Modeling a Large Baseline Optical Stellar Interferometer Using Zernike Polynomials[1]

Theo ten Brummelaar

1. INTRODUCTION

Due to the complex nature of the turbulent atmosphere and the optical systems themselves, analytical solutions to many of the problems associated with a large baseline optical stellar interferometer do not always exist. Thus, to answer some questions regarding the behavior of such a system computer simulations are required.

The most common method of modeling an atmospherically aberrated wavefront is to generate large phase screens containing the correct statistical distribution and move them past the aperture at a given velocity, nominally the current wind speed. These phase screens contain many pixels and must be re-generated for each simulation. For example, for a one second run for two 1m apertures 100 m apart and a wind speed of $10 \, \text{ms}^{-1}$ one would require a phase screen containing 10^7 pixels to sample 1cm patches every 10 ms for one second. There will then be 100×100 pixels within each aperture which must be individually tracked through the instrument. Calculating average phase gradient or the effect of a non-ideal optical surface will require operations on all 10000 pixels for each sample.

The polynomials of Zernike (Wang & Silva 1980, Born & Wolf 1989) are a set of normalized orthogonal functions defined on a circle and can be used to represent the wavefront within a circular aperture. Using Zernike coefficients to represent the wavefront greatly reduces the number of parameters that need to be tracked through an optical system. Many authors have used these functions to investigate and model an atmospherically perturbed wavefront (Fried 1965, Noll 1976, Hogge & Butts 1976, Hu et al. 1989, Roddier 1990); however, these methods are restricted to an analysis of the covariance matrix of the Zernike coefficients. Random series of coefficients generated in this way have the correct statistics but do not have a realistic time evolution.

Wavefronts can also be generated by using the temporal power spectra of the Zernike coefficients In this way a suite of random coefficients is created that not only reflects the correct statistical variances but also has a smooth evolution in time. Furthermore, if the polynomials are used to represent physical path-length, rather than phase, the model becomes wavelength independent. With the coefficients scaled in this way many of the optical

systems within the instrument become easy to represent, reducing the computation time for each optic from a calculation for every pixel to one for each Zernike mode. As shall be shown, complex wavefronts can be well represented by as few as 105 Zernike modes, without the need to choose pixel size, numbers or wavelength in advance.

I shall use the nomenclature set out by Noll (1976), which is summarized in appendix A.

2. ATMOSPHERIC TURBULENCE MODEL

The derivation of the temporal power spectra of Zernike coefficients uses the Tayler hypothesis and the spatial power spectrum of phase fluctuations at ground level (Roddier et al. 1993, ten Brummelaar 1995). The spectrum most often used was defined in Equation 21 of the paper by Noll (1976) where the power is proportional to the spatial frequency $\mathbf{k}$ to the power of $-11/3$. This ignores the effects of the outer scale length of turbulence and so the power goes to infinity for large spatial scales. While this does not seriously affect single aperture models recent measurements (Davis et al. 1995) have shown that this is not a realistic assumption for large arrays. The outer scale of turbulence L_0 can be taken into account by using an exponential model and we will write the phase spatial power spectrum at ground level

$$W_\phi(\mathbf{k}) = 0.023 \left(\frac{R}{r_0}\right)^{5/3} k^{-11/3} \left(1 - e^{-\left(\frac{kL_0}{2\pi R}\right)^2}\right) \text{rad}^2 \tag{1}$$

where R is the aperture radius and r_0 is Fried's parameter.

Following the techniques set out by Roddier et al. (1993) and using Equation 1 for the phase spatial power spectrum, the Zernike coefficient temporal power spectrum can be calculated resulting in

$$
\begin{aligned}
\Phi_{a_j}(f, r_0, R, v_\perp) &= \frac{0.023}{\pi^2}(n+1)\left(\frac{R}{r_0}\right)^{5/3}\frac{R}{v_\perp} \\
&\times \int_{-\infty}^{\infty} dk_y \, (k_x^2 + k_y^2)^{-17/6} \left|J_{n+1}(2\pi\sqrt{k_x^2 + k_y^2})\right|^2 \\
&\times \left(1 - e^{-(k_x^2+k_y^2)\left(\frac{L_0}{2\pi R}\right)^2}\right) \\
&\times \begin{cases} 2 \cos^2 m\phi & m \neq 0, \text{even } j \\ 2 \sin^2 m\phi & m \neq 0, \text{odd } j \\ 1 & m = 0 \end{cases}
\end{aligned}
\tag{2}
$$

where f is the frequency in Hertz, $v_\perp$ is the perpendicular wind speed, assumed to to blow in the direction of the X axis, and k_x and k_y are the Cartesian components of $\mathbf{k}$ such that $k_x = Rf/v_\perp$ and $\tan\phi = k_y/k_x$. In this equation n, m and j are the Zernike polynomial indices.

An approximate form of the Zernike coefficient temporal power spectrum is (ten Brummelaar 1995)

$$\Phi_{a_j}(f, r_0, R, v_\perp) = 0.092 (2\pi)^{\frac{11}{3}} \left(\frac{R}{r_0}\right)^{\frac{5}{3}} \frac{R}{v_\perp}(n+1)\left(\frac{f}{f_0}\right)^{-\frac{8}{3}}\frac{J_{n+1}^2(f/f_0)}{(f/f_0)^2}\left(1 - e^{-\left(\frac{fL_0}{2\pi v_\perp}\right)^2}\right) \tag{3}$$

where $f_0 = v_\perp/(2\pi R)$. Both Equations 2 and 3 produce spectra with equal amounts of power and similar high frequency roll-offs. The approximation is much easier to calculate but does not correctly model the low frequency behavior. Apart from calculation time, little practical difference has been found between the two formulations.

Both methods result in power spectra with the units of rad^2 and are of the form

$$\Phi_{a_j}(f, r_0, R, v_\perp) = \left(\frac{1}{r_0}\right)^{5/3} \Phi'_{a_j}(f, R, v_\perp). \tag{4}$$

Wavelength dependence is part of the r_0 parameter

$$r_0(\lambda) = \left(\frac{\lambda}{\lambda_0}\right)^{6/5} r_{\lambda_0} \tag{5}$$

where λ_0 is a reference wavelength (usually $500\,\mathrm{nm}$) and r_{λ_0} is the Fried parameter at the reference wavelength. Combining Equations 4 and 5 results in

$$\Phi_{a_j}(f, r_0, R, v_\perp) = \left(\frac{\lambda_0}{\lambda}\right)^2 \left(\frac{1}{r_{\lambda_0}}\right)^{5/3} \Phi'_{a_j}(f, R, v_\perp). \tag{6}$$

Note that the power spectra scale with the inverse wavelength squared.

