B

Spherical Harmonics

SPHERICAL harmonics are a frequency-space basis for representing functions defined over the sphere. They are the spherical analogue of the 1D Fourier series. Spherical harmonics arise in many physical problems ranging from the computation of atomic electron configurations to the representation of gravitational and magnetic fields of planetary bodies. They also appear in the solutions of the Schrödinger equation in spherical coordinates. Spherical harmonics are therefore often covered in textbooks from these fields [MacRobert and Sneddon, 1967; Tinkham, 2003].

Spherical harmonics also have direct applicability in computer graphics. Light transport involves many quantities defined over the spherical and hemispherical domains, making spherical harmonics a natural basis for representing these functions. Early applications of spherical harmonics to computer graphics include the work by Cabral et al. [1987] and Sillion et al. [1991]. More recently, several in-depth introductions have appeared in the graphics literature [Ramamoorthi, 2002; Green, 2003; Wyman, 2004; Sloan, 2008].

In this appendix, we briefly review the spherical harmonics as they relate to computer graphics. We define the basis and examine several important properties arising from this definition.

B.1 Definition

A *harmonic* is a function that satisfies Laplace's equation:

$$\nabla^2 f = 0. \tag{B.1}$$

As their name suggests, the spherical harmonics are an infinite set of harmonic functions defined on the sphere. They arise from solving the angular portion of Laplace's equation in spherical coordinates using separation of variables. The spherical harmonic basis functions derived in this fashion take on complex values, but a complementary, strictly real-valued, set of harmonics can also be defined. Since in computer graphics we typically only encounter real-valued functions, we restrict our discussion to the real-valued basis.

If we represent a direction vector $\vec{\omega}$ using the standard spherical parameterization,

$$\vec{\omega} = \left(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta\right), \tag{B.2}$$

then the real spherical harmonic basis functions are defined as:

$$y_l^m(\theta,\phi) = \begin{cases} \sqrt{2}K_l^m \cos(m\phi)P_l^m(\cos\theta) & \text{if } m > 0, \\ K_l^0 P_l^0(\cos\theta) & \text{if } m = 0, \\ \sqrt{2}K_l^m \sin(-m\phi)P_l^{-m}(\cos\theta) & \text{if } m < 0. \end{cases} \tag{B.3}$$

where K_l^m are the normalization constants

$$K_l^m = \sqrt{\frac{(2l+1)}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}}, \tag{B.4}$$

and P_l^m are the associated Legendre polynomials. There are many ways to define the associated Legendre polynomials but the most numerically robust way to evaluate them is using a set of

recurrence relations [Press et al., 1992][1]:

$$P_0^0(z) = 1, \tag{B.5}$$

$$P_m^m(z) = (2m-1)!!(1-z^2)^{m/2}, \tag{B.6}$$

$$P_{m+1}^m(z) = z(2m+1)P_m^m(z), \tag{B.7}$$

$$P_l^m(z) = \frac{z(2l-1)}{l-m}P_{l-1}^m(z) - \frac{(l+m-1)}{l-m}P_{l-2}^m(z). \tag{B.8}$$

The basis functions are indexed according to two integer constants, the *order*, l, and the *degree*, m[2]. These satisfy the constraint that $l \in \mathbb{N}$ and $-l \leq m \leq l$; thus, there are $2l+1$ basis functions of order l.

The order l determines the frequency of the basis functions over the sphere. The spherical harmonics may be written either as trigonometric functions of the spherical coordinates θ and ϕ as above, or alternately as polynomials of the cartesian coordinates x, y, and z. Using the cartesian representation, each y_l^m for a fixed l corresponds to a polynomial of maximum order l in x, y, and z.

[1] We omit the Condon-Shortley phase factor [Condon and Shortley, 1951] of $(-1)^m$ which is sometimes includes in the definition of P_l^m or y_l^m since this simplifies our notation.

[2] The order l is also sometimes refered to as the *band* index, and in quantum mechanics, l and m are refered to as "quantum numbers" and the spherical harmonics "states."

