A User's Guide to Spherical Harmonics

Martin J. Mohlenkamp*

Version: October 18, 2016

This pamphlet is intended for the scientist who is considering using Spherical Harmonics for some application. It is designed to introduce the Spherical Harmonics from a theoretical perspective and then discuss those practical issues necessary for their use in applications.

I expect great variability in the backgrounds of the readers. In order to be informative to those who know less, without insulting those who know more, I have adopted the strategy "State even basic facts, but briefly."

My dissertation work was a fast transform for Spherical Harmonics [6]. After its publication, I started receiving questions about Spherical Harmonics in general, rather than about my work specifically. This pamphlet aims to answer such general questions. If you find mistakes, or feel that important material is unclear or missing, please inform me.

1 A Theory of Spherical Harmonics

In this section we give a development of Spherical Harmonics. There are other developments from other perspectives. The one chosen here has the benefit of being very concrete.

1.1 Mathematical Preliminaries

We define the L^2 inner product of two functions to be

$$\langle f, g\rangle = \int f(s)\bar{g}(s)ds \tag{1}$$

where $\bar{g}$ denotes complex conjugation and the integral is over the space of interest, for example $\int_0^{2\pi} d\theta$ on the circle or $\int_0^{2\pi}\int_0^{\pi} \sin\phi d\phi d\theta$ on the sphere. We define the L^2 norm by $||f|| = \sqrt{\langle f, f\rangle}$. The space L^2 consists of all functions such that $||f|| < \infty$.

A *basis* for L^2 is a set of functions $\{\psi_i\}$ with the properties

Orthogonal: $\langle \psi_i, \psi_j\rangle = 0$ for $i \neq j$.

Normalized: $||\psi_i|| = 1$.

Spanning: Any function in L^2 can be written as a linear combination of the ψ_i's; $f = \sum_i \alpha_i \psi_i$, α_i complex numbers.

The notation $f(x) = \mathcal{O}(g(x))$ means $\lim(|f(x)/g(x)|) < C < \infty$ where the limit is toward the point of interest, usually ∞.

*Department of Mathematics, Ohio University, mohlenka@ohio.edu

1.2 $\mathbf{R}^2$: Fourier Series

Spherical Harmonics can be generated in the same way as Fourier Series, simply in one dimension higher. One cannot understand Spherical Harmonics without first understanding Fourier series. Conversely, a good understanding of Fourier series will get you much of the way through the analysis of Spherical Harmonics. In this section we develop the basic theory of Fourier series, from a perspective that will translate well when we move to Spherical Harmonics. Much of this discussion is based on Stein and Weiss [7, Chapter IV].

We start with something familiar, polynomials in $\mathbf{R}^2$, $\{(x, y) : x, y \in \mathbf{R}\}$. Then we restrict our attention to those polynomials that are *harmonic*, meaning $\Delta_2 p(x, y) = 0$, where Δ_2 is the two-dimensional *Laplacian*

$$\Delta_2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}. \tag{2}$$

For example, $\Delta_2(x + x^2 - y^2) = 0$.

Next we further restrict our consideration to *homogeneous* harmonic polynomials. A function $p(x, y)$ is *homogeneous of degree* n if $p(tx, ty) = t^n p(x, y)$ for all positive real numbers t. Examples of homogeneous harmonic polynomials of degrees 0, 1, 2, and 3 are 6, x, $x^2 - y^2$, and $3x^2y - y^3$. Homogeneous functions are nicely represented in polar coordinates $\{(r, \theta) : r \in \mathbf{R}_+, \theta \in [0, 2\pi)\}$, where $(x, y) = (r \cos \theta, r \sin \theta)$. A function $p_n(x, y)$ that is homogeneous of degree n can be written as $p_n(x, y) = r^n q_n(\theta)$ for some function q_n. In polar coordinates the Laplacian (2) becomes

$$\Delta_2 = \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}. \tag{3}$$

A homogeneous harmonic polynomial of degree n, $(n \geq 0)$ satisfies

$$0 = \Delta_2 p_n(x, y) = n(n-1)r^{n-2}q_n(\theta) + \frac{1}{r}nr^{n-1}q_n(\theta) + \frac{1}{r^2}r^n q_n''(\theta) = r^{n-2}\left(n^2 q_n(\theta) + q_n''(\theta)\right). \tag{4}$$

The operator Δ_2 acts on homogeneous functions in a simple way, separately in r and θ. In r it simply reduces the power by two. In θ it acts using the restriction of Δ_2 to the circle, i.e. the *Circular Laplacian* $\Delta_{2S} = \frac{\partial^2}{\partial \theta^2}$. For (4) to hold, q_n must be an eigenfunction of Δ_{2S}, i.e. $\Delta_{2S} q_n = -n^2 q_n$.

The operator Δ_{2S} is *self-adjoint* on the circle, meaning

$$\langle \Delta_{2S} f(\theta), g(\theta) \rangle = \langle f(\theta), \Delta_{2S} g(\theta) \rangle. \tag{5}$$

This fact follows by integration by parts. Self-adjoint operators have the property that eigenfunctions with different eigenvalues are orthogonal. We have

$$-n^2\langle q_n(\theta), q_m(\theta)\rangle = \langle \Delta_{2S} q_n(\theta), q_m(\theta)\rangle = \langle q_n(\theta), \Delta_{2S} q_m(\theta)\rangle = -m^2\langle q_n(\theta), q_m(\theta)\rangle, \tag{6}$$

so if $n \neq m$, we must have $\langle q_n(\theta), q_m(\theta)\rangle = 0$.

The next question is whether there can be more than one q_n for the same n. The q_n satisfy $q_n''(\theta) = -n^2 q(\theta)$. All solutions of this differential equation are of the form $q_n(\theta) = A\cos(n\theta) + B\sin(n\theta)$ or equivalently $q_n(\theta) = Ce^{in\theta} + De^{-in\theta}$. Thus there are many possibilities for $q_n(\theta)$, but we can generate them all as linear combinations of two elements, $\{e^{in\theta}, e^{-in\theta}\}$, chosen to be orthogonal. The eigenspaces are thus two dimensional. (For $n = 0$ it is only one dimensional.) This construction, now complete, has generated the Fourier series.