For a given aperture radius and wind speed the functions $\Phi'_{a_j}(f, R, v_\perp)$ need only be calculated once for a range of discrete frequencies $(f_1, f_2, \cdots, f_N)$ up to a chosen maximum polynomial order j_{max}. These frequencies need not be evenly spaced nor does N need to be a power of two. Care should be taken when choosing the number and range of frequencies to ensure that the low frequencies have been sufficiently sampled to correctly model the long term and large scale fluctuations and that enough high frequencies are used to include most of the power in the spectra. The high frequency roll-off begins at a harmonic of the frequency f_0 defined above and is given by (ten Brummelaar 1995)

$$f_{\mathrm{knee}} = (n + 1)f_0. \tag{7}$$

Beyond this the power scales as $f^{-17/6}$.

A set of random coefficients with the power spectra defined by Equation 6 can now be created by generating a set of N random phases θ_{ij} between 0 and 2π for each of the j_{max} coefficients and performing an inverse cosine transform. The jth coefficient at time t will then be given by

$$
\begin{aligned}
a_j(t) &= \sqrt{2f_1} \sum_{i=1}^{N} \sqrt{\Phi'_{a_j}(f, r_0, R, v_\perp)} \, \cos(2\pi f_i t - \theta_{ij}) \\
&= \sqrt{2f_1} \, \frac{\lambda_0}{\lambda} \left(\frac{1}{r_{\lambda_0}}\right)^{5/3} \sum_{i=1}^{N} \sqrt{\Phi'_{a_j}(f, R, v_\perp)} \, \cos(2\pi f_i t - \theta_{ij}).
\end{aligned}
\tag{8}
$$

This series will repeat after $\frac{1}{f_1}$ seconds. It is not strictly necessary to make the time steps even, although for practical purposes we will assume that a calculation is performed for time steps of the size τ which must be less than the coherence time of the atmosphere $\tau_0 \approx r_0/v_\perp$.

Using Equation 4 or 6 with Equation 8 yields a set of Zernike coefficients with the units of phase across a normalized aperture. If we now change the definition of the Zernike polynomial expansion from phase to physical path-length by multiplying Equation 8 by $\frac{\lambda}{2\pi}$ this representation becomes wavelength independent. Furthermore, if the coefficients are calculated for $r_{\lambda_0} = 1$ they can be rescaled at anytime for different seeing conditions. It is then possible to generate as many wavefront sets as desired by using a new set of random phases θ_{ij}.

Once a series of atmospherically perturbed wavefronts has been generated it is then necessary to consider the affect of the optical system itself on the Zernike coefficients. For example, a tip/tilt mirror can be implemented by changing the a_2 and a_3 coefficients and a path-length equalizer need only adjust the a_1 coefficient. In order to model an interferometer, or telescope, the following parameters need to be tracked through the system:

1. The current aperture radius.

2. The total physical path-length a_1. It is often necessary to have several such parameters; $a_{j,\text{vacuum}}$ for vacuum path, $a_{j,\text{air}}$ for air path and $a_{j,\text{glass}}$ for physical path within each glass type. The refractive indices can be taken into account when a wavelength is chosen and after the simulation has been completed.

3. Each optic component's affect on the wavefront shape in terms of the Zernike coefficients a_2 and higher.

This model does not take into account diffraction effects, but as shall be shown below, most other optical aberrations can be easily modeled.

This will work well for single apertures, however, when modeling an array with large baselines one more step is required. A set of power spectra and coefficients for a single large aperture that covers the entire area of the array is used to calculate the atmospherically introduced path-lengths at the positions of the individual apertures. These are then added to the a_1 coefficients for each aperture. Only a small number of Zernike coefficients is required for the large aperture as it is only necessary to model the low frequency and large scale fluctuations. It is for these calculations that it is important to include the outer-scale length parameter L_0 since it is coupled strongly to these large scale effects.

3. OPTICAL SURFACE ABERRATIONS

Each optical surface within the instrument will introduce errors into the wavefront. Since the Zernike coefficients now represent physical path-length it is only necessary to have a Zernike model of each surface and add them to those of the wavefront. The coefficients representing an optical surface can be generated by following a methodology similar to that set out by Noll (1976) for atmospheric turbulence.

Consider an ensemble of surfaces, each with different, and random, optical aberrations. We will assume that these aberrations, on average, have the spatial power spectra $W_s(k)$ which depends only on the spatial frequency magnitude k. The covariance of the Zernike coefficients will then be

$$\langle a_j^* a_{j'} \rangle = \int d\boldsymbol{\rho} d\boldsymbol{\rho'} \, W(\rho) \, W(\rho') \, Z_j(\boldsymbol{\rho}) \, C_s(R\rho, R\rho') \, Z_{j'}(\boldsymbol{\rho'}) \tag{9}$$
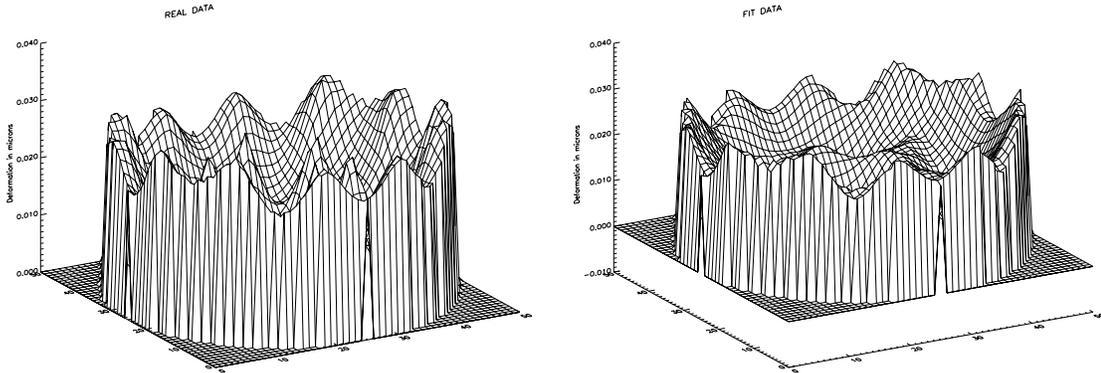
FIGURE 1. Telescope primary mirror deformation as calculated (left) and fitted by 105 Zernike terms (right) for a 1 meter aperture mirror on an Alt/Az mount.

