Wave-front measurement errors from restricted concentric subdomains

Kenneth A. Goldberg and Kevin Geary

Center for X-Ray Optics, Lawrence Berkeley National Laboratory, Berkeley, California 94720

Received December 21, 2000; accepted February 21, 2001

In interferometry and optical testing, system wave-front measurements that are analyzed on a restricted subdomain of the full pupil can include predictable systematic errors. In nearly all cases, the measured rms wave-front error and the magnitudes of the individual aberration polynomial coefficients underestimate the wave-front error magnitudes present in the full-pupil domain. We present an analytic method to determine the relationships between the coefficients of aberration polynomials defined on the full-pupil domain and those defined on a restricted concentric subdomain. In this way, systematic wave-front measurement errors introduced by subregion selection are investigated. Using vector and matrix representations for the wave-front aberration coefficients, we generalize the method to the study of arbitrary input wave fronts and subdomain sizes. While wave-front measurements on a restricted subdomain are insufficient for predicting the wave front of the full-pupil domain, studying the relationship between known full-pupil wave fronts and subdomain wave fronts allows us to set subdomain size limits for arbitrary measurement fidelity. © 2001 Optical Society of America

OCIS codes: 220.1010, 000.3870, 220.4840, 120.3180.

1. INTRODUCTION

In interferometric testing of optical systems, a primary goal is to measure, with the highest possible accuracy, the wave front produced by a test optic. Typically, an interferogram fringe pattern or a series of patterns is projected onto a CCD camera, and one of a host of fringe pattern analysis methods is used to recover the underlying wave-front phase across the measurement area. Wave-front analysis with Zernike polynomials[1–3] requires the definition of a unit-circle coordinate system on a subdomain of the total collected data. Based on the pixels that define the interferogram image, the chosen subdomain is actually a discrete approximation to the unit circle on which the continuous Zernike polynomials are defined.

The selection of an appropriate subdomain is often a subjective procedure that excludes some of the data at the outermost edges of the interferogram. A low signal-to-noise ratio or diffraction from the aperture edges often makes the inclusion of the edge data problematic and thus motivates the selection of a restricted domain size. Yet the exclusion of any data from the full pupil reduces the effective numerical aperture of the measurement and changes the measured wave-front result, often reducing the apparent magnitude of the measured aberrations.

Estimation of the full pupil's wave front based on subdomain measurement is problematic for two main reasons. Foremost is the fact that the behavior of the wave front in the excluded regions is not measured; therefore extrapolation from the measured region introduces high uncertainty. Second, if the aberration polynomials (i.e., the basis set of functions that define individual component aberrations) used in the wave-front fitting are orthogonal on the full domain, yet the data used for the fit come only from a restricted domain, then the fitted values of the aberration polynomial coefficients may be unstable.[1] This result comes primarily from the fact that the full pupil's aberration polynomials are generally not orthogonal on the restricted domain.

When aberration polynomials are not orthogonal on the measurement domain, fitting can become sensitive to small errors in the input. This result is well-known from least-squares error-minimization methods of polynomial fitting.[4,5] (Some authors have shown that the numerical inversion in the least-squares method is not unstable, as is traditionally thought.[2,6] However, for the higher-ordered polynomial terms, the limitations of discrete sampling become especially problematic.) For this reason, in wave-front fitting, it is common first to define an intermediate set of orthogonal aberration polynomials that are appropriate to the discrete subdomain of measurement, then to fit the wave front on that domain by orthogonal projection, and last to transform the resultant coefficients to the desired basis.[1] While our goal is to measure the wave front on the full-pupil domain, since the data are defined only on the subdomain, we must first perform the fitting by using orthogonal basis functions that are appropriate for the subdomain.

While wave-front measurements on a restricted subdomain are generally insufficient for predicting the wave front of the full-pupil domain, studying the relationship between known full-pupil wave fronts and subdomain wave fronts allows us to set a minimum size limit for subdomains to achieve arbitrary measurement fidelity. Quantifying these measurement errors leads to the definition of measurement tolerances on the capture size and on the displacement of the subdomain.

Measurement errors and uncertainties of this sort are always proportional to magnitudes of the input wave-front aberrations. Therefore the tolerances that we set are defined relative to the wave-front aberration magni-

tudes: In absolute terms, measurement errors are always reduced as the wave-front error tends to 0.

