Wavefront Data Reporting and Terminology

Larry N. Thibos, PhD

One of the most important limitations to current research in visual optics is the "jargon gap" that separates clinically trained researchers and practitioners from the technical scientists who inhabit the world of modern optics. To help overcome this communication gap, this brief tutorial summarizes the terminology recommended by the Optical Society of America for reporting wavefront aberration data from eyes and offers some suggestions for visualization of aberration data (Thibos LN, Applegate RA, Schwiegerling JT, Webb R. Standards for reporting the optical aberrations of eyes. Trends in Optics and Photonics 2000. Washington, DC: Optical Society of America). The plan is to 1) define the wavefront aberrations of an eye in English, and then again in the language of mathematics, 2) to show how the aberration structure of any eye, no matter how complicated or idiosyncratic, can be understood by a systematic deconstruction into fundamental building blocks, and 3) to illustrate some methods for graphical display of the full spectrum of these fundamental building blocks in order to gain insight into the nature of ocular aberrations.

To begin, we need a clear picture of the meaning of the term "wavefront aberration." The easiest way to conceive of the wavefront aberration of an eye is to examine the light that is reflected out of an eye due to a laser beam that is focused onto the retina at some point P' (Fig 1A). We can think of this light in terms of the bundle of rays that is reflected through the eye's pupil, perhaps focusing at some point P that would represent the eye's far point, in the case of a myopic individual. Alternatively, we can visualize a surface that is everywhere perpendicular to all the rays in the bundle. This surface is the wavefront of light being reflected out of the eye, and the *shape* of this surface is the essence of a "wavefront aberration function" for the eye. Of course, this wavefront is propagating forward at the speed of light, but if we could freeze time and take a snapshot of an emerging wavefront, then it might look like the illustration in Figure 1.

To understand why the shape of the wavefront is critical, think about the wavefront that would emerge from an optically perfect eye, one that is emmetropic and free of all imperfections. Such an eye would reflect rays that are parallel, intersecting at a far point infinitely far away, and thus the reflected wavefront would be a circular piece of a plane wave with the same diameter as the eye's pupil. The optical imperfections of a real eye, therefore, are revealed by comparing the shape of the actual wavefront reflected from the eye to this ideal plane wave. We can make this comparison quantitatively by measuring the distance between the reflected wavefront and the ideal wavefront, which, for convenience, we place in the (x,y) plane of the eye's entrance pupil (Fig 1B). This distance between the actual wavefront and the pupil plane represents an optical error that varies from point-to-point across the pupil and therefore can be quantified by a function $W(x,y)$, where the letter W stands for wavefront error.

The wavefront error for an eye suffering only from defocus has a parabolic shape (Fig 2A). This shape is defined mathematically by a simple equation:

$$W(x,y) = 2(x^2 + y^2) - 1 \quad \text{(defocus)} \tag{1}$$

that contains the key expression $x^2 + y^2$. There are some other constants in the equation that are not essential, but which provide some convenient features. For example, the -1 in the equation forces the average error to be zero. This corresponds to taking our snapshot of the reflected wavefront at the moment when it passes through the eye's pupil. If any part of the wavefront emerges early, it generates positive errors, and if another part of the wavefront lags behind, it generates negative errors. If we

From the School of Optometry, Indiana University, Bloomington, IN.

Supported by National Institutes of Health grant R01 EY05109.

Presented at the 2nd International Congress of Wavefront Sensing and Aberration-free Refractive Correction, in Monterey, CA, on February 9-10, 2001.