where $C_s(R\rho, R\rho')$ is the covariance function of the optical surface aberrations whose Fourier transform is

$$\Phi_s(k/R, k'/R) = \frac{1}{R^2} W_s(k/R)\, \delta(k - k'). \tag{10}$$

Equation 9 can be written as an integral in Fourier space

$$\langle a_j^* a_{j'} \rangle = \frac{1}{R^2} \int d\mathbf{k} d\mathbf{k}'\, Q_j^*(\mathbf{k})\, W_s(k/R)\, \delta(k - k')\, Q_{j'}(\mathbf{k}') \tag{11}$$

and then calculated to yield

$$\langle a_j^* a_{j'} \rangle = \frac{2}{\pi R^2} \sqrt{(n+1)(n'+1)}\, (-1)^{(n+n'-2m)/2}\, \delta_{mm'} \int_0^\infty W_s(k/R) \frac{J_{n+1}(2\pi k) J_{n'+1}(2\pi k)}{k} dk \tag{12}$$

for $j - j'$ even and zero for $j - j'$ odd.

With Equation 12 it is possible to generate sets of coefficients that display the correct statistics using the technique set out by Roddier (1990). In most cases, however, the cross terms will be small and it is sufficient to use Gaussian random variables with a variance of

$$\sigma_{a_j}^2 = \frac{2(n+1)}{\pi R^2} \int_0^\infty W_s(k/R) \frac{J_{n+1}^2(2\pi k)}{k} dk. \tag{13}$$

An example function for use as the spatial power spectrum is based on the gamma statistical distribution

$$W_s(k) = \frac{\sigma_{\text{rms}}^2}{, (k_0)}\, k^{k_0 - 1}\, e^{-k}. \tag{14}$$

This distribution has no energy at the zero spatial frequency point, a maximum at the frequency k_0, a slow roll-off at frequencies higher than k_0 and a total power, or variance, of σ_{rms}^2. The total power can be set to that of the surface quality, usually expressed in terms of a number of waves at a reference wavelength but easily converted to physical path. The peak spatial frequency must then be chosen, with high peak frequency implying a wide distribution across the spatial frequency scale.

With large mirrors, such as those used in telescopes or beam reducers, other deformations can be introduced due to mirror sag which are not random and can therefore not be modeled in the same way as the surface quality. Mirror deformations like these, however, can be accurately predicted using finite element analysis and then cast in terms of the Zernike coefficients. Figure 1 contains an example of such a fit for a one meter telescope mirror.

4. TIP/TILT SERVO MODEL

The action of a tip/tilt mirror can be implemented by taking into account the current aperture size and changing the a_2 and a_3 coefficients. When modeling a mirror in this way it is important to remember that $Z_2(1,0) = Z_3(1,\frac{\pi}{2}) = 2$ and so only half of the required tilt should be added to the coefficients. The dynamic behavior of the mirror can be modeled by using standard servo analysis techniques (ten Brummelaar & Tango 1994)

The tip/tilt detector could also be modeled in this way although the a_2 and a_3 parameters do not directly correspond to the wavefront tilt. All the Zernike terms for which $m = 1$ contribute to wavefront tilt and the average phase gradient in physical path per normalized aperture radius in the direction ϵ is given by (ten Brummelaar 1995)

$$\Theta_{\text{tilt}}(\epsilon) = \sum_{m=1} a_j(t) Z_j(1,\epsilon) \, \text{path/radius}. \tag{15}$$

This can be used to represent the signal of a quadrant detector.

The expression for the error associated with a quadrant detector at the measurement wavelength λ is (Tyler & Fried 1982)

$$\sigma_{\Theta_{\text{tilt}}} = \pi \left[\left(\frac{3}{32} \right)^2 + \left(\frac{o}{16} \right)^2 \right]^{\frac{1}{2}} \frac{\left(\frac{\lambda}{R} \right)}{\text{SNR}} \, \text{rad}. \tag{16}$$

Here o is the angular subtense of the object divided by the diffraction angle $(\lambda/2R)$ of the optical system and SNR is the signal to noise ratio of the four detectors summed to act as a single detector. Few target objects will be resolved by a single aperture and so we can say that $o \ll 1$. Furthermore, the signal to noise ratio of the four detectors summed will be primarily determined by the Poisson statistics of the photon events so we can write the error in angular position measurement of the quadrant detectors as

$$\sigma_{\Theta_{\text{tilt}}} = \frac{3\pi}{32} \frac{\lambda}{R} \frac{1}{\sqrt{N_{\text{ph}}}} \, \text{rad} \tag{17}$$

where N_{ph} is the total number of counts received in all four quadrants. This expression needs to be re-cast into physical path per normalized aperture radius to become

$$\sigma_{\Theta_{\text{tilt}}} = \frac{3\pi}{32} \frac{\lambda}{\sqrt{N_{\text{ph}}}} \, \text{path/radius}. \tag{18}$$

Both Equations 15 and 18 are independent of aperture radius, however, the aperture radius will have to be taken into account when modeling the dynamics of the moving mirror itself.

Due to the greater number of parameters involved, a generalized model is less tractable for an optical path-length equalizer (OPLE). Like the tip/tilt system, however, it is possible to

model an OPLE by varying just a few Zernike coefficients. Moving the OPLE is the same as changing either the $a_{1,\text{vacuum}}$ or $a_{1,\text{air}}$ coefficients depending on whether the OPLE cart is within a vacuum chamber or not. The dynamical behavior is strongly coupled to the mechanical design but a simple lag filter with an e^{-1} response of less than a millisecond is a good approximation (Shao 1994). Other optical effects such as wavefront tilt due to misalignment of the cart or defocus caused by movement of the secondary in a cat's eye can be modeled by adjusting the a_2, a_3, and a_4 coefficients. Fringe tracking detectors need to modeled once a wave-band has been chosen and fringes calculated as set out in the following section.

5.   STREHL RATIOS AND FRINGES

Once two or more beams have been tracked through the atmosphere and the optical system a wave-length, optical bandwidth and pixilation size must be chosen. The values of the Zernike modes used at each pixel location can be calculated once, using Equation 50 and stored for later use. With this done, visibilities and Strehls can be worked out for various seeing conditions and wavelengths without the need to re-run most of the simulation.