The first few spherical harmonics, in both spherical and cartesian coordinates, expand to:

		Spherical	**Cartesian**
$l=0$	$y_0^0(\theta,\phi) =$	$\sqrt{\frac{1}{4\pi}}$	$\sqrt{\frac{1}{4\pi}}$,
$l=1$	$y_1^{-1}(\theta,\phi) =$	$\sqrt{\frac{3}{4\pi}}\sin\phi\sin\theta$	$\sqrt{\frac{3}{4\pi}}x$,
	$y_1^0(\theta,\phi) =$	$\sqrt{\frac{3}{4\pi}}\cos\theta$	$\sqrt{\frac{3}{4\pi}}z$,
	$y_1^1(\theta,\phi) =$	$\sqrt{\frac{3}{4\pi}}\cos\phi\sin\theta$	$\sqrt{\frac{3}{4\pi}}y$,
$l=2$	$y_2^{-2}(\theta,\phi) =$	$\sqrt{\frac{15}{4\pi}}\sin\phi\cos\phi\sin^2\theta$	$\sqrt{\frac{15}{4\pi}}xy$,
	$y_2^{-1}(\theta,\phi) =$	$\sqrt{\frac{15}{4\pi}}\sin\phi\sin\theta\cos\theta$	$\sqrt{\frac{15}{4\pi}}yz$,
	$y_2^0(\theta,\phi) =$	$\sqrt{\frac{5}{16\pi}}(3\cos^2\theta-1)$	$\sqrt{\frac{5}{16\pi}}(3z^2-1)$,
	$y_2^1(\theta,\phi) =$	$\sqrt{\frac{15}{4\pi}}\cos\phi\sin\theta\cos\theta$	$\sqrt{\frac{15}{8\pi}}xz$,
	$y_2^2(\theta,\phi) =$	$\sqrt{\frac{15}{16\pi}}(\cos^2\phi-\sin^2\phi)\sin^2\theta$	$\sqrt{\frac{15}{32\pi}}(x^2-y^2)$.

We illustrate the first basis functions in Figure B.1.

B.2 Projection and Expansion

The spherical harmonics define a complete basis over the sphere. Thus, any real-valued spherical function f may be *expanded* as a linear combination of the basis functions

$$f(\vec{\omega}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} y_l^m(\vec{\omega})\, f_l^m, \tag{B.9}$$

where the coefficients f_l^m are computed by *projecting* f onto each basis function y_l^m

$$f_l^m = \int_{\Omega_{4\pi}} y_l^m(\vec{\omega}) f(\vec{\omega}) d\vec{\omega}. \tag{B.10}$$

Just as with the Fourier series, this expansion is exact as long as l goes to infinity; however, this requires an infinite number of coefficients. By limiting the number of bands to $l = n-1$ we retain

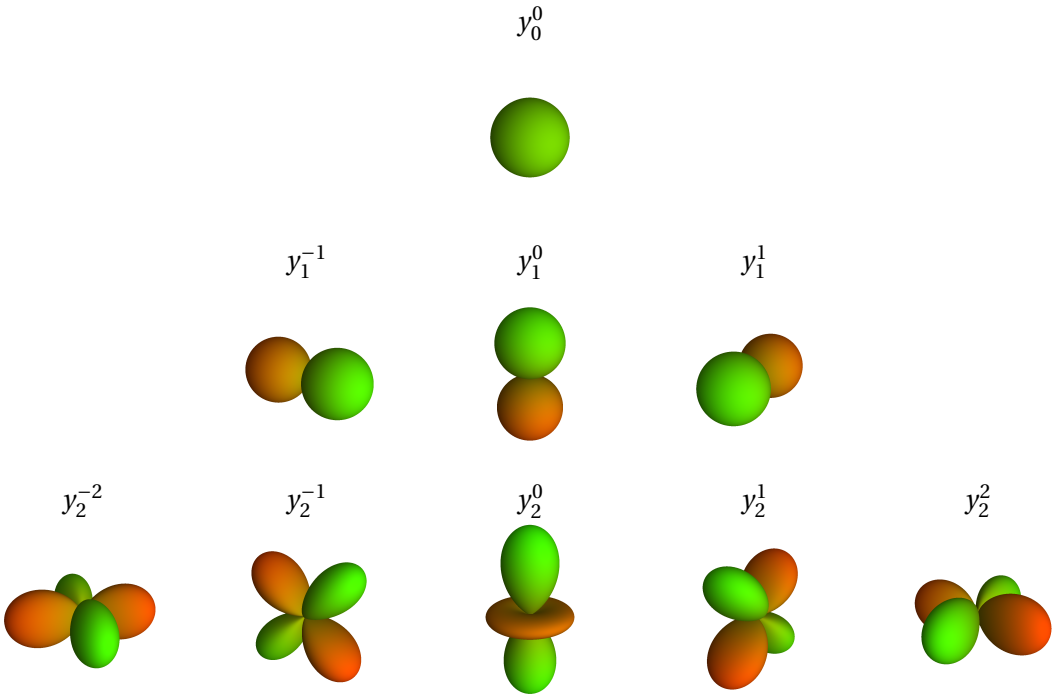

Figure B.1: Plots of the real-valued spherical harmonic basis functions. Green indicates positive values and red indicates negative values.