Correspondence: Larry N. Thibos, PhD, School of Optometry,

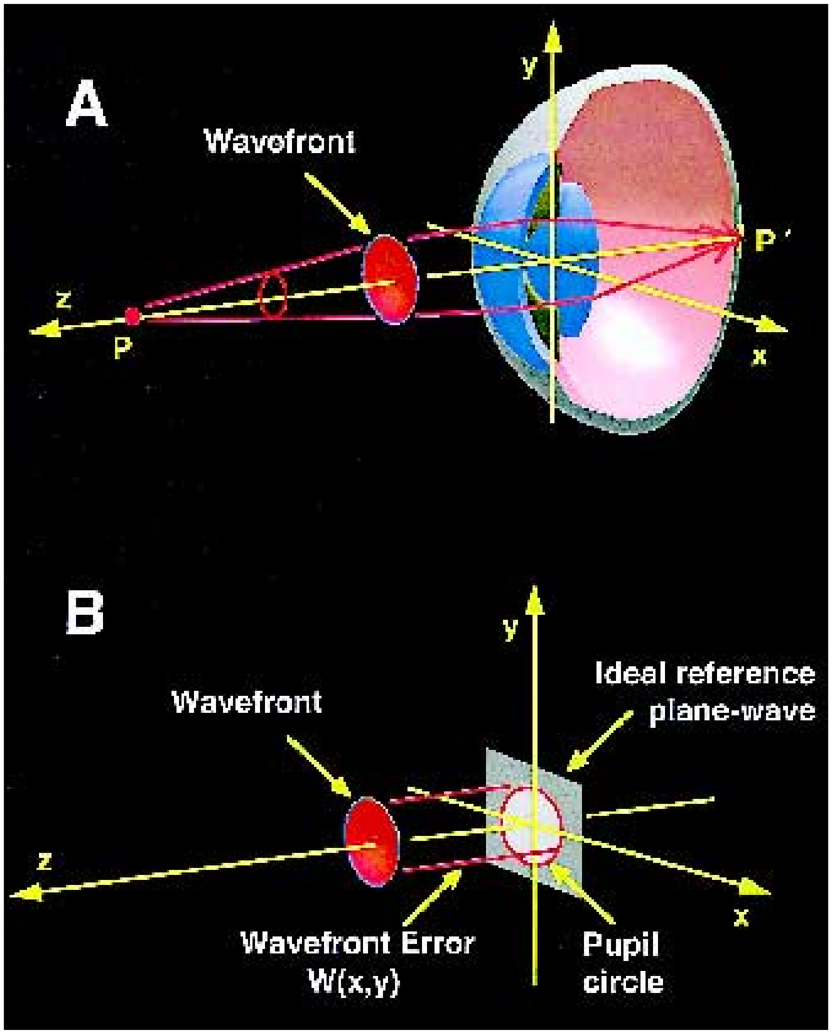

Figure 1. A) The shape of the reflected wavefront of light due to a point source at P' determines the wavefront aberration of the eye. **B)** The wavefront aberration W(x,y) is defined as the distance between each point on the reflected wavefront and some ideal reference wavefront. For an emmetropic eye, the ideal reference is a plane wave in the pupil plane. In this case the shape of the reflected wavefront is also the aberration function W(x,y).

catch the wavefront at just the right moment, the positive errors will exactly cancel the negative errors and the mean will be zero, just like this equation says.

It turns out that the equations used to represent wavefront errors of eyes often look much simpler if written not in terms of the x-y coordinates of a rectangular coordinate system in the pupil plane, but instead in terms of the polar coordinates r and u. An eye with astigmatism, for example, will reflect a saddle-shaped wavefront, as shown in Figure 2B, which has a simple algebraic equation when written in polar form

$$W(r,u) = r^2\cos 2u \quad \text{(astigmatism)} \qquad (2)$$

Now that we see how to conceive of the eye's wavefront aberration function and how to describe it mathematically with an equation, the next step is to combine simple wavefronts like those illustrated in Figure 2 to make more complex wavefronts that can describe the aberration structure of real eyes. To do this we need a catalog of basic shapes that we can add together. There are lots of ways to make such a catalog, but the most popular scheme is credited to the optical scientist, Zernike. Since the basic building blocks of Zernike will be the basis for describing the aberration structure of eyes, they are known as *basis functions*. Each Zernike basis function is the product of two other functions, one of which depends only on radius and the other, which depends only on meridian. For example, the astigmatism wavefront described by Eq (2) contains a polynomial in the radius variable r, in this case a second-order polynomial, and a sinusoidal term involving the meridian variable u, in this case with