The Strehl ratio of a single beam is defined to be the intensity in the center of the aberrated image divided by the central intensity of an Airy disk. Ignoring scintillation effects, we can write the wavefront across the aperture as $\sqrt{I_0}e^{i\varphi(R\rho,\theta)}$ where I_0 is the classical intensity. The Strehl ratio will then be

$$
\begin{aligned}
S &= \left| \int d\boldsymbol{\rho} W(\rho) e^{i\varphi(R\rho,\theta)} \right|^2 \\
&= \left[\int d\boldsymbol{\rho} W(\rho)\cos(\varphi(R\rho,\theta)) \right]^2 + \left[\int d\boldsymbol{\rho} W(\rho)\sin(\varphi(R\rho,\theta)) \right]^2 ,
\end{aligned}
\tag{19}
$$

where the wavefront phase is now given by

$$
\varphi(R\rho,\theta) = \frac{2\pi}{\lambda} \sum_{j=2}^{j_{\max}} a_j Z_j(\rho,\theta) \ .
\tag{20}
$$

Note that we do not include the first Zernike mode of piston phase as it represents the mean phase across the aperture and does not contribute to Strehl.

It is also possible to approximate Strehl using the relation

$$
S \approx e^{-\sigma_\varphi^2}
\tag{21}
$$

where σ_φ^2 is the phase variance across the aperture. This can be calculated using

$$
\begin{aligned}
\sigma_\varphi^2 &= \int d\boldsymbol{\rho} W(\rho)\varphi^2(R\rho,\theta) \\
&= \frac{4\pi^2}{\lambda^2} \int d\boldsymbol{\rho} W(\rho) \left(\sum_{j=2}^{j_{\max}} a_j Z_j(\rho,\theta) \right)^2 \\
&= \frac{4\pi^2}{\lambda^2} \int d\boldsymbol{\rho} W(\rho) \left[\sum_{j=2}^{j_{\max}} a_j^2 Z_j^2(\rho,\theta) + 2\sum_{j\neq j'} a_j a_{j'} Z_j(\rho,\theta) Z_{j'}(\rho,\theta) \right]
\end{aligned}
$$

$$= \frac{4\pi^2}{\lambda^2} \sum_{j=2}^{j_{\max}} a_j^2 \tag{22}$$

due to the orthogonality of the Zernike polynomials. Calculating the phase variance then becomes a matter of squaring and summing up the Zernike terms and converting to the correct dimensions of radians squared. The Strehl can then calculated for each frame using

$$S \approx \exp\left(\frac{4\pi^2}{\lambda^2} \sum_{j=2}^{j_{\max}} a_j^2 \right). \tag{23}$$

The long term average of the Strehl will be

$$\langle S \rangle \approx e^{\langle -\sigma_\varphi^2 \rangle} = \exp\left(\frac{4\pi^2}{\lambda^2} \sum_{j=2}^{j_{\max}} \langle a_j^2 \rangle \right). \tag{24}$$

Expressions for the coefficient variances $\langle a_j^2 \rangle$ for atmospheric effects alone can be found in the work of Noll (1976).

The calculation of visibility works in a similar fashion. The classical intensity of two beams combined in the aperture plane at the wavenumber $\nu = \frac{1}{\lambda}$ and an optical bandwidth of $\Delta\nu$ is

$$I(\rho, \theta) = 1 + V \frac{\sin(\pi\,\mathrm{OPD}(R\rho,\theta)\,\Delta\nu)}{\pi\,\mathrm{OPD}(R\rho,\theta)\,\Delta\nu}\,\cos(2\pi\,\nu\,\mathrm{OPD}(R\rho,\theta) + \phi_V) \tag{25}$$

where V is the visibility magnitude, ϕ_V is the visibility phase and $\mathrm{OPD}(R\rho,\theta)$ is the optical path-length. The OPD can be calculated using the a_1 parameters for the vacuum, glass and air paths modified by the appropriate refraction coefficients:

$$\mathrm{OPD}(R\rho,\theta) = a_{1,\mathrm{vacuum}} + n_{\mathrm{air}}a_{1,\mathrm{air}} + n_{\mathrm{glass}}a_{1,\mathrm{glass}} + \sum_{j=2}^{j_{\max}} a_j Z_j(\rho,\theta). \tag{26}$$

The total classical intensity can then be used as a probability density function to generate random photon events using Poisson statistics. The measured visibility magnitude and phase can be calculated using the techniques described by Tango & Twiss (1980) and fringe tracking algorithms such as group delay tracking (Lawson 1994) can be implemented. Other measurement systems can be modeled in a similar manner. Figure 2 contains some examples of simulated fringes using 105 Zernike modes, 10 cm seeing, an outer scale length of 200 m and at a wavelength of 500 nm. Each frame contains 128×128 pixels, is 5ms apart and the wind speed was set to $10\,\mathrm{ms}^{-1}$.

Like Strehl ratio, visibility magnitude can also be approximated using the phase variance defined in Equation 22 and the approximation to visibility given by Tango & Twiss (1980)

$$\eta^2 = 1 - 2\sigma_\varphi^2 = 1 - \frac{8\pi^2}{\lambda^2} \sum_{j=2}^{j_{\max}} a_j^2 \tag{27}$$

where η is the visibility transfer factor, that is, the measured visibility is $V_{\mathrm{meas}} = \eta\,V$.