only the frequencies of the function up to some threshold, obtaining an n^{th} order band-limited approximation $\tilde{f}$ of the original function f:

$$\tilde{f}(\vec{\omega}) = \sum_{l=0}^{n-1} \sum_{m=-l}^{l} y_l^m(\vec{\omega})\, f_l^m. \tag{B.11}$$

Low-frequency functions can be well approximated using only a few bands, and as the number of coefficients increases, higher frequency signals can be approximated more accurately.

It is often convenient to reformulate the indexing scheme to use a single parameter $i = l(l+1)+m$. With this convension it is easy to see that an n^{th} order approximation can be reconstructed using n^2 coefficients,

$$\tilde{f}(\vec{\omega}) = \sum_{i=0}^{n^2-1} y_i(\vec{\omega})\, f_i. \tag{B.12}$$

B.3 Properties

The spherical harmonic functions have many basic properties that make them particularly convenient for use in computer graphics.

1. **Convolution.** Since the spherical harmonic basis is effectively a Fourier domain basis defined over the sphere, it inherits a similar frequency space convolution property. If $h(z)$ is a circularly symmetric kernel, then the *convolution* $h \star f$ is equivalent to weighted *multiplication* in the SH domain

$$(h \star f)_l^m = \sqrt{\frac{4\pi}{2l+1}} h_l^0 f_l^m. \tag{B.13}$$

   The convolution property allows for efficient computation of prefiltered environment maps and irradiance environment maps [Ramamoorthi and Hanrahan, 2001].

2. **Orthonormality.** The spherical harmonics are orthogonal for different l and different m. This means that the inner product of any two distinct basis functions is zero. Furthermore, the normalization constant K_l^m ensures that the inner product of a basis function with itself is one. This can be expressed mathematically as

$$\int_{\Omega_{4\pi}} y_i(\vec{\omega})\, y_j(\vec{\omega})\, d\vec{\omega} = \delta_{ij}, \tag{B.14}$$

   where δ_{ij} is the Kronecker delta function.

   The efficient projection and expansion operations described above are made possible by the fact that the SH basis is orthonormal. Many other useful and efficient operations also result from this important property.

3. **Double Product Integral.** The orthonormality property provides for a very simple expression to compute the integrated product of two functions represented in the SH basis. The

integral product of two SH functions $\tilde{a}(\vec{\omega})$ and $\tilde{b}(\vec{\omega})$ can be expanded as

$$\int_{\Omega_{4\pi}} \tilde{a}(\vec{\omega})\,\tilde{b}(\vec{\omega})\,d\vec{\omega} = \int_{\Omega_{4\pi}} \left(\sum_i a_i\, y_i(\vec{\omega})\right)\left(\sum_j b_j\, y_j(\vec{\omega})\right) d\vec{\omega}, \tag{B.15}$$

$$= \sum_i \sum_j a_i\, b_j \underbrace{\int_{\Omega_{4\pi}} y_i(\vec{\omega})\, y_j(\vec{\omega})\, d\vec{\omega}}_{C_{ij}}, \tag{B.16}$$

where C_{ij} are called the **coupling coefficients**, which, due to the definition of orthonormality in Equation B.14, are simply $C_{ij} = \delta_{ij}$. This simple form for the coupling coefficients introduces significant sparsity in the expression above, leading to the simplification:

$$\int_{\Omega_{4\pi}} \tilde{a}(\vec{\omega})\,\tilde{b}(\vec{\omega})\,d\vec{\omega} = \sum_i \sum_j a_i\, b_j\, C_{ij}, \tag{B.17}$$

$$= \sum_i \sum_j a_i\, b_j\, \delta_{ij}, \tag{B.18}$$

$$= \sum_i a_i\, b_i. \tag{B.19}$$

This expression states that the integrated product of two SH functions is simply the dot product of their coefficient vectors. The double product integral is of particular interest in computer graphics since it means lighting can be computed very efficiently in the frequency domain. If both the lighting and the cosine-weighted BRDF are represented in the SH basis, then the lighting integral can be computed using a simple dot product. This property is exploited by many PRT techniques [Sloan et al., 2002; Kautz et al., 2002].