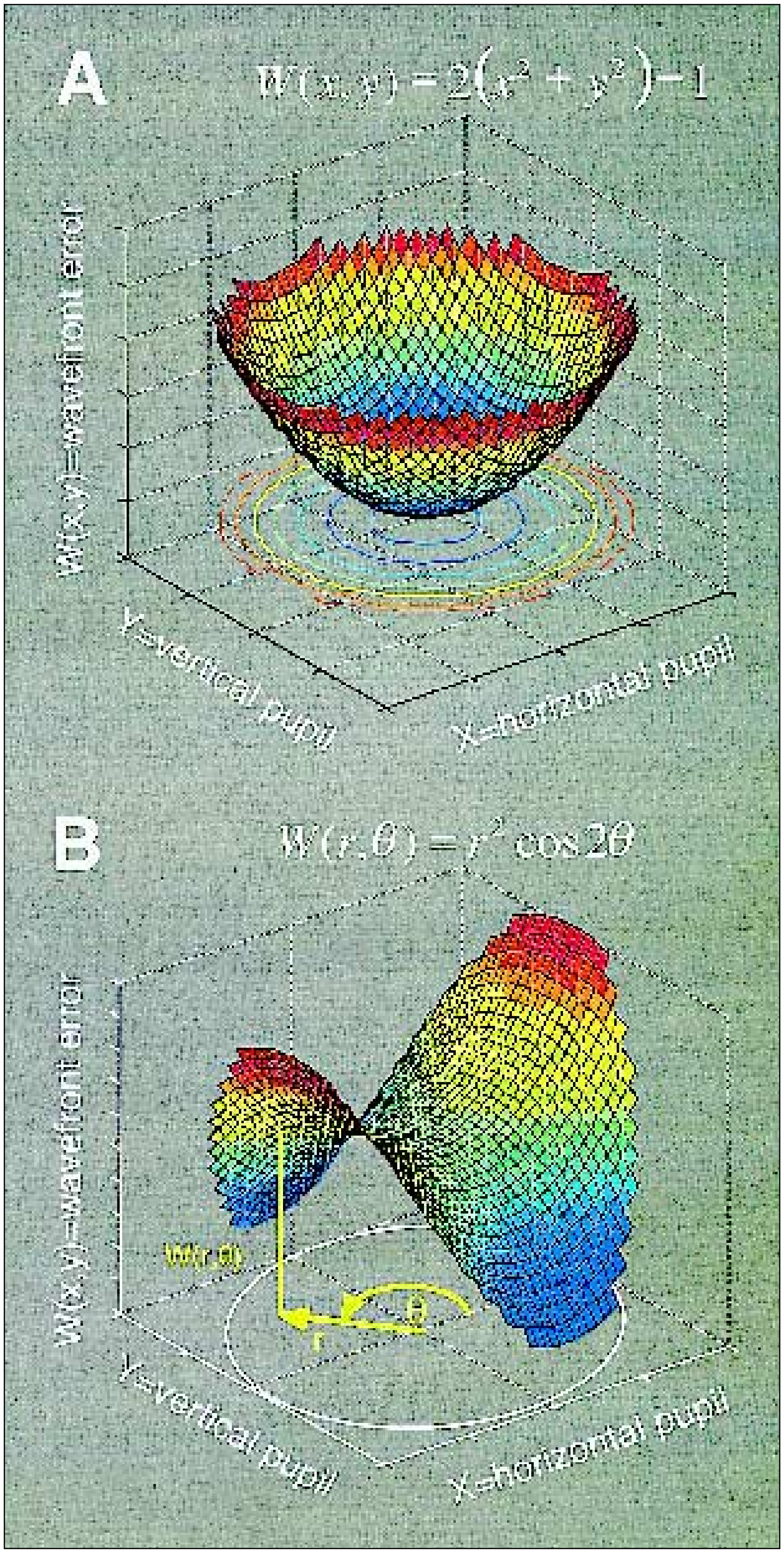

Figure 2. Examples of wavefront aberration functions: **A)** defocus, specified in rectangular coordinates, and **B)** astigmatism, specified in polar coordinates.