Figure 3 shows an example of the results of a series of simulations for various D/r_0 values. The solid lines represent the visibility and Strehl predicted by Tango & Twiss (1980), the
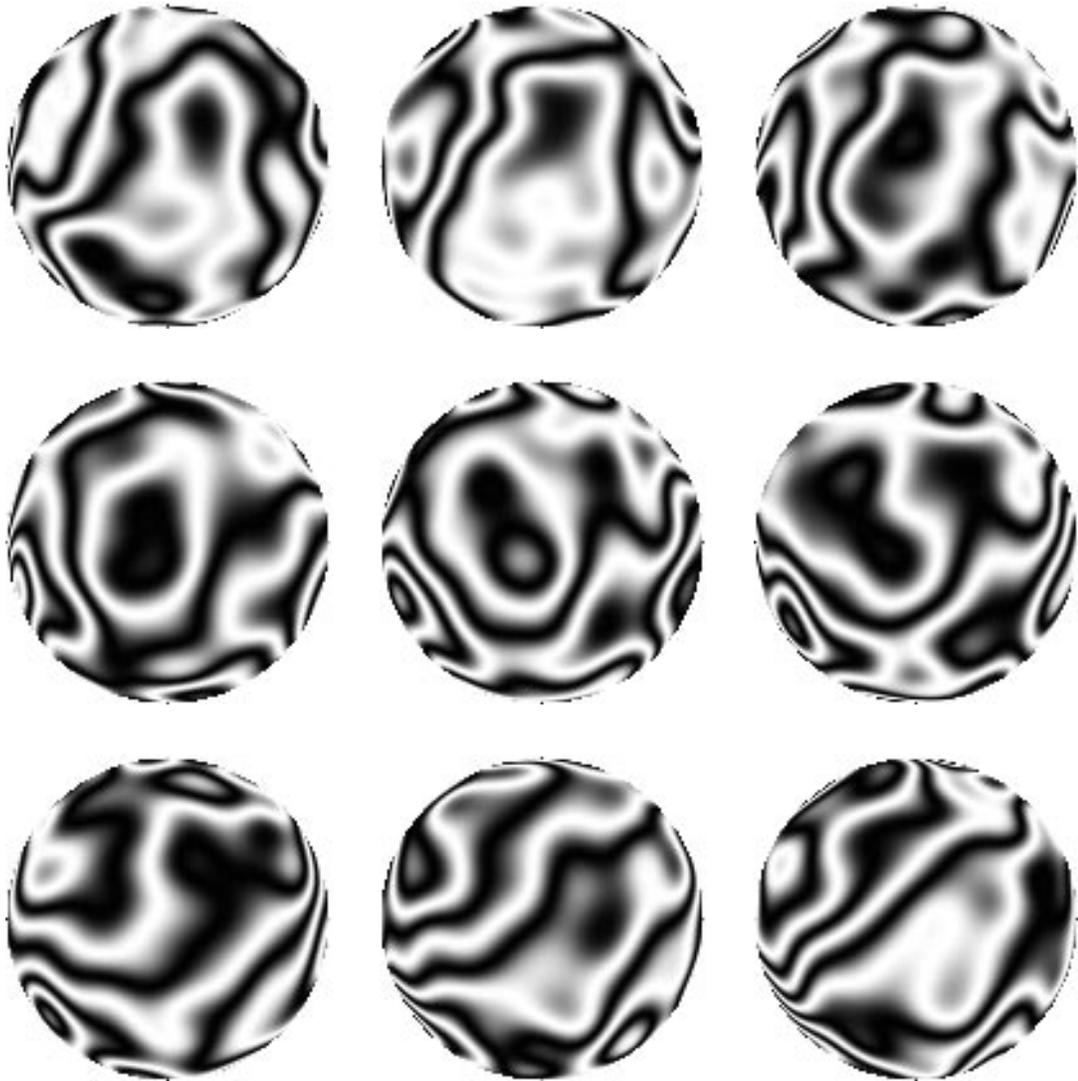
FIGURE 2. Results of a simulation run using the first 105 Zernike polynomials. Each frame represents a sample 5 milliseconds apart from top left to bottom right for two 1 meter apertures 100 meters apart. The seeing was set to r_0=10cm and L_0=200m and the calculations are for monochromatic light at a wavelength of 0.5 microns.
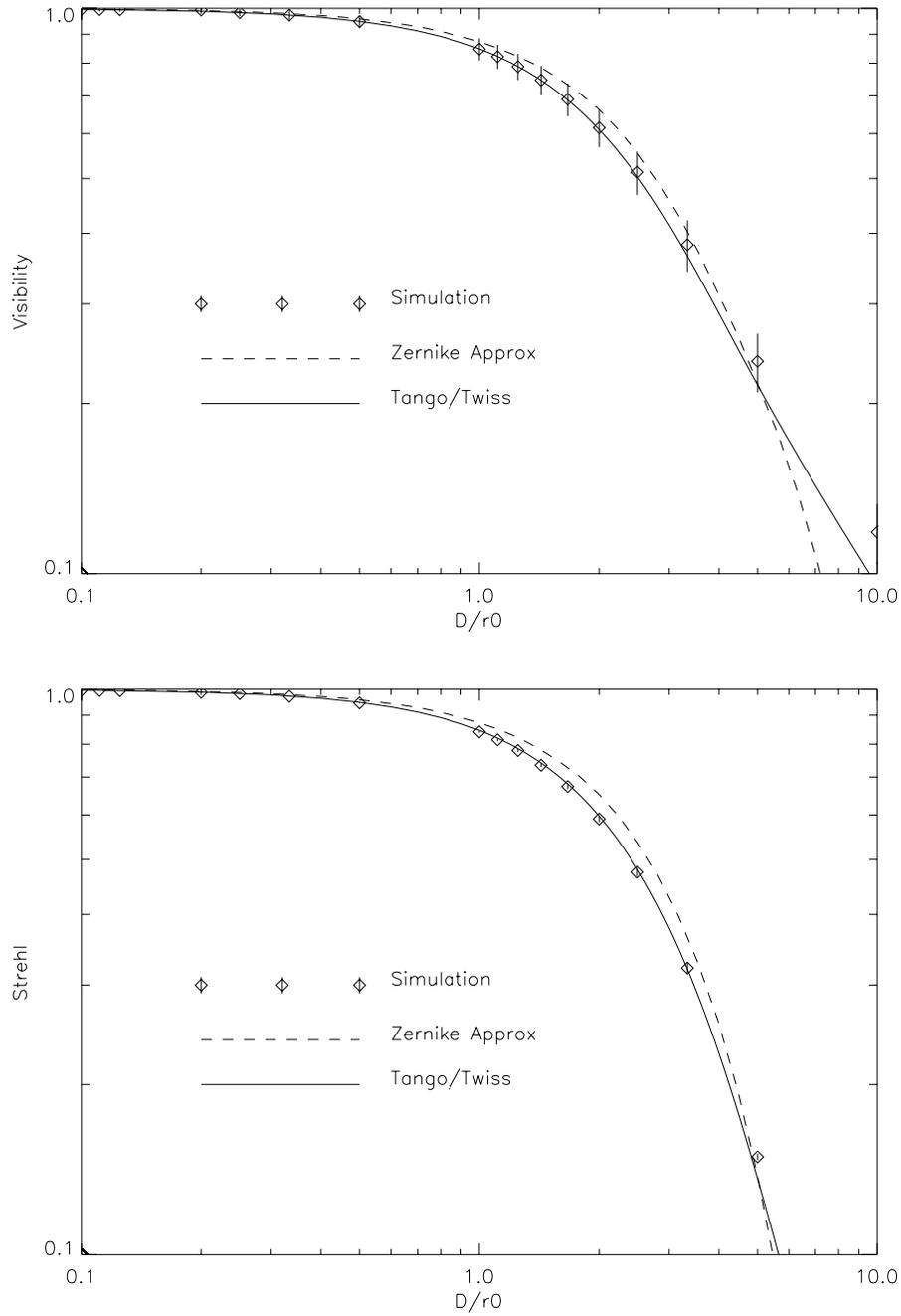
FIGURE 3. Simulation results compared to the Zernike approximation set out in the text and the analytical formulation by Tango & Twiss (1980). The top plot shows how the visibility changes with D/r_0 and the bottom plot shows Strehl ratio changes with D/r_0. The simulations were performed using 105 Zernike terms, a wind speed of 10 ms^{-1} and an infinite outer-scale length. The a tip/tilt system was modeled based on the SUSI device (ten Brummelaar & Tango 1994) and applied to the wavefront data before Strehl ratios or visibilities were measured.