Let us consider next what we have given up by using only homogeneous, harmonic polynomials.

Theorem 1 ([7, p.139])**.** *Any polynomial of degree n, when restricted to the circle, can be written as a sum of homogeneous harmonic polynomials of degree at most n.*

On the circle itself, we lose nothing by our restrictions. Again restricted to the circle, polynomials can approximate any continuous function, and thus any L^2 function, to any desired degree of accuracy. (Polynomials are dense in $L^2(S^1)$.) The exponentials $\{e^{in\theta}\}_{n\in\mathbf{Z}}$ thus span L^2. After normalization, they will be a basis. There are of course independent proofs that the Fourier series are a basis for $L^2(S^1)$.

To recap, we have constructed the Fourier series basis for the circle by

1. restricting Δ_2 to the circle,
2. decomposing into the eigenspaces of Δ_{2S}, and then
3. taking a convenient basis within each eigenspace.

1.3 $\mathbf{R}^3$: Ordinary Spherical Harmonics

The construction in Section 1.2 can be done in any dimension. The sphere S^1 in $\mathbf{R}^2$ is normally called the circle, but we could equally well call it a sphere and say the Fourier Series are Spherical Harmonics. The usual usage for Spherical Harmonics refers to the Surface Spherical Harmonics on the sphere S^2 in $\mathbf{R}^3$. We shall follow this usage and examine this case in this section.

In $\mathbf{R}^3$, $\{(x, y, z) : x, y, z \in \mathbf{R}\}$, the Laplacian is

$$\Delta_3 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{7}$$

In spherical coordinates $\{(r, \theta, \phi) : r \in \mathbf{R}_+, \theta \in [0, 2\pi), \phi \in [0, \pi]\}$, where $(x, y, z) = (r \cos\theta \sin\phi, r \sin\theta \sin\phi, r \cos\phi)$, this means

$$\Delta_3 = \csc^2\phi \frac{\partial^2}{\partial\theta^2} + \frac{\partial^2}{\partial\phi^2} + \cot\phi \frac{\partial}{\partial\phi} + 2r\frac{\partial}{\partial r} + r^2\frac{\partial^2}{\partial r^2}\,. \tag{8}$$

Suppose, as above, we have a homogeneous harmonic polynomial of degree n, $p_n(x, y, z) = r^n q_n(\theta, \phi)$. We then have

$$0 = \Delta_3 p_n = \left[\csc^2\phi \frac{\partial^2}{\partial\theta^2} + \frac{\partial^2}{\partial\phi^2} + \cot\phi \frac{\partial}{\partial\phi}\right] r^n q_n(\theta, \phi) + 2rnr^{n-1} q_n(\theta, \phi) + r^2 n(n-1) r^{n-2} q_n(\theta, \phi) \tag{9}$$

$$= r^n \left(\left[\csc^2\phi \frac{\partial^2}{\partial\theta^2} + \frac{\partial^2}{\partial\phi^2} + \cot\phi \frac{\partial}{\partial\phi}\right] + n(n+1)\right) q_n(\theta, \phi)\,. \tag{10}$$

Defining the *Spherical Laplacian* by

$$\Delta_{3S} = \csc^2\phi \frac{\partial^2}{\partial\theta^2} + \frac{\partial^2}{\partial\phi^2} + \cot\phi \frac{\partial}{\partial\phi}\,, \tag{11}$$

from (10) we have

$$\Delta_{3S} q_n(\theta, \phi) = -n(n+1) q_n(\theta, \phi)\,. \tag{12}$$

This q_n is therefore an eigenfunction of the Spherical Laplacian. Any such eigenfunction is called a *Spherical Harmonic*.

As before, we note that Δ_{3S} is self-adjoint, which implies that the eigenspaces Λ_n are orthogonal. The space Λ_n consists of the homogeneous harmonic polynomials of degree n restricted to the sphere, and has dimension $2n + 1$. On the sphere, the homogeneous harmonic polynomials span the set of all polynomials, which in turn are dense in L^2. Our spherical harmonics therefore span L^2. If we take a basis within each eigenspace then this collection will give a basis for L^2 of the sphere. Thus Spherical Harmonics arise in $\mathbf{R}^3$ in the same way Fourier series arise in $\mathbf{R}^2$. They are consequently sometimes called 'Fourier series on the Sphere'. (Other less deserving objects sometimes also go by this name in the literature.)

For fixed n, we can organize a basis for Λ_n as

$$\left\{ \frac{e^{im\theta}}{\sqrt{2\pi}} \tilde{P}_n^m(\cos\phi) \right\}_{-n \le m \le n} . \tag{13}$$

It is not immediately obvious that we can separate variables and assume exponential functions in the θ direction. We are able to do this essentially because the lines of fixed ϕ are circles. We could also simply assume this form and show the construction succeeds. This organization is not forced, but separating the variables is so useful that there are no competitive options. A disadvantage of this organization is that it makes the poles into special points.

We would like to find the conditions on $\tilde{P}_n^m$ to make (13) a set of (smooth) spherical harmonics. From (12), we need to have

$$\Delta_{3S} \frac{e^{im\theta}}{\sqrt{2\pi}} \tilde{P}_n^m(\cos\phi) = \left[\frac{\partial^2}{\partial\phi^2} + \cot\phi \frac{\partial}{\partial\phi} - \frac{m^2}{\sin^2\phi} \right] \frac{e^{im\theta}}{\sqrt{2\pi}} \tilde{P}_n^m(\cos\phi) = -n(n-1) \frac{e^{im\theta}}{\sqrt{2\pi}} \tilde{P}_n^m(\cos\phi). \tag{14}$$

This equation, and the condition that $\tilde{P}_n^m(1) \neq \pm\infty$, identifies the $\tilde{P}_n^m$'s up to a constant as the *Associated Legendre Functions (of the first kind) of order m and degree n.* We use the tilde ($\tilde{\cdot}$) to indicate the L^2 normalized version of the classically defined Associated Legendre functions, denoted P_n^m. $\tilde{P}_n^m$ can be constructed explicitly, and we do this in Section 3.1.

