# Optimal Spherical Harmonics Projection

WILLEM H. DE BOER
Mucky Foot Productions Ltd.
e-mail: reply@whdeboer.com

We introduce the reader to the mathematics behind projection of $n$-dimensional vectors into a basis on $n$-dimensional space, where $n$ can be anything upto and including infinity. We show how ideally one wants to project into the duals of a basis if this basis is not orthonormal, and provide the mathematics to formulate this operation in matrix form. The second part of the article discusses spherical harmonics projection of real-valued scalar functions on $S^2$, used in real-time global illumination applications and conclude that projection into the dual basis is equal to calculating a least-squares solution.

Categories and Subject Descriptors: I.3.7 [**Computer Graphics**]: Three-dimensional graphics and realism

General Terms: Algorithms, human factors, performance

Additional Key Words and Phrases: Spherical Harmonics, Projection, Basis

## 1. INTRODUCTION

Spherical harmonics have been used in the field of computer graphics for providing simplified solutions to the rendering equation, which is central to global illumination problems. One of the key concepts of using spherical harmonics is projecting a real-valued scalar function on $S^2$ (such as the incoming radiance of a particular point on the surface of an object) into the spherical harmonics. A projection of a function $f$ into the spherical harmonics[1] $\{Y_i\}$ yields coefficients $\{c_i\}$, which are the weights used in a linear combination of $\{Y_i\}$ to represent $f$. In many papers that treat spherical harmonics in computer graphics (including [Snyder 2002] [Ramamoorthi and Hanrahan 2001]) the theory behind spherical harmonics and projection is only briefly outlined. We feel that many people in the games field will benefit from a more in-depth discussion of the theory behind projection. An elementary background in linear algebra, and multi-variable calculus is assumed of the reader.

**Goal** - Our goal is to provide an insight into spherical harmonics projection of functions, by first drawing an analogy with the finite-dimensional case and later on providing the machinery to optimally project an arbitrary function on $S^2$ into the spherical harmonics.

**Overview** - We begin by considering projection in the finite-dimensional vector space $R^n$. Section 2 shows how projection works in $R^n$, when it fails, and how to solve this. In section 3 we consider the infinite-dimensional space. The first part

[1]We will denote the *real-valued* spherical harmonics by $\{Y_i\}$. Their complex equivalents are generally not directly used in computer graphics.

of this section discusses projection of functions into the spherical harmonics, and the second part describes optimal projection. In section 4 we will briefly discuss an example of other applications in computer graphics that may see benefit from projection into the dual basis.

## 2. PROJECTION IN A FINITE DIMENSIONAL SPACE

Consider the finite $n$-dimensional Euclidean vector space, $R^n$. A vector $\mathbf{v}$ in this space can be uniquely written as a linear combination of a complete set of basis vectors on $R^n$:

$$\mathbf{v} = v_1\bar{e}_1 + v_2\bar{e}_2 + ... + v_n\bar{e}_n = \sum_{i=1}^{n} v_i\bar{e}_i$$

Where $\{\bar{e}_i\}$ are the basis vectors, and $\{v_i\}$ are the coordinates of $\mathbf{v}$ in $\{\bar{e}_i\}$. There are certain problems in computer graphics that require us to be able to move the other way: How can we compute $\{v_i\}$ when given a basis? Let us choose the basis $\{\bar{e}_i\}$ as this given basis and let us assume that this basis is orthonormal, so $\langle\bar{e}_i|\bar{e}_j\rangle = 1$ only if $i = j$, 0 if $i \neq j$ (where $i, j = 1, ..., n$), and $\langle\ |\ \rangle$ denotes the dot product. Due to orthonormality of $\{\bar{e}_i\}$, we can obtain the component $v_i$ of $\mathbf{v}$ by performing a dot product with the appropriate basis vector:

$$\langle\mathbf{v}|\bar{e}_i\rangle = \langle(\sum_{j=1}^{n} v_j\bar{e}_j)\ |\ \bar{e}_i\rangle = \sum_{j=1}^{n} v_j\langle\bar{e}_j|\bar{e}_i\rangle = v_i \tag{1}$$

Doing this for every $\bar{e}_i$ in the basis results in a vector $\mathbf{v} = [v_1, v_2, ..., v_n]^t$ of coordinates of $\mathbf{v}$ in the basis $\{\bar{e}_i\}$. Equation (1) is called projection: We are projecting a vector $\mathbf{v}$ into a basis $\{\bar{e}_i\}$, to obtain the coordinates of $\mathbf{v}$ in that basis.