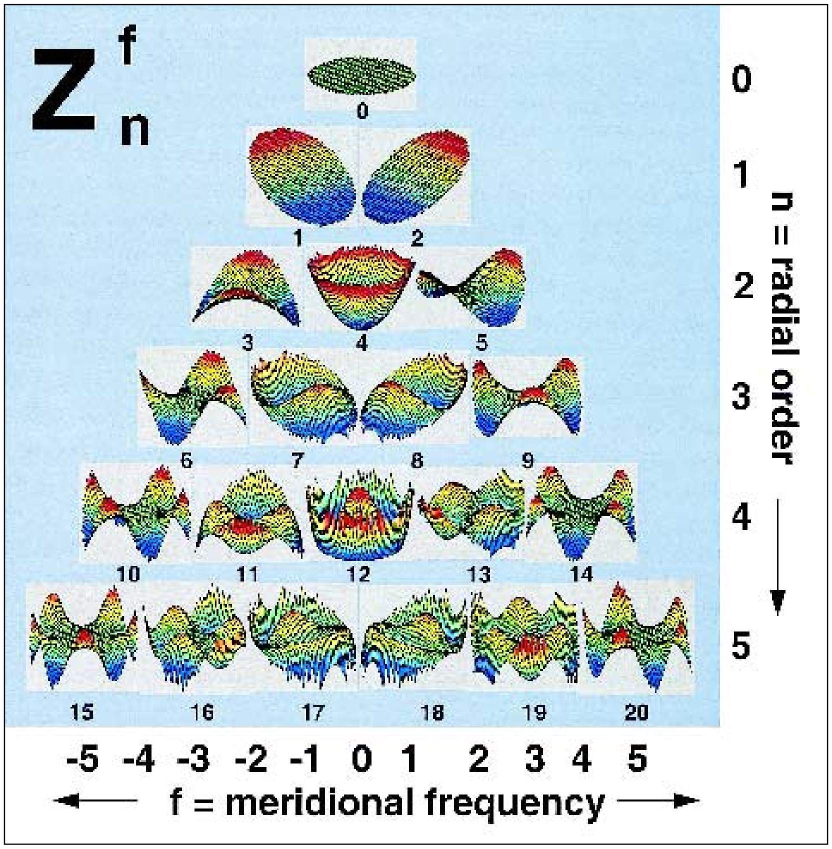

Figure 3. A periodic table of Zernike basis functions. Subscript *n* indicates radial order, which gives the row number in the table. Superscript *f* indicates meridional frequency, which gives the column number in the table.

a frequency of 2. This basic pattern of a polynomial times a sinusoid occurs for all of the Zernike functions. Another example is a basis function called "coma," which is the product of a third-order polynomial and a sinusoidal harmonic of frequency 1.

$$W(r,u) = (3r^3 - 2r)\sin u \quad \text{(coma)} \qquad (3)$$

One of the attractive features of the set of Zernike functions is that they are mutually orthogonal, which means they are independent of each other mathematically. Another convenient feature is that every mode except the first has a zero mean, and they all are scaled so as to have unit variance (Thibos LN, Applegate RA, Schwiegerling JT, Webb R. Standards for reporting the optical aberrations of eyes. Trends in Optics and Photonics 2000. Washington, DC: Optical Society of America). This puts all of the modes on a common basis so their relative magnitudes can be compared easily. The orthogonality of the Zernike basis functions makes it easy to calculate the total variance in a wavefront as the sum of the variances in the individual components.

The Zernike basis functions, or "modes" as they are often called, are systematically arranged into a periodic table with the shape of a pyramid (Fig 3). Each row in the pyramid corresponds to a given order of the polynomial component of the function and each column corresponds to a different meridional frequency. By convention, harmonics in cosine phase are assigned positive frequencies and harmonics in sine phase are assigned negative frequencies. Although each mode can be assigned a single reference number, a more natural numbering system is a double-script notation, which designates each basis function according to its order and frequency. The radial order is used as a subscript and the meridional frequency is used as a superscript to unambiguously and conveniently identify each mode.