Note that for each fixed m, the set $\{\tilde{P}_n^m(\cos\phi)\}_{n \ge |m|}$ will need to be orthogonal with respect to the measure $\sin\phi d\phi$, so that harmonics from different Λ_n will be orthogonal.

1.4 ...and Why Do We Care?

In this section we consider some of the applications of Spherical Harmonics. We do this mainly by showing some of the applications of Fourier Series and seeing how they extend. These results are well known.

Fourier Series are a basis for the circle. They have the nice property that the eigenspaces Λ_n are rotation-invariant. Since the Laplacian does not care where the coordinates are, its eigenspaces do not either. If we restrict ourselves to functions generated by the eigenspaces up to degree N, this restriction is 'fair' in the sense that all portions of the circle are treated equally. The rotation-invariance of Λ_n also implies that if we know the expansion of a function, we can compute the expansion of any of its rotations by applying a sparse matrix. This matrix is zero except for 2×2 blocks along the diagonal, one block for each n. If we have a band-width limitation $n < N$, this means we can apply this matrix in $2N$ time, instead of $4N^2$ time.

Spherical Harmonics work similarly. The eigenspaces are also rotation-invariant, though now they are not of the same size. The matrix of a rotation in this basis consists of blocks along the diagonal, with a $(2n+1) \times (2n+1)$ block for Λ_n. (These are called Wigner rotation matrices.) If we again have a limit $n < N$, we can apply the matrix in $\sum(2n+1)^2 = (4N^3 - N)/3$ time, instead of $[\sum(2n+1)]^2 = N^4$ time. A decomposition using Λ_n for $n < N$ is again 'fair' to all points on the sphere.

Next we consider operators which are linear and rotation-invariant. Many operators which correspond to physical processes (gravity, heat, electricity) are coordinate-invariant. In particular, if the process is taking place on a circle or sphere, these operators commute with rotations. On the circle, rotating $e^{in\theta}$ through an angle ν is the same as multiplying by $e^{in\nu}$. For such an operator L, this means

$$(Le^{in\theta})(\xi) = (Le^{in(\theta-\nu)})(\xi+\nu) = e^{-in\nu}(Le^{in\theta})(\xi+\nu) \,. \tag{15}$$

Since the left side of (15) is independent of ν, the right side must also be is independent of ν, which implies $(Le^{in\theta})(\xi+\nu) = \lambda_n e^{in(\xi+\nu)}$ for some constant λ_n. Thus the Fourier Series are eigenfunctions for *all* such operators. Since they are also a basis, this means they diagonalize such operators, making them trivial to apply. For the Spherical Harmonics, we have the same result, and the proof is essentially the same.

Example 1 (Convolution)**.** *The well used property that Fourier series convert convolution to multiplication is equivalent to noting the Fourier Series are eigenfunctions, as we did above. For Spherical Harmonics we have the similar result*

$$(f * g)_n^m = f_n^m g_n^0. \tag{16}$$

See e.g. [3, p.210] for its discussion.

2 Understanding the Spherical Harmonics

There is a method for estimating the solutions to some types of differential equations, in particular Schrödinger equations, known variously as the Quasi-Classical, Semi-Classical, or WKB method, or more generally as an asymptotic analysis. See e.g. Landau and Lifschitz [5, Chapter VII] or Bender and Orszag [2]. We will use this method to get a qualitative description of the Associated Legendre functions.

Supposing we have a function $\psi(x)$ which satisfies a Schrödinger equation

$$\frac{d^2}{dx^2}\psi(x) = -V(x)\psi(x), \tag{17}$$

this method approximates ψ by $\exp(\int^x \pm\sqrt{-V(x)}dt)$. There is a significant amount of theory surrounding this approximation, but the basic idea is that if taking two derivatives has the affect of multiplying by a function, then taking one derivative should be like multiplying by the square root of that function. If $V(x) < 0$ then the root is real, and we will have exponential growth or decay, depending on the choice of $\pm$. If $V(x) > 0$ then the root is imaginary, and we will have oscillations of 'instantaneous frequency' $\sqrt{V(x)}$. Points where $V(x) = 0$ are called *turning points*.

We know our Associated Legendre functions satisfy the differential equation (14). This is not quite in the right form, so we construct the functions $\{\psi_n^m(\phi) = \sqrt{\sin\phi}\tilde{P}_n^m(\cos\phi)\}$. Since we multiplied by 'half' of the measure $\sin\phi$, $\{\psi_n^m(\phi)\}$ is a basis with measure $d\phi$ for fixed m. From (14) we can deduce the differential equation

$$\frac{d^2}{d\phi^2}\psi_n^m(\phi) = -\left[(n+1/2)^2 - \frac{m^2-1/4}{\sin^2\phi}\right]\psi_n^m(\phi)\,, \tag{18}$$

which is a nice Schrödinger equation, to which we can apply the Quasi-Classical approximation. This approximation yields 'instantaneous frequency'

$$\nu_n^m(\phi) = \sqrt{(n+1/2)^2 - \frac{m^2-1/4}{\sin^2\phi}}\,, \tag{19}$$

valid when the argument of the root is positive, and the approximation $\psi_n^m(\phi) \approx \exp\left(i\int^\phi \nu_n^m(t)dt\right)$.

These estimates tell us that at the edges of the interval (when $V(x) < 0$) our functions decay rapidly and smoothly. As we move toward the center of the interval, they have instantaneous frequency increasing and concave down (see Figures 1 and 2). Globally these functions are complicated, but locally they look very much like trigonometric functions. These estimates can also be used to model the entire matrix $(\psi_n^m(\phi_j))_{n,j}$ for m fixed (see Figure 3).

