics with positive m are given below. Those with negative m can be obtained using Eq. (13.37).

For $l = 0$, $\quad Y_{00} = \dfrac{1}{\sqrt{4\pi}}$.

For $l = 1$, $\quad Y_{10} = \sqrt{\dfrac{3}{4\pi}}\cos\theta, \qquad Y_{11} = -\sqrt{\dfrac{3}{8\pi}}e^{i\varphi}\sin\theta.$

For $l = 2$,
$$Y_{20} = \sqrt{\frac{5}{16\pi}}\left(3\cos^2\theta - 1\right),$$
$$Y_{21} = -\sqrt{\frac{15}{8\pi}}e^{i\varphi}\sin\theta\cos\theta,$$
$$Y_{22} = \sqrt{\frac{15}{32\pi}}e^{2i\varphi}\sin^2\theta.$$

For $l = 3$,
$$Y_{30} = \sqrt{\frac{7}{16\pi}}\left(5\cos^3\theta - 3\cos\theta\right),$$
$$Y_{31} = -\sqrt{\frac{21}{64\pi}}e^{i\varphi}\sin\theta\left(5\cos^2\theta - 1\right),$$
$$Y_{32} = \sqrt{\frac{105}{32\pi}}e^{2i\varphi}\sin^2\theta\cos\theta,$$
$$Y_{33} = -\sqrt{\frac{35}{64\pi}}e^{3i\varphi}\sin^3\theta.$$

From Eqs. (13.13), (13.18), and (13.34) and the fact that $\alpha = l(l+1)$ for some nonnegative integer l, we obtain

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\left[P_l^m e^{im\varphi}\right]+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\left[P_l^m e^{im\varphi}\right]+l(l+1)P_l^m e^{im\varphi}$$
$$=0,$$

which gives

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dP_l^m}{d\theta}\right)-\frac{m^2}{\sin^2\theta}P_l^m+l(l+1)P_l^m=0.$$

As before, we let $u=\cos\theta$ to obtain

$$\frac{d}{du}\left[\left(1-u^2\right)\frac{dP_l^m}{du}\right]+\left[l(l+1)-\frac{m^2}{1-u^2}\right]P_l^m=0. \tag{13.39}$$

This is called the **associated Legendre differential equation**. Its solutions, the associated Legendre functions, are given in closed form in Eq. (13.35). For $m=0$, Eq. (13.39) reduces to the Legendre differential equation whose solutions, again given by Eq. (13.35) with $m=0$, are the Legendre polynomials encountered in Chap. 8. When $m=0$, the spherical harmonics become φ-independent. This corresponds to a physical situation in which there is an explicit azimuthal symmetry. In such cases (when it is obvious that the physical property in question does not depend on φ) a Legendre polynomial, depending only on $\cos\theta$, will multiply the radial function.

associated Legendre differential equation

13.4.1 Expansion of Angular Functions

The orthonormality of spherical harmonics can be utilized to expand functions of θ and φ in terms of them. The fact that these functions are complete will be discussed in a general way in the context of Sturm-Liouville systems. Assuming completeness for now, we write

$$f(\theta,\varphi)=\begin{cases}\sum_{l=0}^{\infty}\sum_{m=-l}^{l}a_{lm}Y_{lm}(\theta,\varphi) & \text{if } l \text{ is not fixed,}\\ \sum_{m=-l}^{l}a_{lm}Y_{lm}(\theta,\varphi) & \text{if } l \text{ is fixed,}\end{cases} \tag{13.40}$$

where we have included the case where it is known a priori that $f(\theta,\varphi)$ has a given fixed l value. To find a_{lm}, we multiply both sides by $Y_{lm}^*(\theta,\varphi)$ and integrate over the solid angle. The result, obtained by using the orthonormality relation, is

$$a_{lm}=\iint d\Omega f(\theta,\varphi)Y_{lm}^*(\theta,\varphi), \tag{13.41}$$

where $d\Omega\equiv\sin\theta\,d\theta\,d\varphi$ is the element of solid angle. A useful special case of this formula is

$$\begin{aligned}a_{l0}^{(f)}&=\iint d\Omega f(\theta,\varphi)Y_{l0}^*(\theta,\varphi)\\ &=\sqrt{\frac{2l+1}{4\pi}}\iint d\Omega f(\theta,\varphi)P_l(\cos\theta),\end{aligned} \tag{13.42}$$