Given this catalog of fundamental building blocks, we may now describe the aberration structure of an eye mathematically as the weighted sum of Zernike basis functions.

$$W(r,u) = \underset{\text{order}}{(}\ \underset{\text{frequency}}{(}\ c_n^f Z_n^f \qquad (4)$$

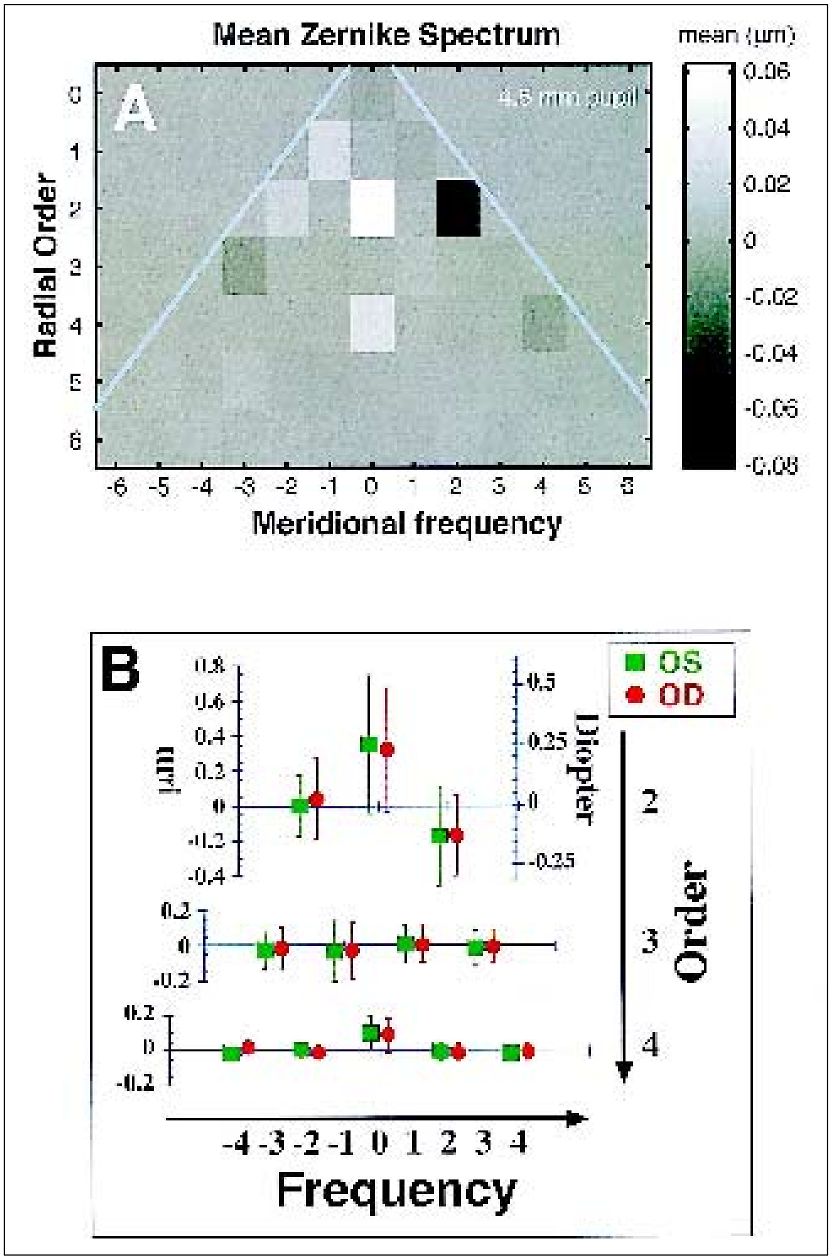

Figure 4. Two methods for displaying the Zernike spectrum. The pyramidal layout of data for individual modes corresponds to the periodical table in Figure 3. **A)** Each Zernike mode is represented by a rectangle, the intensity of which represents the coefficient value. **B)** Each mode is represented by a subgraph that shows the mean (symbol) and standard deviation (error bars) for right and left eyes. Data source is the Indiana Aberration Study, 2000.