In (18) we saw that $\psi_n^m(\phi) = \sqrt{\sin\phi}\tilde{P}_n^m(\cos\phi)$ satisfied a Schrödinger equation with

$$V(\phi) = (n+1/2)^2 - \frac{m^2-1/4}{\sin^2\phi}. \tag{20}$$

Instead we could consider

$$V(x) = (n+1/2)^2 - \frac{m^2-1/4}{x^2} \tag{21}$$

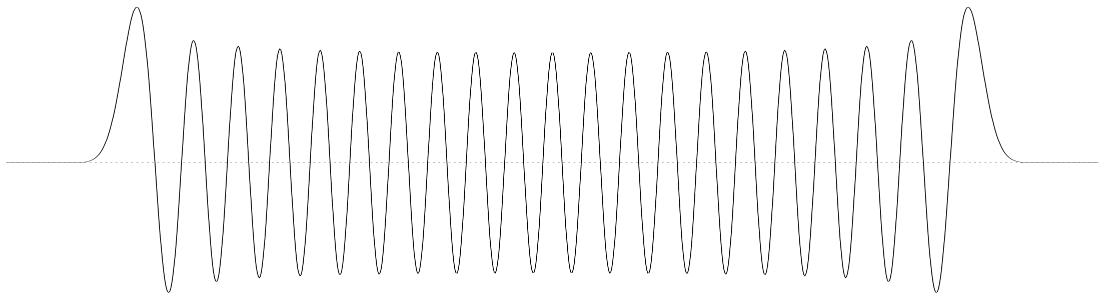

Figure 1: $\psi_{60}^{20}(\phi) = \sqrt{\sin\phi}\tilde{P}_{60}^{20}(\cos\phi)$ on $[0,\pi]$.

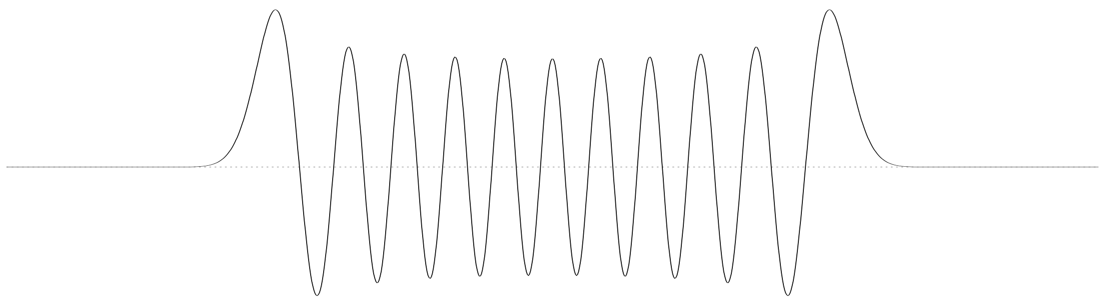

Figure 2: $\psi_{60}^{40}(\phi) = \sqrt{\sin\phi}\tilde{P}_{60}^{40}(\cos\phi)$ on $[0,\pi]$.

and ψ would be the Bessel function $\sqrt{x}J_m((n+1/2)x)$. Qualitatively, these behave the same as the Associated Legendre functions on $[0,\pi/2]$. See Figure 4 for graphs comparing the behavior of an Associated Legendre function and the corresponding Bessel function. Classical theorems relating the zeroes of Bessel functions and Associated Legendre functions in the limit as $n \to \infty$ may be found in [8, p.127].

We note that for a fixed interval in ϕ that avoids 0 and π, and for fixed m, as $n \to \infty$, ν_n^m tends toward the constant function $n+1/2$. Thus ψ_n^m tends toward a cosine. Many other asymptotic relationships may be found in [1, Section 8].

3 Computing with Spherical Harmonics

In this section we outline what is needed in order to do computations with Spherical Harmonics. In Section 3.1 we discuss how to construct these functions. In Section 3.2 we discuss the ways to discretize the sphere and their implications. In Section 3.3 we synthesize these concepts and produce the algorithm for expanding in or evaluating Spherical Harmonic series.

3.1 Generating the $\tilde{P}_n^m$'s

In this section we present a methods for constructing the Associated Legendre Functions. There are many other representations for the Associated Legendre functions and formulas involving them [1] [4]. The only method we have found satisfactory for numerical work is the normalized recurrence (29) below.

The Associated Legendre functions may be constructed as an amplitude times a Jacobi polynomial,

$$\tilde{P}_n^m(\cos\phi) = (\sin\phi)^m \tilde{P}_{n-m}^{(m,m)}(\cos\phi) \quad \text{for } m \geq 0\,, \tag{22}$$

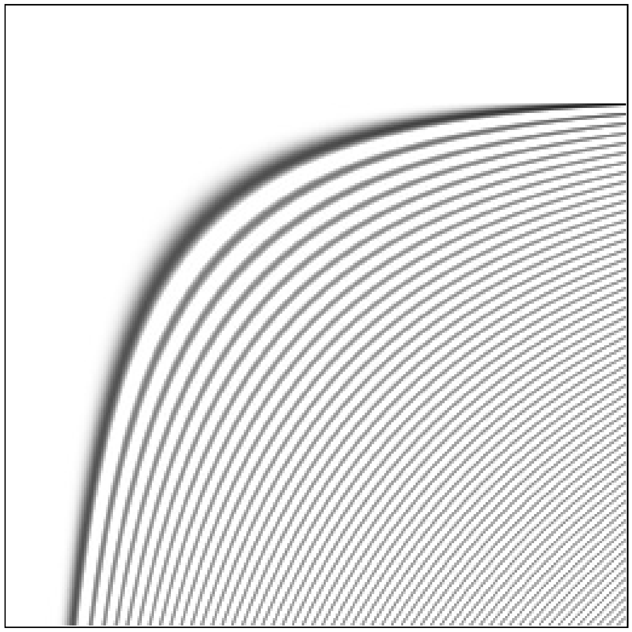

Figure 3: The positive part of $\psi_n^{40}(\phi) = \sqrt{\sin\phi}\tilde{P}_n^{40}(\cos\phi)$ for n in $[0, 256]$ and ϕ in $[0, \pi/2]$. The n axis points down, and the ϕ axis to the right.

and for $m < 0$ using $\tilde{P}_n^m(\cos\phi) = \tilde{P}_n^{-m}(\cos\phi)$. This formula is immensely useful, but we must be cautious in its application. If we write out the Jacobi polynomial as a polynomial, it will have huge coefficients. If we try to evaluate this representation at a point we are attempting to subtract huge numbers to yield small numbers, which will lead to (catastrophic) loss of precision. The problem is that the monomials x^k are nearly colinear.