dashed line is the approximations given in Equations 23 and 27 while the points represent full simulations using Equations 19 and 25. Once again 105 Zernike modes were used in a 128×128 pixel aperture, a wind speed of $10\,\mathrm{ms}^{-1}$ and a monochromatic light at $500\,\mathrm{nm}$. Each simulation was for 5 seconds using $1\,\mathrm{ms}$ samples. The tip/tilt servo model was based upon the system build for the Sydney University Stellar Interferometer (ten Brummelaar & Tango 1994) and the outer scale length was set to infinity. The error bars represent the standard deviation of the Strehl and visibility measurements.

6. USING MULTIPLE SUB-APERTURES

Choosing the optimum aperture size to use is a compromise between photon statistics and atmospheric turbulence. If the aperture is too small too few photons are available for good signal to noise, while if the aperture is too large wavefront distortions will reduce the visibility and signal to noise ratio. Buscher (1988) has shown that for a noise limited detector the optimum aperture diameter is approximately $3r_0$. There are two ways of taking advantage of apertures larger than this:

- Correct the tip/tilt across the entire aperture and then divide it into many sub-apertures.

- Employ high order adaptive optics (AO) across the entire aperture. The simplest form of this would be to have a tip/tilt servo for each sub-aperture.

The former will be less costly while the later may have the greatest potential benefit for very large apertures.

While it is possible to calculate sub-apertures directly from the Zernike coefficients it is sometimes more useful to transform to a new aperture with a new set of coefficients. The first step in this process is to show that the statistics of the Zernike coefficients are independent of an aperture rotation, thereby decreasing the number of parameters required. We begin with an aperture A with radius R whose Zernike coefficients are a_j and rotate the wavefront $\varphi_A(R\rho', \theta')$ by an angle α to yield a new aperture B with wavefront $\varphi_B(R\rho, \theta)$ whose Zernike coefficients are b_j. The wavefront shape across the aperture is then

$$\varphi_B(R\rho,\theta) = \varphi_A(R\rho',\theta'+\alpha) = \sum_{j=1}^{j_{\max}} a_j Z_j(\rho',\theta'+\alpha) = \sum_{j=1}^{j_{\max}} b_j Z_j(\rho,\theta). \qquad (28)$$

Therefore, using Equation 49 the new coordinate system b_j is related to the old coordinate system a_j via

$$
\begin{aligned}
b_j &= \int d\boldsymbol{\rho}\, W(\rho)\, \varphi_B(R\rho,\theta)\, Z_j(\rho,\theta) \\
&= \int d\boldsymbol{\rho}\, W(\rho) \left(\sum_{j'=1}^{j_{\max}} a_{j'} Z_{j'}(\rho,\theta+\alpha) \right) Z_j(\rho,\theta) \\
&= \sum_{j'=1}^{j_{\max}} a_{j'} I_{1jj'}
\end{aligned}
\qquad (29)
$$

where

$$I_{1jj'} = \int d\boldsymbol{\rho}\, W(\rho)\, Z_{j'}(\rho, \theta + \alpha) Z_j(\rho, \theta). \tag{30}$$

Using the orthogonality properties of the Zernike polynomials it can be shown that

$$I_{1jj'} = \begin{cases} 0 & \text{for } m \neq m' \text{ or } n \neq n', \\ 1 & \text{for } n = n' \text{ and } m = m' = 0, \\ \cos m\alpha & \text{for } n = n',\ m = m' \neq 0 \text{ and both } j \text{ and } j' \text{ odd or even}, \\ \sin m\alpha & \text{for } n = n',\ m = m' \neq 0,\ j \text{ even and } j' \text{ odd}, \\ -\sin m\alpha & \text{for } n = n',\ m = m' \neq 0,\ j \text{ odd and } j' \text{ even}. \end{cases} \tag{31}$$

The only modes of importance are those for which $n = n'$ and $m = m'$ and we need consider only three cases: $m = 0$, $m \neq 0$ and j even, and $m \neq 0$ and j odd. Combining Equations 29 and 31 results in

$$b_{j,m=0} = a_{j',m=0}, \tag{32}$$
$$b_{j,m\neq0,j\text{ even}} = a_{j'\text{ even}} \cos m\alpha + a_{j'\text{ odd}} \sin m\alpha \text{ and} \tag{33}$$
$$b_{j,m\neq0,j\text{ odd}} = a_{j'\text{ odd}} \cos m\alpha - a_{j'\text{ even}} \sin m\alpha. \tag{34}$$

Combining this with the results of Noll (1976) then gives

$$\langle b_j^2 \rangle = \langle a_j^2 \rangle. \tag{35}$$

The variance of the coefficients in the rotated aperture are therefore the same as those of the original aperture and it will only be necessary to consider radius change and position change in a single direction when studying a sub-aperture.

The sub-aperture will have a radius fR where $0 < f < 1$ and will lie a distance ρ_0 from the aperture center along the X axis. Once again the original aperture will have the wavefront $\varphi_A(R\rho', \theta')$ and the coefficients a_j while the sub-aperture has the wavefront $\varphi_B(fR\rho, \theta)$ and the coefficients b_j. The coordinate systems are displayed in Figure 4 and are related via the equations

$$\sin \theta' = \frac{f\rho \sin \theta}{\rho'},$$
$$\cos \theta' = \frac{\rho_0 + f\rho \cos \theta}{\rho'} \text{ and}$$
$$\rho'^2 = \rho_0^2 + f^2\rho^2 + 2\rho\rho' f \cos \theta. \tag{36}$$

The new coefficients b_j are related to the old coefficients $a_{j'}$ via

$$b_j = \sum_{j'=1}^{j_{\max}} a_{j'} \int d\boldsymbol{\rho} W(\rho) Z_j(\rho, \theta) Z_{j'}(\rho', \theta') = \sum_{j'=1}^{j_{\max}} a_{j'} I_{2jj'}. \tag{37}$$