Such a description is called a Zernike expansion of the wavefront aberration. The weight c_n^f, which must be applied to each basis function when computing the sum, is called an aberration coefficient. Each aberration coefficient is just a number, with physical units typically specified in micrometers, or sometimes reported in the units of the wavelength of light. The aberration coefficients of a Zernike expansion are analogous to the Fourier coefficients of a Fourier expansion, which are in turn analogous to the energy spectrum of a light source. Thus, it is common to speak of the "Fourier spectrum" of a waveform, and in the same way we may speak of the "Zernike spectrum" of the eye's optical system.

This development leads to our final topic—to explore various ways of displaying the spectrum of aberration coefficients associated with a Zernike expansion. To visualize the two-dimensional Zernike spectrum as a pyramid, we can assemble a collection of subregions in our graph that correspond to the pyramid of Zernike basis functions. Within each sub-region, we can show information about the corresponding Zernike mode. For example, we can display the value of the Zernike coefficients of a single eye, or the mean of a population of eyes, using a pyramid of rectangular blocks. As shown in Figure 4A, the intensity of each block indicates the value of the coefficient for the corresponding mode. White signifies a large positive value, black signifies a large negative value, and grey signifies zero. The advantage of this visualization scheme is that it permits an immediate visual assessment of the relative magnitude of the various modes that make up the aberration structure of an eye.

To show the mean and standard deviation of each coefficient in a population of eyes, we might use each subregion in the pyramid to plot a small graph (Fig 4B). There is ample room in such a display for two symbols that separately report results from a population of right and left eyes. Such a graph makes it easier to see which frequencies and which orders have the largest coefficients and therefore have the largest effect on the quality of the retinal image. Incidentally, it is common practice to omit the zero-order and first-order coefficients from the pyramid spectrum because they correspond to prismatic deviations and constant offsets, which have no bearing on the optical quality of the retinal image for monochromatic light and consequently are not of interest in most applications.

A more adventurous scheme might use the subregions of the pyramid to plot frequency histograms to show the distribution of Zernike coefficients in greater detail. Alternatively, we may wish to simplify the picture by reducing the full, two-dimensional Zernike spectrum to one dimension by summing across rows or down columns of the pyramid. The proper way to perform this summing operation is to compute the square root of the sum of the squares of the various Zernike coefficients. This accounting scheme is made possible by the orthogonality property of the Zernike basis functions and corresponds to the summing of the variances contributed by each mode to the total variance.

To simplify the presentation even more, it may be useful to define a quantity M_e called "equivalent defocus," which is the amount of defocus required to produce the same wavefront variance as found in one or more higher-order aberrations. A simple formula allows us to compute equivalent defocus in diopters if we know the total wavefront variance in the Zernike modes in question

$$M_e = 4\pi\sqrt{3}\,\frac{RMS\ Error}{Pupil\ Area} \quad (5)$$

In this equation, "RMS error" is shorthand for the square root of wavefront variance. Of course, we must keep in mind that the kind of optical blur produced by higher order aberrations is not exactly the same as the blur produced by defocus. Nevertheless, this concept of equivalent defocus helps interpret the Zernike coefficients in familiar dioptric terms. The basis of the equivalent defocus concept is the notion that the imaging quality of an eye is determined primarily by wavefront variance, and it doesn't matter which Zernike mode produces that variance. At present, this assumption is an article of faith amongst aberrometrists. However, we don't yet know whether wavefront variance is the best predictor of visual performance. There are many other metrics of image quality that can be derived from the wavefront aberration function that may prove to be more useful. For example, metrics associated with the point-spread function and the optical transfer function of the eye have a good track record in vision science for correlating optics and vision, and these may turn out to be more useful than wavefront variance in the clinic. Hopefully, answers to this important question will be featured in future meetings of the International Wavefront Congress.