In this analysis, we derive the effect of the restricted subdomain size on the measurement fidelity for each individual Zernike polynomial term and then generalize the results by linear superposition to fit arbitrary wave fronts. Vector and matrix notation is used to simplify the description and the calculations. The method is applied to circular pupils, for which the well-known Zernike polynomials are an appropriate orthogonal basis for the definition of wave-front aberrations. A similar derivation may be performed for arbitrary pupil shapes.

The treatment of wave-front fitting on nonconcentric, or displaced, subdomains will be the subject of a separate paper by the authors.

2. DEFINITION OF ZERNIKE POLYNOMIALS

We begin with a strictly real definition of the Zernike polynomials $U_n^m(r, \theta)$, functions of the polar coordinates r and θ. The indices m and n describe the azimuthal and radial order of the polynomials.

$$U_n^m(r, \theta) = \begin{cases} R_n^m(r)\cos m\theta, & m \leqslant 0 \\ R_n^m(r)\sin m\theta, & m > 0 \end{cases}, \tag{1}$$

with

$$R_n^m(r) \equiv \sum_{s=0}^{n'} (-1)^s \binom{n'}{s} \binom{n-s}{n'} r^{n-2s}$$

$$= \sum_{s=0}^{n'} (-1)^s \frac{(n-s)!}{s!\left(\frac{n-m}{2}-s\right)!\left(\frac{n+m}{2}-s\right)!} r^{n-2s}, \tag{2}$$

where

$$n' \equiv \frac{n-|m|}{2} = \begin{cases} \frac{n-m}{2}, & m > 0 \\ \frac{n+m}{2}, & m \leqslant 0 \end{cases}. \tag{3}$$

The binomial coefficient is defined in the standard way:

$$\binom{a}{b} \equiv \frac{a!}{b!(a-b)!}. \tag{4}$$

For these definitions, the following conditions hold: $n \geqslant |m| \geqslant 0$, and $n - m$ is even. In later equations, the functional dependence of U_n^m on r and θ is implicit; R_n^m is a function of r alone.

It is important to note that the polynomial magnitudes used here are *not* scaled to normalize the rms magnitude of the individual terms, a convention commonly used elsewhere. Rather, the polynomials are bounded on the range $[-1, 1]$, and this is referred to as unit magnitude.

The expression(s) for the radial component of the Zernike polynomials can be simplified by using a series coefficient $C_n^m(s)$, defined as

$$C_n^m(s) \equiv (-1)^s \binom{n'}{s} \binom{n-s}{n'}$$

$$= (-1)^s \binom{\frac{n-|m|}{2}}{s} \binom{n-s}{\frac{n-|m|}{2}}$$

$$= (-1)^s \frac{(n-s)!}{s!(n'-s)!(n-s-n')!}$$

$$= (-1)^s \frac{(n-s)!}{s!\left(\frac{n-|m|}{2}-s\right)!\left(\frac{n+|m|}{2}-s\right)!}$$

$$= (-1)^s \frac{(n-s)!}{s!\left(\frac{n-m}{2}-s\right)!\left(\frac{n+m}{2}-s\right)!}. \tag{5}$$

Note that $C_n^m = C_n^{-m}$. The radial function can now be written concisely as

$$R_n^m(r) = \sum_{s=0}^{n'} C_n^m(s) r^{n-2s}. \tag{6}$$

3. WAVE-FRONT FITTING ON THE SUBDOMAIN

Consider a centered circular subdomain of the unit circle on which the Zernike polynomials are defined; let this subdomain have a radius p such that $0 < p \leqslant 1$. This is shown in Fig. 1. Given an input wave front on the full unit circle, our goal is to find a general expression for the wave front defined on the subdomain with the new unit radius scaled to the radius p. We can first find a solution for the subdomain Zernike polynomial series that fits a single arbitrary input Zernike polynomial term on the full domain. Then, by linear superposition, we generalize the result for an arbitrary input wave front, represented as a Zernike polynomial series. To avoid confusion, we will refer to the polynomials on the full domain as the "input" and on the restricted subdomain as the "fit."