### 2.1 The Duals

Obtaining the coordinates of a vector $\mathbf{v}$ in a basis by performing a projection works because of the assumed orthonormality of the particular basis we are projecting into, as we have just shown. What if we want to project $\mathbf{v}$ into a basis that is not orthonormal? Suppose $\{\bar{e}_i\}$ is not orthonormal. Then using Eq. (1) to calculate the component of $\mathbf{v}$ will not work, i.e. $\langle\bar{e}_i|\bar{e}_j\rangle$ will not in general vanish for $i \neq j$. In order to obtain the components of $\mathbf{v}$ in a non-orthonormal basis we must use the dual of $\{\bar{e}_i\}$, denoted by $\{\bar{e}_i^*\}$. The *defining property* of the dual is that $\langle\bar{e}_i^*|\bar{e}_j\rangle = 1$ if $i = j$, 0 if $i \neq j$ (where $i, j = 1, ..., n$). By using $\bar{e}_i^*$ rather than $\bar{e}_i$ in Eq. (1), we can obtain the components of $\mathbf{v}$ in $\{\bar{e}_i\}$ even if this basis is not orthonormal:

$$\langle\mathbf{v}|\bar{e}_i^*\rangle = \langle(\sum_{j=1}^{n} v_j\bar{e}_j)\ |\ \bar{e}_i^*\rangle = \sum_{j=1}^{n} v_j\langle\bar{e}_j|\bar{e}_i^*\rangle = v_i \tag{2}$$

Another interesting result, which is very useful to infinite-dimensional projection (as we will shall see further on) is the following:

$$\langle\mathbf{v}|\bar{e}_i\rangle = \langle(\sum_{j=1}^{n} v_j\bar{e}_j)\ |\ \bar{e}_i\rangle = \sum_{j=1}^{n} v_j\langle\bar{e}_j|\bar{e}_i\rangle \tag{3}$$

$$\langle \mathbf{v}|\bar{e}_i\rangle = \langle(\sum_{j=1}^{n} w_j \bar{e}_j^*) \mid \bar{e}_i\rangle = w_i \tag{4}$$

Note how $\{v_j\}$ are the components of $\mathbf{v}$ expressed in $\{\bar{e}_j\}$, and $\{w_j\}$ are the components of $\mathbf{v}$ expressed in $\{\bar{e}_j^*\}$. By equating Eq. (3) and Eq. (4) we obtain:

$$w_i = \sum_{j=1}^{n} v_j \langle \bar{e}_j|\bar{e}_i\rangle = v_1\langle \bar{e}_1|\bar{e}_i\rangle + v_2\langle \bar{e}_2|\bar{e}_i\rangle + ... + v_n\langle \bar{e}_n|\bar{e}_i\rangle$$

Or, the component $w_i$ can be obtained by calculating a weighted sum of the components $\{v_i\}$. In matrix notation:

$$\begin{aligned} \mathbf{w} &= \mathbf{A}\mathbf{v} \\ \mathbf{v} &= \mathbf{A}^{-1}\mathbf{w} \end{aligned} \tag{5}$$

Where $\mathbf{A}$ is an $n \times n$ matrix, with components $a_{ij} = \langle \bar{e}_i|\bar{e}_j\rangle$. What does this mean? It means that if we have the components of $\mathbf{v}$ in $\{\bar{e}_i\}$, we can calculate the components of $\mathbf{v}$ in $\{\bar{e}_i^*\}$ by left-multiplying[2] $\mathbf{v}$ by the matrix $\mathbf{A}$. Conversely, the components of $\mathbf{v}$ in $\{\bar{e}_i\}$ can be obtained by left multiplying $\mathbf{w}$ by the inverse $\mathbf{A}^{-1}$. Note that if $\{\bar{e}_i\}$ is an orthonormal basis, the matrix $\mathbf{A}$ is the identity matrix, and therefore $\{\bar{e}_i\}$ is equal to its dual $\{\bar{e}_i^*\}$.