Another aspect to note is that while $\sqrt{\sin\phi}\tilde{P}_n^m(\cos\phi)$ is of nearly constant amplitude (Figures 1 and 2), $(\sin\phi)^m$ can be very small. This means the Jacobi polynomials are very large in this region, although size 1 near $\phi = \pi/2$. Algorithms which ignore this conditioning problem in the polynomial part of the Associated Legendre functions (such as [3]) will be unstable.

The Jacobi Polynomials can be generated by a three term recurrence relation. The traditional recurrence is [8, p.68,71]

$$P_{-1}^{(m,m)}(x) = 0\,, \tag{23}$$

$$P_0^{(m,m)}(x) = 1\,, \quad \text{and} \tag{24}$$

$$\begin{aligned} k(k+2m)(k+m-1)P_k^{(m,m)}(x) =& (2k+2m-1)(k+m)(k+m-1)xP_{k-1}^{(m,m)}(x) \\ &- (k+m-1)^2(k+m)P_{k-2}^{(m,m)}(x). \end{aligned} \tag{25}$$

To normalize to obtain $\tilde{P}_k^{(m,m)}$ we must divide by

$$\sqrt{\frac{2^{2m+1}\Gamma^2(k+m+1)}{(2k+2m+1)\Gamma(k+1)\Gamma(k+2m+1)}}. \tag{26}$$

The Γ functions (factorials) in (26) can cause overflows for even moderate m and k.

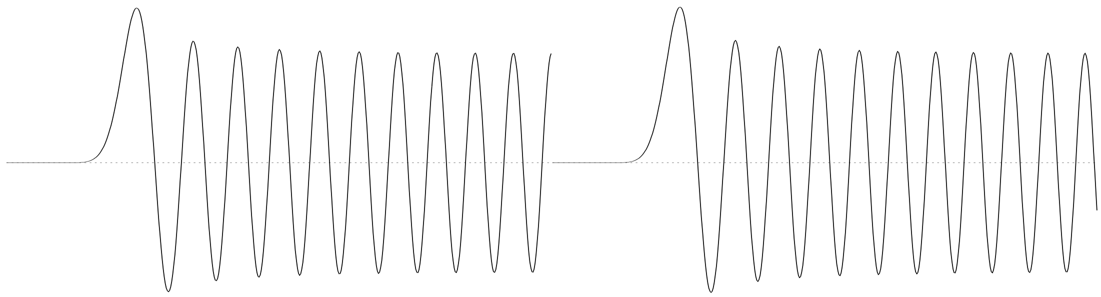

Figure 4: $\sqrt{\sin\phi}\tilde{P}_{60}^{20}(\cos\phi)$ and $\sqrt{r}J_{20}(60r)$ on $[0,\pi/2]$

Using the recurrence relation (25) and normalization (26) above we can construct $\tilde{P}_k^{(m,m)}$ with the recurrence

$$\tilde{P}_{-1}^{(m,m)}(x) = 0\,, \tag{27}$$

$$\tilde{P}_0^{(m,m)}(x) = \sqrt{\frac{\Gamma(2m+2)}{2}}\frac{1}{2^m\Gamma(m+1)} = \frac{1}{\sqrt{2}}\prod_{j=1}^{m}\sqrt{1+\frac{1}{2j}}\,, \quad \text{and} \tag{28}$$

$$\begin{aligned}\tilde{P}_k^{(m,m)}(x) =& 2x\tilde{P}_{k-1}^{(m,m)}(x)\left(1+\frac{m-1/2}{k}\right)^{1/2}\left(1-\frac{m-1/2}{k+2m}\right)^{1/2} \\ &- \tilde{P}_{k-2}^{(m,m)}(x)\left(1+\frac{4}{2k+2m-3}\right)^{1/2}\left(1-\frac{1}{k}\right)^{1/2}\left(1-\frac{1}{k+2m}\right)^{1/2}.\end{aligned} \tag{29}$$

For large m this recurrence is poorly conditioned, especially when $m \le n \le 2m$. The solution we want corresponds to the larger singular value, however, so the conditioning acts in our favor. The recurrence is also prone to underflows when m is a few thousand. We can fix this by using scientific notation, $\sqrt{\sin\phi}\tilde{P}_n^m(\cos\phi) = A2^B$, $1/2 < A \le 1$, $B \in \mathbf{Z}$, and storing A and B separately.

Remark 1. *The formulas above give $\tilde{P}_n^m$ with a particular sign convention, where $\tilde{P}_n^m(\cos\phi)$ is positive for small ϕ. For translation into other conventions, see [1].*

3.2 Discretizing the Problem

One would like to have an uniform discretization for the sphere, with all portions equally represented. From such an uniform discretization we could construct a platonic solid. It is known, however, that there are only a few platonic solids, and the largest number of faces is 20 (icosohedron) and largest number of vertices is 20 (dodecahedron). If we want to discretize the sphere with many points, we cannot do it uniformly.

Instead we set the goal of using the fewest points to resolve the Spherical Harmonics up to some degree. Since the Spherical Harmonics themselves are "fair" and "uniform", this gives a good representation for functions on the sphere.

3.2.1 Principles of Quadrature

A quadrature is a rule for converting an integral into a sum. Quadratures may be exact or approximate. Here we restrict our attention to rules that are exact for some given class of functions. We consider integrals on $[-1,1]$ and find sample points $\{x_j\}_1^N$ and weights $\{w_j\}_1^N$ such that

$$\int_{-1}^{1} f(x)dx = \sum_{j=1}^{N} w_j f(x_j)\,. \tag{30}$$

First suppose $f(x)$ is a line. Given any two sample values of f, we can determine this line, and thus the value of the integral. The line can be determined as a Lagrange interpolating polynomial by

$$\frac{(x-x_2)}{(x_1-x_2)}f(x_1) + \frac{(x-x_1)}{(x_2-x_1)}f(x_2) \tag{31}$$

and then the integral by

$$\int_{-1}^{1} \frac{(x-x_2)}{(x_1-x_2)}dx f(x_1) + \int_{-1}^{1} \frac{(x-x_1)}{(x_2-x_1)}dx f(x_2) = w_1 f(x_1) + w_2 f(x_2)\,. \tag{32}$$

Any two sample points will work in theory, but if x_1 and x_2 are both close to 1, for example, then the quadrature will be numerically unstable since w_1 and w_2 will be large and of opposite sign.