4. **Triple Product Integral.** In many applications, the product integral of not just two, but three SH functions is of particular interest. This product can be expanded as

$$\int_{\Omega_{4\pi}} \tilde{a}(\vec{\omega})\,\tilde{b}(\vec{\omega})\,\tilde{c}(\vec{\omega})\,d\vec{\omega} = \int_{\Omega_{4\pi}} \left(\sum_i a_i\, y_i(\vec{\omega})\right)\left(\sum_j b_j\, y_j(\vec{\omega})\right)\left(\sum_k c_k\, y_k(\vec{\omega})\right) d\vec{\omega}, \tag{B.20}$$

$$= \sum_i \sum_j \sum_k a_i\, b_j\, c_k \int_{\Omega_{4\pi}} y_i(\vec{\omega})\, y_j(\vec{\omega})\, y_k(\vec{\omega})\, d\vec{\omega}, \tag{B.21}$$

$$= \sum_i \sum_j \sum_k a_i\, b_j\, c_k\, C_{ijk}, \tag{B.22}$$

which gives rise to the **tripling coefficients**, C_{ijk}. Unlike double product integrals, which have an incredibly simple form and reduce to a single dot product, the triple product integral is seemingly much more complicated. Fortunately, the set of tripling coefficients is also sparse, so not all the individual coefficients need to be explicitly computed and stored. For spherical harmonics, the C_{ijk} correspond to Clebsch-Gordan coefficients, whose analytic values and properties are well studied [Tinkham, 2003].

5. **Double Product Projection.** The tripling coefficients also arise when computing the product of two spherical harmonic functions directly in the SH basis. We can compute the i^{th} coefficient of the SH projection of the product $c(\vec{\omega}) = a(\vec{\omega})b(\vec{\omega})$ as

$$c_i = \int_{\Omega_{4\pi}} y_i(\vec{\omega})\, c(\vec{\omega}) d\vec{\omega}, \tag{B.23}$$

$$= \int_{\Omega_{4\pi}} y_i(\vec{\omega})\, a(\vec{\omega})\, b(\vec{\omega}) d\vec{\omega}, \tag{B.24}$$

$$= \int_{\Omega_{4\pi}} y_i(\vec{\omega}) \left(\sum_j a_j\, y_j(\vec{\omega})\right) \left(\sum_k b_k\, y_k(\vec{\omega})\right) d\vec{\omega}, \tag{B.25}$$

$$= \sum_j \sum_k a_j\, b_k \int_{\Omega_{4\pi}} y_i(\vec{\omega})\, y_j(\vec{\omega})\, y_k(\vec{\omega}) d\vec{\omega}, \tag{B.26}$$

$$= \sum_j \sum_k a_j\, b_k\, C_{ijk}. \tag{B.27}$$

This expression states that the i^{th} coefficient of c is a linear combination of the, up to, $j \times k$ coefficients from a and b. The weighting of these terms is determined by the triping coefficients, which are *independent* of the particular choice of a and b. In effect, if we wish to efficiently compute the product projection of many pairs of functions, we only need to compute the tripling coefficients once.

The product projection simplifies further if we know one of the functions beforehand. For instance, if $b(\vec{\omega})$ is fixed, then we can construct a *transfer matrix*, $\mathbf{M}$, which directly transforms coefficients of any arbitrary function $\tilde{a}$ into the coefficients of the product $\tilde{c}$

using a single vector-matrix multiplication:

$$c_i = \int_{\Omega_{4\pi}} y_i(\vec{\omega})\, a(\vec{\omega})\, b(\vec{\omega}) d\vec{\omega}, \tag{B.28}$$

$$= \int_{\Omega_{4\pi}} y_i(\vec{\omega}) \left(\sum_j a_j\, y_j(\vec{\omega}) \right) b(\vec{\omega}) d\vec{\omega}, \tag{B.29}$$