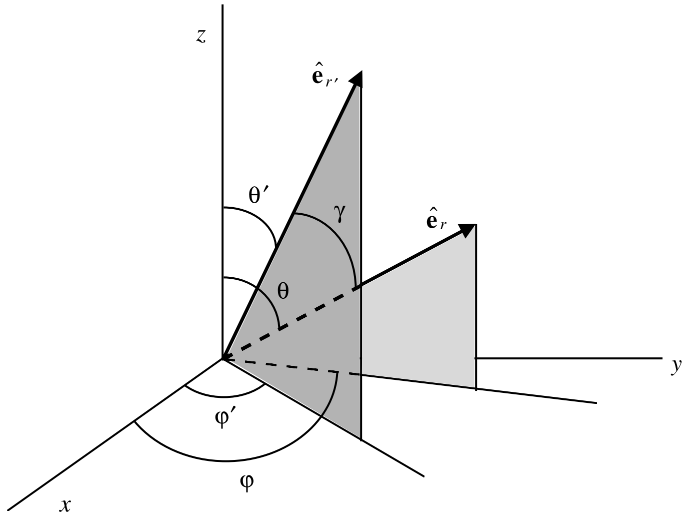

Fig. 13.1 The unit vectors $\hat{\mathbf{e}}_r$ and $\hat{\mathbf{e}}_{r'}$ with their spherical angles and the angle γ between them

where we have introduced an extra superscript to emphasize the relation of the expansion coefficients with the function being expanded. Another useful relation is obtained when we let $\theta = 0$ in Eq. (13.40):

$$f(\theta,\varphi)\big|_{\theta=0} = \begin{cases} \sum_{l=0}^{\infty}\sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta,\varphi)|_{\theta=0} & \text{if } l \text{ is not fixed,} \\ \sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta,\varphi)|_{\theta=0} & \text{if } l \text{ is fixed.} \end{cases}$$

From Eqs. (13.35) and (13.34) one can show that

$$Y_{lm}(\theta,\varphi)\big|_{\theta=0} = \delta_{m0} Y_{l0}(0,\varphi) = \delta_{m0}\sqrt{\frac{2l+1}{4\pi}}.$$

Therefore,

$$f(\theta,\varphi)\big|_{\theta=0} = \begin{cases} \sum_{l=0}^{\infty} a_{l0}^{(f)} \sqrt{\frac{2l+1}{4\pi}} & \text{if } l \text{ is not fixed,} \\ a_{l0}^{(f)} \sqrt{\frac{2l+1}{4\pi}} & \text{if } l \text{ is fixed.} \end{cases} \tag{13.43}$$

13.4.2 Addition Theorem for Spherical Harmonics

addition theorem for spherical harmonics

An important consequence of the expansion in terms of Y_{lm} is called the **addition theorem** for spherical harmonics. Consider two unit vectors $\hat{\mathbf{e}}_r$ and $\hat{\mathbf{e}}_{r'}$ making spherical angles (θ,φ) and (θ',φ'), respectively, as shown in Fig. 13.1. Let γ be the angle between the two vectors. The addition theorem states that

$$P_l(\cos\gamma) = \frac{4\pi}{2l+1}\sum_{m=-l}^{l} Y_{lm}^*(\theta',\varphi') Y_{lm}(\theta,\varphi). \tag{13.44}$$

We shall not give a proof of this theorem here and refer the reader to an elegant proof on page 974 which uses the representation theory of groups. The addition theorem is particularly useful in the expansion of the frequently occurring expression $1/|\mathbf{r}-\mathbf{r}'|$. For definiteness we assume $|\mathbf{r}'| \equiv r' < |\mathbf{r}| \equiv r$. Then, introducing $t = r'/r$, we have

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \frac{1}{(r^2+r'^2-2rr'\cos\gamma)^{1/2}} = \frac{1}{r}\left(1+t^2-2t\cos\gamma\right)^{-1/2}.$$

Recalling the generating function for Legendre polynomials from Chap. 8 and using the addition theorem, we get

$$\begin{aligned}\frac{1}{|\mathbf{r}-\mathbf{r}'|} &= \frac{1}{r}\sum_{l=0}^{\infty} t^l P_l(\cos\gamma) = \sum_{l=0}^{\infty} \frac{r'^l}{r^{l+1}} \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_{lm}^*\left(\theta',\varphi'\right) Y_{lm}(\theta,\varphi) \\ &= 4\pi \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \frac{1}{2l+1} \frac{r'^l}{r^{l+1}} Y_{lm}^*\left(\theta',\varphi'\right) Y_{lm}(\theta,\varphi).\end{aligned}$$

It is clear that if $r < r'$, we should expand in terms of the ratio r/r'. It is therefore customary to use $r_<$ to denote the smaller and $r_>$ to denote the larger of the two radii r and r'. Then the above equation is written as

expansion of $1/|\mathbf{r}-\mathbf{r}'|$ in spherical coordinates

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = 4\pi \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \frac{1}{2l+1} \frac{r_<^l}{r_>^{l+1}} Y_{lm}^*\left(\theta',\varphi'\right) Y_{lm}(\theta,\varphi). \tag{13.45}$$

This equation is used frequently in the study of Coulomb-like potentials.