The same procedure works for higher degree polynomials. Given N sample points we can construct a Lagrange interpolating polynomial as in (31) and then a quadrature as in (32). If f has degree less than N we have an exact quadrature rule.

The *Gaussian quadrature* is a trick that gives us an exact quadrature for polynomials of degree less than $2N$ using only N points. The Gaussian quadrature is built from the Legendre Polynomials P_n (coincidentally $\tilde{P}_n = \tilde{P}_n^0$). The Legendre polynomials are constructed by a Gramm-Schmidt orthonormalization procedure on $[-1,1]$ starting with the sequence $1, x, x^2, \ldots$. The effect of this is that $\tilde{P}_n$ is of degree n and is orthogonal to all polynomials p_k with degree $k < n$, i.e.

$$\int_{-1}^{1} \tilde{P}_n(x) p_k(x) dx = 0\,. \tag{33}$$

If f is of degree $2N-1$ we can use polynomial division to write $f(x) = P_N(x)p_{N-1}(x) + r_{N-1}(x)$ where the subscript denotes degree. Because of (33) we know $\int f = \int r$. Since r_{N-1} is of degree $N-1$, its integral may be done using (any) N quadrature points. Our challenge is to pick the points so that $\int P_N p_{N-1} = 0$. But $\tilde{P}_N(x)$ has exactly N zeroes in $[-1,1]$, so we pick these as our sample points.

Remark 2. *Transformed to the sphere using $x_j = \cos\phi_j$ we note that ϕ_j are nearly equispaced, gathering slightly near the poles. One can use true equispaced nodes in ϕ, and these correspond to Chebychev (spelling varies) nodes in x. The Chebychev nodes have several nice properties, but require twice as many points as the Gaussian nodes.*

The principle for quadratures of trigonometric polynomials is slightly different. We pick N sample points, equally spaced: $\{j2\pi/N : 0 \le j < N\}$. For $f = \sum_{|k|<N} a_k e^{ik\theta}$, we have $\int f = a_0$ and

$$\frac{1}{N}\sum_j \sum_{|k|<N} a_k e^{ikj2\pi/N} = \frac{1}{N}\sum_{|k|<N} a_k \sum_j e^{ikj2\pi/N}\,. \tag{34}$$

If $0 < |k| < N$ then the set $\{e^{ikj2\pi/N}\}$ of points in the complex plane is symmetric about 0, and so must sum to 0. When $k = 0$ (or lN), $e^{i0j2\pi/N} = 1$, so the sum is $(1/N)a_0 N = a_0$. Thus for trigonometric polynomials, the quadrature weights are all 1 (really $1/N$) and the quadrature points are an equispaced sampling.

3.2.2 Spherical Harmonics

In order to expand in Spherical Harmonics (Section 3.3) we will need to take inner products of Spherical Harmonics. Since these inner products are integrals, the key point of our discretization is that it gives a quadrature valid for the product of two Spherical Harmonics. We need to compute the integrals

$$\int_0^\pi \int_0^{2\pi} \tilde{P}_{n'}^{m'}(\cos\phi)\frac{e^{im'\theta}}{\sqrt{2\pi}}\tilde{P}_n^m(\cos\phi)\frac{e^{-im\theta}}{\sqrt{2\pi}}\sin\phi d\phi d\theta\,. \tag{35}$$

In θ the Spherical Harmonics are just exponentials, and so is their product, so we use the rule for trigonometric polynomials and evaluate using $2N$ points equispaced in θ. If $m \neq m'$ we get 0, otherwise we are then left with the integral

$$\int_0^\pi \tilde{P}_{n'}^m(\cos\phi)\tilde{P}_n^m(\cos\phi)\sin\phi d\phi\,. \tag{36}$$

For $n, n' < N$, $\tilde{P}_n^m(\cos\phi)\tilde{P}_{n'}^m(\cos\phi) = P(\cos\phi)$ is a polynomial in $\cos\phi$ of degree at most $2N-2$. One can verify this by using their explicit construction in terms of Jacobi polynomials (Section 3.1) and noting $\sin^{2m}\phi = (1-\cos^2\phi)^m$. Our integral then reduces to

$$\int_{-1}^1 P(x)dx\,. \tag{37}$$

Using Gaussian nodes and weights (in x) we only need N points to obtain an exact quadrature formula.

One could instead use equally spaces points in ϕ. The advantage of equally spaced points is their simplicity. Sometimes data is provided on an equispaced grid, or a data processing or display program assumes it. The disadvantage is that they are less efficient. Equispaced points in ϕ are equivalent to Chebychev nodes in x, so we will need $2N$ points. One can translate the usual results for Chebychev weights to our case, or follow the derivation in Driscoll and Healy [3] to obtain

$$w_{2N}(j) = \frac{\sqrt{2}}{N}\sum_{l=0}^{N-1}\frac{1}{2l+1}\sin((2l+1)\phi_j)\,. \tag{38}$$

3.2.3 Thinning the Grid

As we can see in Figure 1, the Associated Legendre Functions become very small near the ends of the interval, which correspond to the poles. As we increase m, this effect becomes stronger, as we can see in Figure 2. Although the function does not become zero, at some point it becomes small enough that we can neglect it. Although it is difficult to give an analytic expression for when this occurs, it is easy to to track the turning point (inflection point), where the rapid decay begins. From the differential equation (18), we see that the inflection point occurs when $(n+1/2)^2 - (m^2 - 1/4)(\sin^2\phi)$, so $\sin\phi \approx m/n$.