If no tip/tilt correction is made the variances of the coefficients in the new aperture will depend only on the current seeing conditions so scale with radius to the five thirds power

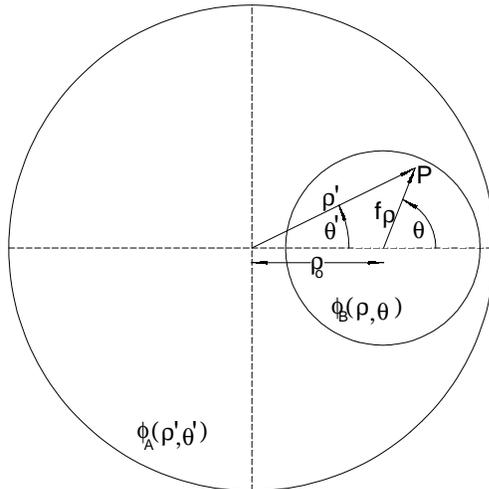$$\langle b_j^2 \rangle = f^{5/3} \langle a_j^2 \rangle. \tag{38}$$

FIGURE 4. Definition of the axes for creating a sub-aperture. The primary aperture has the phase profile $\Phi_A(\rho', \theta')$ and a radius of R. The sub-aperture lies along the x axis at a distance ρ_0 from the origin and has a radius of fR and the phase profile $\Phi_B(\rho, \theta)$.

In most optical arrays, whether using AO or not, the wavefront tilt across the entire aperture will be servo-ed out by the tip/tilt system and the piston terms removed by the OPLEs. Thus we need to know how a reduction in the variances $\langle a_1^2 \rangle$, $\langle a_2^2 \rangle$ and $\langle a_3^2 \rangle$ affect the variances of the b_j coefficients. The contribution to $\langle b_j^2 \rangle$ by the power in the j'th mode of the original aperture is given by

$$\text{VAR}(b_j, a_{j'}) = \langle a_{j'}^2 \rangle I_{2jj'} + 2 \sum_{k=1}^{j_{\max}} \langle a_k a_{j'} \rangle I_{2jk} I_{2jj'}. \tag{39}$$

Using Equations 36 and 37 it can be shown that

$$I_{2j1} = \begin{cases} 1 & \text{for } j = 1 \\ 0 & \text{otherwise} \end{cases}, \tag{40}$$

$$I_{2j2} = \begin{cases} 2\rho_0 & \text{for } j = 1 \\ f & \text{for } j = 2 \\ 0 & \text{otherwise} \end{cases} \tag{41}$$

and

$$I_{2j3} = \begin{cases} f & \text{for } j = 3 \\ 0 & \text{otherwise} \end{cases}. \tag{42}$$

With Equations 38 through 42 it is now possible to write the mode variances in the sub-aperture, given that the piston and tilt terms have been removed from the main aperture

$$\langle b_1^2 \rangle = f^{5/3} \langle a_1^2 \rangle - 4\rho_0^2 \langle a_2^2 \rangle - 4\sqrt{2}\rho_0(3\rho_0^3 - 2\rho_0 + 3f^2\rho_0)\langle a_2 a_8 \rangle \tag{43}$$

$$\langle b_2^2 \rangle = (f^{5/3} - f^2)\langle a_2^2 \rangle - \sqrt{2}(9\rho_0^2 f^2 + 2f^4 - 2f^2)\langle a_2 a_8 \rangle \tag{44}$$

$$\langle b_3^2 \rangle = (f^{5/3} - f^2)\langle a_3^2 \rangle - \sqrt{2}(3\rho_0^2 f^2 + 2f^4 - 2f^2)\langle a_3 a_7 \rangle \tag{45}$$
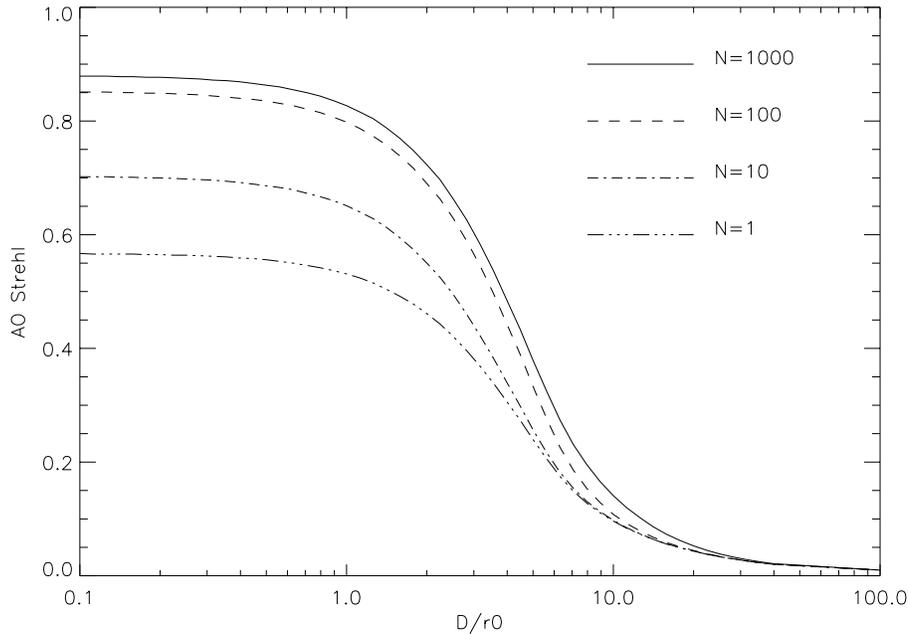
FIGURE 5. The Strehl ratio required of an adaptive optics system to match the signal to noise performance of using seven optimized sub-apertures for a range of seeing conditions and photon fluxes.

$$\langle b_{j>3}^2 \rangle \quad = \quad f^{5/3} \langle a_j^2 \rangle \tag{46}$$

and the total phase variance within the sub-aperture

$$\sigma_{\varphi_B}^2 = f^{5/3} \sigma_{\varphi_A}^2 - \frac{4\pi^2}{\lambda^2} \left(2f^2 \langle a_2^2 \rangle - \sqrt{2}(12\rho_o^2 f^2 + 4f^4 - 4f^2) \langle a_2 a_8 \rangle \right) . \tag{47}$$

Using the approximate forms for Strehl and visibility from section 5 and Equation 47 it is possible to investigate the performance of an AO system as compared to using multiple sub-apertures. An example is given in Figure 5 for a one meter aperture where the Strehl ratio required of an AO to match the signal to noise performance of seven sub-apertures has been plotted for various seeing conditions and photon fluxes. The signal to noise expression from Tango & Twiss (1980) has been used and it has been assumed that the AO system requires 50% of the photon flux for the wavefront sensor. The approximation that Strehl and visibility have similar values has also been employed (ten Brummelaar, Bagnuolo, & Ridgway 1995).