$$= \sum_j a_j \int_{\Omega_{4\pi}} y_i(\vec{\omega})\, y_j(\vec{\omega})\, b(\vec{\omega}) d\vec{\omega}, \tag{B.30}$$

$$= \sum_j a_j\, \mathbf{M}_{ij}. \tag{B.31}$$

6. **Rotational Invariance.** The SH basis functions are *rotationally invariant*, which means that if g is a rotated copy of f, i.e.,

$$g(\vec{\omega}) = f(\mathbf{R}\vec{\omega}), \tag{B.32}$$

for any 3×3 rotation matrix $\mathbf{R}$, then

$$\tilde{g}(\vec{\omega}) = \tilde{f}(\mathbf{R}\vec{\omega}). \tag{B.33}$$

This property means that, in order to evaluate a rotated SH function $\tilde{g}$, we can either rotate the *lookup* into the unrotated approximation $\tilde{f}$ or lookup directly into the rotated approximation $\tilde{g}$. This property implies that spherical harmonic projection produces no aliasing.

7. **Rotation.** Spherical harmonics also support efficient rotation. This means that if we know $\tilde{f}$, we can compute the SH coefficients of the rotated function $\tilde{g}$ *exactly* by just applying a linear transformation to the projection coefficients of $\tilde{f}$. This linear transformation is itself a higher-dimensional rotation matrix, $\tilde{\mathbf{R}}$, where the i^{th} coefficient of the rotated function g

is simply:

$$g_i = \sum_j f_j \tilde{\mathbf{R}}_{ij}. \tag{B.34}$$

Due to the rotation invariance property, the coefficients in one band of f only influence the same band of coefficients in the rotated representation g. This leads to the following block-sparse structure:

$$\tilde{\mathbf{R}} = \left[\begin{array}{c|ccc|ccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline 0 & \mathbf{X} & \mathbf{X} & \mathbf{X} & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & \mathbf{X} & \mathbf{X} & \mathbf{X} & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & \mathbf{X} & \mathbf{X} & \mathbf{X} & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline 0 & 0 & 0 & 0 & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \cdots \\ 0 & 0 & 0 & 0 & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \cdots \\ 0 & 0 & 0 & 0 & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \cdots \\ 0 & 0 & 0 & 0 & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \cdots \\ 0 & 0 & 0 & 0 & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \cdots \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right]. \tag{B.35}$$

Several methods have been proposed for efficiently computing this coefficient transformation. Analytic forms can be derived for axis-aligned rotations, which can then be combined using Euler angle decomposition to construct rotations about arbitrary axes [Kautz et al., 2002]. This approach is only practical for low-order coefficients, and rotations of higher order SH functions are more efficiently implemented using recurrence relations [Ivanic and Ruedenberg, 1996, 1998; Blanco et al., 1997; Choi et al., 1999; Pinchon and Hoggan, 2007]. Křivánek et al. [2005c] instead proposed a fast approximate SH rotation method for small angles, which uses a truncated Taylor expansion of the rotation matrix. Green [2003] provides a more detailed summary of available methods.

If the function f has circular symmetry about the z-axis, then its projection consists only of zonal harmonics (only the $m = 0$ SH basis functions). Zonal harmonics (ZH) are an important subset of the full set of SH basis functions since they often lead to more efficient

operations. Furthermore, circularly symmetric functions are common in graphics. Most phase functions discussed in Chapter 4, for instance, exhibit circular symmetry.

Since ZH functions only have one non-zero coefficient per band, they can be rotated much more efficiently. This means that only one column of each band-matrix is needed:

$$\tilde{\mathbf{R}} = \left[\begin{array}{c|cccc|ccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline 0 & 0 & \mathbf{X} & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & \mathbf{X} & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & \mathbf{X} & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{X} & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{X} & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{X} & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{X} & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{X} & 0 & 0 & \cdots \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right]. \tag{B.36}$$

The non-zero elements in this matrix have an easy analytic form. To rotate a zonal harmonic function into direction $\vec{d}$ we simply need to apply the formula:

$$g_l^m = \sqrt{\frac{4\pi}{2l+1}} f_l^0 y_l^m(\vec{d}). \tag{B.37}$$

Note the similarity of this equation to the convolution in Equation B.13. In effect, rotation of a circularly symmetric function is the same as convolving a kernel with a delta function at the desired rotation axis.