When computing the integral (36) we can therefore discard some quadrature points, essentially those that satisfy $\sin\phi_j < m/N$. As m increases the number of discarded points increases. For a given ϕ, we therefore only need to compute the integral in θ for $m < m_\phi \leq N$ and so can use $2m_\phi$ points instead of $2N$. This effect allows us to "thin" the grid by discarding some points near the poles. After this thinning, each point corresponds to an approximately equal area. In practice, as we move to the pole, thinning is usually only done when the number of points in θ can be reduced by a factor of 2.

3.2.4 Other Considerations

One nice feature of Fourier series is that it is also a discrete basis: using N values, we expand into N basis functions and so lose nothing. For Spherical Harmonics, however, we have to assume something about the function f we wish to expand. We needed $2N \times N = 2N^2$ sample points to expand to degree N, yielding only $2N \times N/2 = N^2$ Spherical Harmonics. We thus cannot hope to capture an arbitrary function on $2N^2$ points.

If we try to represent a function using too few points, we will have aliasing. With Fourier series, the function $\exp(i(N+k)\theta)$ and the function $\exp(ik\theta)$ are indistinguishable on the grid $\{j2\pi/N : 0 \leq j < N\}$, since

$$\exp(i(N+k)j2\pi/N) = \exp(ij2\pi)\exp(ikj2\pi/N) = \exp(ikj2\pi/N)\,. \tag{39}$$

This is periodic aliasing since frequencies k and $N+k$ are confused. Another type of aliasing is reflective aliasing, where frequencies $N+k$ and $N-k$ are confused. This occurs, for example, with the cosine basis $\{\cos(kx)\}_{k=0}^{N-1}$ on $x \in [0,\pi]$, with the discretization $\{j\pi/N : 0 \leq j < N\}$. By trigonometric formulas, we have

$$\begin{aligned} \cos((N+k)j\pi/N) &= \cos((2N+(k-N)j\pi/N) \\ &= \cos(2j\pi)\cos((k-N)j\pi/N) - \sin(2j\pi)\sin((k-N)j\pi/N) \\ &= \cos((N-k)j\pi/N)\,. \end{aligned} \tag{40}$$

The aliasing of spherical harmonics follows the Fourier rules in θ, and is similar to the reflective aliasing in ϕ. It is not, however, true aliasing, because the undersampled function does not appear to be a single lower frequency function, but rather as a linear combination of lower frequency functions. The inner product (36) will integrate correctly as long as $n+n' < 2N$. If $n = N+k$, then the inner products for $n' \geq N-k$ will be incorrect. When $n+n'$ is odd, the odd symmetry will make the inner product correctly zero, but in general the computed inner products will be nonsense.

3.3 Expansions and Evaluations

There are three basic things one does with Spherical Harmonics: expand into them, evaluate some operator in them, and evaluate a series of them. The operator step depends on the problem you are dealing with, so we will not consider it, beyond pointing out the diagonalization property in Section 1.4. In this section we go through the algorithms for expanding in or evaluating a Spherical Harmonic series. First we give the basic algorithm and then several improvements. We treat the expansion problem in detail, and then sketch the (simpler) evaluation problem.

3.3.1 The Basic Expansion Algorithm

In the expansion problem for Spherical Harmonics we are given the sampled values $f(\phi_j,\theta_k)$ of a function of the form

$$f(\phi,\theta) = \sum_{\mathcal{I}_N} \alpha_n^m \tilde{P}_n^m(\cos\phi)\frac{e^{im\theta}}{\sqrt{2\pi}}\,. \tag{41}$$

The samples are on a grid in $\phi \times \theta$ with $2N$ equispaced points in θ and N Gaussian nodes in $\cos\phi$:

$$\left\{(\phi_j,\theta_k) = \left(\cos^{-1}(g_j^N), 2\pi\frac{k}{2N}\right) \;:\; j,k \in \mathbf{Z}; 0 \leq j < N, 0 \leq k < 2N\right\}\,. \tag{42}$$

The coefficient indices are assumed to be in the set

$$\mathcal{I}_N = \{(m,n) \in \mathbf{Z} \times \mathbf{Z} \;:\; 0 \leq n < N, -n \leq m \leq n\}\,. \tag{43}$$

From the $\mathcal{O}(N^2)$ input values $f(\phi_j,\theta_k)$ we wish to compute the $\mathcal{O}(N^2)$ output values α_n^m in an efficient, numerically stable way.

For this expansion problem we have learned first that these Spherical Harmonics are an orthonormal set, and so we should compute α_n^m by

$$\alpha_n^m = \int_0^\pi \int_0^{2\pi} f(\phi,\theta)\tilde{P}_n^m(\cos\phi)\frac{e^{-im\theta}}{\sqrt{2\pi}}\sin\phi d\phi d\theta \tag{44}$$

$$= \int_0^\pi \left(\int_0^{2\pi} f(\phi,\theta)\frac{e^{-im\theta}}{\sqrt{2\pi}}d\theta\right)\tilde{P}_n^m(\cos\phi)\sin\phi d\phi. \tag{45}$$

With our assumption (41) on the form of f, we only need be able to compute inner products of P_n^m's (m fixed).

Our next task is to discretize the integrals. The inner integral is a Fourier series expansion. With our assumption (41) on the true nature of f, the sampling theorem says

$$f(\phi,\hat{m}) = \int_0^{2\pi} f(\phi,\theta)\frac{e^{-im\theta}}{\sqrt{2\pi}}d\theta = \sum_{k=0}^{2N-1} f(\phi,\theta_k)e^{-im\theta_k}\frac{\sqrt{2\pi}}{2N}\,. \tag{46}$$

For fixed m and $n, n' < N$, $\tilde{P}_n^m(\cos\phi)\tilde{P}_{n'}^m(\cos\phi) = (\sin\phi)^{2m}\tilde{P}_{n-m}^{(m,m)}(\cos\phi)\tilde{P}_{n'-m}^{(m,m)}(\cos\phi) = P(\cos\phi)$ is a polynomial in $\cos\phi$ of degree at most $2N-2$. Thus using Gaussian nodes and weights in $\cos\phi$, we can perform this integral exactly using N points (see Section 3.2.1), to obtain

$$\alpha_n^m = \int_0^\pi f(\phi,\hat{m})\tilde{P}_n^m(\cos\phi)\sin\phi d\phi = \sum_{j=0}^{N-1} f(\phi_j,\hat{m})\tilde{P}_n^m(\cos\phi_j)\sin\phi_j w_N(j)\,. \tag{47}$$