## 3. PROJECTION IN INFINITE DIMENSIONAL SPACE

### 3.1 Spherical Harmonics Projection

Recent research into global illumination solutions that are usable in real-time applications has seen a lot of focus on spherical harmonics. Spherical harmonics form a complete basis on $S^2$ (the unit spere in $R^3$), the most important function domain for global illumination. A function defined on $S^2$ can be represented by a linear combination of spherical harmonics $\{Y_i\}$. Projection of a real-valued scalar function $f$ on $S^2$ into $\{Y_i\}$ is equivalent to Eq. (1), except we are working in an infinite-dimensional space, and the way we perform a dot product in this particular space is by evaluating an integral:

$$\begin{aligned} \langle f|Y_i\rangle &= \int_0^{2\pi}\int_0^{\pi} f(\theta,\phi)\, Y_i(\theta,\phi) \sin\theta \, d\theta d\phi \\ &= \int_0^{2\pi}\int_0^{\pi} (\sum_{j=1}^{\infty} c_j\, Y_j(\theta,\phi))\, Y_i(\theta,\phi) \sin\theta \, d\theta d\phi \\ &= \sum_{j=1}^{\infty} c_j\, (\int_0^{2\pi}\int_0^{\pi} Y_j(\theta,\phi)\, Y_i(\theta,\phi) \sin\theta \, d\theta d\phi) \\ &= \sum_{j=1}^{\infty} c_j\, \langle Y_j|Y_i\rangle = c_i \end{aligned} \tag{6}$$

[2]Note that we assume vectors to be column vectors, so left-multiplication by a matrix makes sense.

Where $\theta$ represents elevation over the sphere, and $\phi$ represents azimuth. Note how Eq. (6) is the same as Eq. (1), only $f$ has replaced $\mathbf{v}$, and $Y_i$ has replaced $\bar{e}_i$. The above equation works just like Eq. (1), because the spherical harmonics form an orthonormal basis on $S^2$ by their construction: $\langle Y_i|Y_j\rangle = 1$ if $i = j$, or 0 if $i \neq j$.

In the case of projecting over the *hemisphere* $\mathbf{H}$, the spherical harmonics do not form an orthonormal basis, so ideally we would want to project $f$ over the duals $\{Y_i^*\}$, like we did in Eq. (2) for the $n$-dimensional case. For computer graphics purposes, it is generally not easy or desirable to calculate the duals directly. It is easier to perform a matrix multiplication of the coefficients of a function in the spherical harmonics. We proceed as follows: We project $f$ into $\{Y_i\}$, to obtain the coefficients in the dual basis $\{Y_i^*\}$ (as in Eq. (4)):

$$\begin{aligned}
\langle f|Y_i\rangle &= \int_0^{2\pi}\int_0^{\frac{1}{2}\pi}(\sum_{j=1}^{\infty} b_j\, Y_j^*(\theta,\phi))\, Y_i(\theta,\phi)\sin\theta\, d\theta d\phi \\
&= \sum_{j=1}^{\infty} b_j\, (\int_0^{2\pi}\int_0^{\frac{1}{2}\pi} Y_j^*(\theta,\phi)\, Y_i(\theta,\phi)\sin\theta\, d\theta d\phi) \\
&= \sum_{j=1}^{\infty} b_j\, \langle Y_j^*|Y_i\rangle = b_i
\end{aligned} \tag{7}$$

Where $\{b_i\}$ are the coefficients of $f$ in the dual basis. Note how the domain of integration over $\theta$ is $[0, \frac{1}{2}\pi]$, because we are integrating over the hemisphere. Also note that we never need to evaluate $\{Y_j^*\}$ directly; we just perform an integration of $f$ with the spherical harmonic $Y_i$. We then calculate the components of the matrix $\mathbf{A}$ as follows:

$$a_{ij} = \langle Y_i|Y_j\rangle = \int_0^{2\pi}\int_0^{\frac{1}{2}\pi} Y_i(\theta,\phi)\, Y_j(\theta,\phi)\sin\theta\, d\phi d\theta \tag{8}$$