During times of good seeing ($r_0 > 10$cm) an AO system would need to achieve Strehl ratios of 0.15 or better in order to out-perform seven sub-apertures. Apart from in the infra-red, this is a tall order for an AO system and it is doubtful that it would be worth the expense. During times of bad seeing an AO system need not achieve very high Strehls, although, in the worst seeing conditions it is not clear that an AO system would work at all. These results will scale with aperture, although for large apertures many more sub-apertures are possible.

7. CONCLUSION

In the absence of an analytical solution computer simulations are a powerful method of investigating the behavior of complex astronomical optical systems. Unfortunately these simulations are often computationally expensive and take a long time to calculate. Using the polynomials of Zernike to represent the physical path of the wavefront in each aperture, rather than phase, it is possible to perform most of these simulations without the need to choose a wavelength or pixilation scale. Once a series of Zernike coefficients has been generated one need only modify the appropriate modes, rather than perform a calculation on each pixel, when modeling the optical system. The affects of each optic, including surface defects and mirror sag, are easily modeled in this way as is a tip/tilt servo and path-length equalizer. Visibility and Strehl ratio can then be calculated directly, after choosing a wavelength and pixilation scale, or approximations based on the Zernike coefficients can be used.

8. REFERENCES

Born, M. & Wolf, E. 1989, *Principles of Optics* (Pergamon Press, Elmsford, New York)

Buscher, D. 1988, "Optimizing a ground-based optical interferometer for sensitivity at low light levels," MNRAS, **235**, 1203-1226

ten Brummelaar, T.A. 1995, "The Contribution of High Order Zernike Modes to Wavefront Tilt.," Optics Comm., **115**, 417-424

ten Brummelaar, T.A. & Tango, W.J. 1994, "A Wavefront Tilt Correction Servo for the Sydney University Stellar Interferometer," Experimental Astronomy, **4**, 297-315

ten Brummelaar,T.A., Bagnuolo, W.G. Jr., & Ridgway, S.T. 1995, "Strehl Ratio and Coherence Loss in Long Baseline Stellar Interferometry," Optics Letters, **20**, 521-523

Davis, J.D., Lawson, P.R., Booth, A.J., Tango, W.J., & Thorvaldson, E.D. 1995, "Atmospheric path variations for baselines up to 80m measured with the Sydney University Stellar Interferometer',' MNRAS, **273**, L53-L58

Fried, D.L. 1965, "Statistics of a Geometric Representation of Wavefront Distortion," JOSA., **55**, 1427-1435

Hogge, C.B. & Butts, R.R. 1976, "Frequency Spectra for the Geometric Representation of Wavefront Distortions Due to Atmospheric Turbulence," IEEE Trans. Antennas Propagat., **AP-24**, 144-154

Hu, P.H., Stone, J., & Stanley, T. 1989, "Application of Zernike Polynomials to Atmospheric Propagation Problems," JOSA-A, **6**, 1595-1608

Lawson, P.R. 1994, "Group delay tracking in optical stellar interferometry using the Fast Fourier Transform," JOSA-A, **12**, 366-374

Noll, R.J. 1976, "Zernike Polynomials and Atmospheric Turbulence," JOSA, **66**, 207-211

Roddier, F., Northcott, M.J., Graves, J.E., & McKenna, D.L. 1993, "One-Dimensional Spectra of Turbulence-Induced Zernike Aberrations: Time-Delay and Isoplanicity Error in Partial Adaptive Compensation," JOSA-A, **10**, No. 5, 957-965

Roddier, N. 1990, "Atmospheric Wavefront Simulation and Zernike Polynomials," SPIE Amplitude and Intensity Spatial Interferometry, **1237**, 668 - 679

Shao, M. 1994, JPL, private communication

Tango, W.J. & Twiss, R.Q. 1980, "Michelson Stellar Interferometry," Progress in Optics, **XVII**, 239-277

Tyler, G.A. & Fried, D.L. 1982, "Image-Position Error Associated with a Quadrant Detector," JOSA, **72**, 804-808

Wang, J.Y. & Silva, D.E. 1980, "Wave-Front Interpretation with Zernike Polynomials," Applied Optics, **19**, No. 9, 1510-1518

A. ZERNIKE POLYNOMIAL DEFINITION

Using polar coordinates ρ and θ, normalized for an unobscured aperture of radius R, the phase of the wavefront across the aperture can be written

$$\varphi(R\rho,\theta) = \sum_j a_j Z_j(\rho,\theta) \tag{48}$$

where a_j are the expansion coefficients given by

$$a_j = \int d\boldsymbol{\rho}\, W(\rho)\, \varphi(R\rho,\theta)\, Z_j(\rho,\theta). \tag{49}$$

The Zernike polynomials themselves are given by

$$Z_j(\rho,\theta) = \sqrt{n+1}\, R_n^m(\rho) \times \begin{cases} \sqrt{2}\cos m\theta & m \neq 0, \quad j \text{ even} \\ \sqrt{2}\sin m\theta & m \neq 0, \quad j \text{ odd} \\ 1 & m = 0 \end{cases} \tag{50}$$

where

$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s!\,[(n+m)/2 - s]!\,[(n-m)/2 - s]!}\, \rho^{n-2s}. \tag{51}$$

The constants m and n are integers such that $m \leq n$ and $n - |m|$ is even. The index j is used to order the modes.

The weighting function

$$W(\rho) = \begin{cases} 1/\pi & \rho \leq 1 \\ 0 & \rho > 1 \end{cases} \tag{52}$$

is added so that the integrals can be taken over all space.

Zernike polynomials follow the orthogonality relation

$$\int d\mathbf{r}\, W(r) Z_j(\mathbf{r}) Z_{j'}(\mathbf{r}) = \delta_{jj'}. \tag{53}$$

and have the Fourier transform

$$Q_j(k,\phi) = \sqrt{n+1}\, \frac{J_{n+1}(2\pi k)}{\pi k} \times \begin{cases} (-1)^{(n-m)/2}\, i^m\, \sqrt{2}\cos m\phi & m \neq 0, \text{even } j \\ (-1)^{(n-m)/2}\, i^m\, \sqrt{2}\sin m\phi & m \neq 0, \text{odd } j \\ (-1)^{n/2} & m = 0 \end{cases}. \tag{54}$$