Remark 3. *Instead of using the Gaussian nodes in (42) and weights in (47), one can use equispaced nodes and their corresponding weights. One must use $2N$ points instead of N, however. Otherwise everything works the same.*

If we perform the double sum (integral) for each (m,n), we do $\mathcal{O}(N^4)$ operations. As we noted in the form of (45), however, the variables separate. Each of the inner integrals in (45) takes $\mathcal{O}(N^2)$ to compute, and there are $\mathcal{O}(N)$ or them, for cost $\mathcal{O}(N^3)$. To compute the integral (47) costs $\mathcal{O}(N)$ operations, and we must do this for $\mathcal{O}(N^2)$ values for (m,n), so the cost for this step is also $\mathcal{O}(N^3)$. The total cost is thus also $\mathcal{O}(N^3)$.

3.3.2 Improvement: Precompute the $\tilde{P}_n^m$s

The first step in making the basic algorithm efficient is to precompute as much as possible. If you are only going to do one transform this will not help, but if you will do several transforms of the same size, this will improve performance significantly. The things to precompute are the Gaussian nodes $\cos^{-1}(g_j^N)$ and weights w_j^N, the values of $\tilde{P}_n^m(g_j^N)$, and possibly the measure $\sin(\cos^{-1}(g_j^N))$ (this could be incorporated into w_j^N).

For stability, one might wish to incorporate 'half' the measure onto the Associated Legendre functions and store $\sqrt{\sin(\cos^{-1}(g_j^N))}\tilde{P}_n^m(g_j^N)$. This factor keeps the amplitude nearly constant (see Figures 1 and 2).

3.3.3 Improvement: Use the FFT in θ

The integral in (46) is a Fourier series expansion. The sum can be computed for each ϕ in $\mathcal{O}(N\log N)$ time with the Fast Fourier Transform (FFT), to yield $f(\phi,\hat{m})$. The cost for this step is reduced from $\mathcal{O}(N^3)$ to $\mathcal{O}(N^2\log N)$, although the overall order of the algorithm is unchanged.

3.3.4 Improvement: Use the Symmetry in ϕ

One property of $\tilde{P}_n^m(\cos\phi)$ that is only apparent from its explicit construction in Section 3.1 is that it is either even or odd across $\phi = \pi/2$ as $n - m$ is even or odd. We can perform the sum in (47) separately for the even and odd components on $\phi \in (0, \pi/2)$ if we first perform simple reflections.

We construct the even portion of the function

$$f_e(\phi_j, \hat{m}) = f(\phi_j, \hat{m}) + f(\phi_{N-1-j}, \hat{m}) \tag{48}$$

for $0 \leq j < N/2$ and compute

$$\alpha_n^m = \sum_{j=0}^{N/2-1} f_e(\phi_j, \hat{m}) \tilde{P}_n^m(\cos\phi_j) \sin\phi_j w_N(j) \tag{49}$$

for $m - n$ even. The odd portion is similar.

For fixed m, the computation of (47) took $(N - m)N$ operations, but using these symmetries it takes $2(N/2) + 2((N - m)/2) = N + (N - m)N/2$ operations, which is about half as many.

3.3.5 Fast Transforms

The algorithm so far has computational complexity $\mathcal{O}(N^3)$. For large N this becomes unwieldy and there is a call for a faster algorithm, analogous to the FFT. In [6] there is an algorithm with theoretical time $\mathcal{O}(N^2(\log N)^2)$ and a more practical algorithm with time $\mathcal{O}(N^{5/2})$. The $\mathcal{O}(N^{5/2})$ algorithm breaks even with the algorithm described above at about $N = 100$, and cuts computational time by about three at $N = 512$. For problems smaller than $N = 512$ it should not be considered, but for problems much larger that $N = 512$ it can give significant cost savings. Software is available; search for "libftsh".

3.3.6 The Evaluation Problem

We are given a set of coefficients $\{\alpha_n^m\}_{(m,n)\in\mathcal{I}_N}$ and we form the sum

$$f(\phi, \theta) = \sum_{\mathcal{I}_N} \alpha_n^m \tilde{P}_n^m(\cos\phi) e^{im\theta}/\sqrt{2\pi}\,. \tag{50}$$

This sum is evaluated on a grid in $\phi \times \theta$ with $2N$ equispaced points in θ and N Gaussian nodes in $\cos\phi$. The evaluation problem for Spherical Harmonics is to compute the $\mathcal{O}(N^2)$ output values $f(\phi_j, \theta_k)$ from the $\mathcal{O}(N^2)$ input values α_n^m in an efficient, numerically stable way.

This sum could be evaluated directly, but the improvements in Sections 3.3.1 to 3.3.5 also apply.

References

[1] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions*, volume 55 of *Applied Math Series*. National Bureau of Standards, 1964.

[2] C. M. Bender and S. A. Orszag. *Advanced mathematical methods for scientists and engineers*. International series in pure and applied Mathematics. McGraw–Hill, New York, 1978.

[3] James R. Driscoll and Dennis M. Healy, Jr. Computing Fourier transforms and convolutions on the 2-sphere. *Adv. in Appl. Math.*, 15(2):202–250, 1994. `doi:10.1006/aama.1994.1008`.

[4] E. W. Hobson. *The Theory of Spherical and Ellipsoidal Harmonics*. Chelsea, New York, 1931.

[5] L. D. Landau and E. M. Lifshitz. *Quantum Mechanics (Non–relativistic Theory)*. Pergamon Press, New York, 3rd edition, 1977.

[6] Martin J. Mohlenkamp. A fast transform for spherical harmonics. *J. Fourier Anal. Appl.*, 5(2-3):159–184, 1999. doi:10.1007/BF01261607.

[7] Elias Stein and Guido Weiss. *Fourier Analysis on Euclidean Spaces*. Princeton Univ. Press, Princeton NJ, 1971.

[8] G. Szegö. *Orthogonal Polynomials*. AMS, Providence, RI, 4th edition, 1975.