And left-multiply the vector $\mathbf{b} = [b_1, ..., b_n]^t$ by the inverse $\mathbf{A}^{-1}$, to obtain the coefficients $\{c_i\}$ in $\{Y_i\}$. Note how $\mathbf{A}$ is equal to the matrix $A = \{A_{jk}\}$ in Sloan [Snyder 2003] used for calculating the least-squares optimal projection of a function, and how $\mathbf{b}$ is equal to the vector of coefficients $b = \{b_i\}$ in the same paper. We can therefore re-interpret $b$ as the coefficients of $f$ in the spherical harmonics dual basis $\{Y_i^*\}$. Then the matrix $A$ can be interpreted as the representation of a change of basis transformation, equal to the one in Eq. (5). Therefore, we can conclude that least-squares optimal projection is equal to our method. Also note that if we take the domain of integration to be $S^2$ entirely, the matrix $\mathbf{A}$ is equal to the identity matrix, by the very construction of the spherical harmonics.

[3]Note that typically, $m = 9$ or $m = 25$ for diffuse cases [Snyder 2002].

### 3.2 Optimal SH Projection Overview

The general algorithm for projecting a function $f$ on the hemisphere $H$ into the spherical harmonics for $m$ coefficients[3] is then:

(1) Calculate the inverse $\mathbf{A}^{-1}$ of the $m \times m$ matrix $\mathbf{A}$, where each entry $a_{ij}$ is calculated using Eq. (8)

(2) Calculate the $m$ projection coefficients $\{b_i\}$ of $f$ in the dual basis, using Eq. (7)

(3) Left-multiply $\mathbf{b} = [b_1, ..., b_m]^t$ by $\mathbf{A}^{-1}$ to obtain $\mathbf{c} = [c_1, ..., c_m]^t$: The coefficients of the projection of $f$ into $\{Y_i\}$.

Note that we can pre-calculate $\mathbf{A}^{-1}$ and use it to project any function defined on the hemisphere into $\{Y_i\}$. Any integration scheme can be used for 1. and 2. Green [Green 2003] recommends using Monte-Carlo integration with stratified sampling.

## 4. DISCUSSION

Projecting into the dual basis does not only benefit spherical harmonics projection on the hemisphere. It can benefit numerous other applications in computer graphics that work with a non-orthonormal basis. One example is texture reconstruction. Textures are typically point-sampled representations of a continuous image signal, and ideally their reconstruction involves evaluating the convolution of the texture with the $sinc(x)$ filter in the spatial domain. In real-time computer graphics, however, reconstruction is almost always performed using bilinear basis functions. This is especially true of all current graphics hardware. It is therefore tempting to replace the original construction process with projection into the bilinear basis, $\{B_i\}$. Due to the overlap of neighbouring basis functions, it is easy to see that we would need to project into the duals, $\{B_i^*\}$:

$$\langle I|B_i^* \rangle = \int_0^1 \int_0^1 (\sum_{j=1}^{M \times N} I_j \, B_j(x,y)) \, B_i^*(x,y) \, dydx = I_i$$

Where $\{I_j\}$ are the coefficients of the image signal in the bilinear basis (ie, the final texel values), and $M$ and $N$ are the width and height of the texture respectively.

## 5. CONCLUSION

We have presented the mathematics behind projection of a vector in $n$-dimensional space, where $n$ is any positive number upto and including infinity. Using elementary linear algebra we have shown that projection of a vector into the dual of a non-orthonormal basis can be formulated as a matrix multiplication. One application of projection in computer graphics is global illumination using spherical harmonics, and we have shown how to optimally project an arbitrary real-valued scalar function on $S^2$ into the spherical harmonics. We have also very briefly discussed another application in computer graphics that would ideally see a projection into the dual basis.

## 6. ACKNOWLEDGEMENTS

A very big thanks to Peter-Pike Sloan for his valuable suggestions and discussions.

### REFERENCES

GREEN, R. 2003. Spherical harmonic lighting: The gritty details. *Game Developers Conference Proceedings*.

RAMAMOORTHI, R. AND HANRAHAN, P. 2001. An efficient representation of irradiance environment maps. *SIGGRAPH Proceedings*.

SNYDER, P. S. J. H. J. H. J. 2003. Clustered principal components for precomputed radiance transfer. *SIGGRAPH Proceedings*.

SNYDER, P. S. J. K. J. 2002. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency environments. *SIGGRAPH Proceedings*.