13.5 Problems

13.1 By applying the *operator* $[x_j, \mathbf{p}_k]$ to an arbitrary function $f(\mathbf{r})$, show that $[x_j, \mathbf{p}_k] = i\delta_{jk}$.

13.2 Use the defining relation $\mathbf{L}_i = \epsilon_{ijk} x_j \mathbf{p}_k$ to show that $x_j\mathbf{p}_k - x_k\mathbf{p}_j = \epsilon_{ijk}\mathbf{L}_i$. In both of these expressions a sum over the repeated indices is understood.

13.3 For the angular momentum operator $\mathbf{L}_i = \epsilon_{ijk} x_j \mathbf{p}_k$, show that the commutation relation $[\mathbf{L}_j, \mathbf{L}_k] = i\epsilon_{jkl}\mathbf{L}_l$ holds.

13.4 Evaluate $\partial f/\partial y$ and $\partial f/\partial z$ in spherical coordinates and find $\mathbf{L}_y$ and $\mathbf{L}_z$ in terms of spherical coordinates.

13.5 Obtain an expression for $\mathbf{L}^2$ in terms of θ and φ, and substitute the result in Eq. (13.12) to get the Laplacian in spherical coordinates.

13.6 Show that $\mathbf{L}^2 = \mathbf{L}_+\mathbf{L}_- + \mathbf{L}_z^2 - \mathbf{L}_z$ and $\mathbf{L}^2 = \mathbf{L}_-\mathbf{L}_+ + \mathbf{L}_z^2 + \mathbf{L}_z$.

13.7 Show that $\mathbf{L}^2$, $\mathbf{L}_x$, $\mathbf{L}_y$, and $\mathbf{L}_z$ are hermitian operators in the space of square-integrable functions.

13.8 Verify the following commutation relations:

$$\left[\mathbf{L}^2, \mathbf{L}_\pm\right] = 0, \qquad [\mathbf{L}_z, \mathbf{L}_\pm] = \pm\mathbf{L}_\pm, \qquad [\mathbf{L}_+, \mathbf{L}_-] = 2\mathbf{L}_z.$$

13.9 Show that $\mathbf{L}_-|Y_{\alpha\beta}\rangle$ has $\beta - 1$ as its eigenvalue for $\mathbf{L}_z$, and that $|Y_{\alpha,\beta_\pm}\rangle$ cannot be zero.

13.10 Show that if the $|Y_{jm}\rangle$ are normalized to unity, then with proper choice of phase, $\mathbf{L}_-|Y_{jm}\rangle = \sqrt{j(j+1) - m(m-1)}|Y_{j,m-1}\rangle$.

13.11 Derive Eq. (13.36).

13.12 Starting with $\mathbf{L}_x$ and $\mathbf{L}_y$, derive the following expression for $\mathbf{L}_\pm$:

$$\mathbf{L}_\pm = e^{\pm i\varphi}\left(\pm\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}\right).$$

13.13 Integrate $dP/d\theta - l\cot\theta P = 0$ to find $P(\theta)$.

13.14 Verify the following differential identity:

$$\left(\frac{d}{d\theta} + n\cot\theta\right)f(\theta) = \frac{1}{\sin^n\theta}\frac{d}{d\theta}\left[\sin^n\theta f(\theta)\right].$$

13.15 Let $l = l'$ and $m = m' = 0$ in Eq. (13.31), and substitute for Y_{l0} from Eq. (13.29) to obtain $A_l = \sqrt{(2l+1)/4\pi}$.

13.16 Show that

$$\mathbf{L}_+^k Y_{l,-l}(u,\varphi) = C_l'\frac{(-1)^k e^{-i(l-k)\varphi}}{(1-u^2)^{(l-k)/2}}\frac{d^k}{du^k}\left[\left(1-u^2\right)^l\right].$$

13.17 Derive the relations $Y_{l,-m}(\theta,\varphi) = (-1)^m Y_{l,m}^*(\theta,\varphi)$ and

$$P_l^{-m}(\theta) = (-1)^m\frac{(l-m)!}{(l+m)!}P_l^m(\theta).$$

13.18 Show that

$$\sum_{m=-l}^{l}\left|Y_{lm}(\theta,\varphi)\right|^2 = \frac{2l+1}{4\pi}.$$

Verify this explicitly for $l = 1$ and $l = 2$.

13.19 Show that the addition theorem for spherical harmonics can be written as

$$P_l(\cos\gamma) = P_l(\cos\theta)P_l\bigl(\cos\theta'\bigr) + 2\sum_{m=1}^{l}\frac{(l-m)!}{(l+m)!}P_l^m(\cos\theta)P_l^m\bigl(\cos\theta'\bigr)\cos\bigl[m\bigl(\varphi-\varphi'\bigr)\bigr].$$