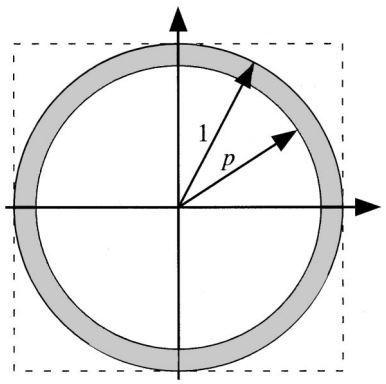

Fig. 1. Centered circular subdomain of the unit circle, with radius p.

A. Fitting a Single Input Zernike Polynomial

Consider a single arbitrary input Zernike polynomial term U_n^m, defined on the full unit circle. Let r' be the radius defined on the full domain. Making the substitution $r \equiv r'/p$ scales the radius on the subdomain, r, to reach unity at the subdomain edge. In this way, $r' = pr$. The expression for the radial dependence of the input wave front across the subdomain is

$$R_n^m(pr) = \sum_{s=0}^{n'} C_n^m(s)(pr)^{n-2s} = \sum_{s=0}^{n'} C_n^m(s)p^{n-2s}r^{n-2s}. \tag{7}$$

Orthogonality dictates that the polynomial terms in the fit wave front must match the azimuthal dependence of the input wave front exactly: The fit terms all have the same m value as the input wave front. (This point is easily understood by separating azimuthal components of the aberration polynomials from the radial parts and applying the orthogonality conditions.) This fact simplifies the calculations significantly, allowing us to concentrate only on the radial dependence of the fit wave front.

The fit wave front W' may be written as a Zernike polynomial series with coefficients a_{n_1m}:

$$W' = \sum_{n_1}\sum_{m} a_{n_1m}U_{n_1}^m = \sum_{n_1} a_{n_1m}U_{n_1}^m. \tag{8}$$

Here the sum over m reduces to one term, the m value of the input wave front. If the radial dependence alone is considered, Eqs. (7) and (8) yield the following coupled equations:

$$\sum_{s=0}^{n'} C_n^m(s)p^{n-2s}r^{n-2s} = \sum_{n_1} a_{n_1m}R_{n_1}^m = \sum_{n_1} a_{n_1m}\sum_{s=0}^{n'} C_{n_1}^m(s)r^{n_1-2s}. \tag{9}$$

A solution can be found for the fit wave-front Zernike series coefficients by matching the coefficients of the powers of r:

$$\begin{bmatrix} C_n^m(0) & 0 & 0 & 0 & \cdots \\ C_n^m(1) & C_{n-2}^m(0) & 0 & 0 & \cdots \\ C_n^m(2) & C_{n-2}^m(1) & C_{n-4}^m(0) & 0 & \cdots \\ \vdots & \vdots & \vdots & & \ddots \end{bmatrix} \begin{pmatrix} a_{n,m} \\ a_{n-2,m} \\ a_{n-4,m} \\ \vdots \end{pmatrix} = \begin{pmatrix} p^nC_n^m(0) \\ p^{n-2}C_n^m(1) \\ p^{n-4}C_n^m(2) \\ \vdots \end{pmatrix}. \tag{10}$$

Each row corresponds to the coefficients of a different power of r. The first row represents the coefficients of r^n, the second row is for r^{n-2}, the third row is for r^{n-4}, and so forth, down to terms of order $n - |m|$ in r. We call the matrix $\mathbf{\Gamma}$, and the two column vectors are $\mathbf{a}$ and $\mathbf{p}$, respectively. Thus $\mathbf{\Gamma a} = \mathbf{p}$, and the elements of $\mathbf{a}$ can be found by matrix inversion: $\mathbf{a} = \mathbf{\Gamma}^{-1}\mathbf{p}$. Since there is a separate $\mathbf{\Gamma}$ matrix for each allowed pair of n and $|m|$ values, the notation $\mathbf{\Gamma^{nm}}$ will be used to remove ambiguity. Since C_n^m equals C_n^{-m}, $\mathbf{\Gamma^{nm}}$ is equivalent to $\mathbf{\Gamma^{n\,-m}}$.

The element of $\mathbf{\Gamma^{nm}}$ in the ith row and the jth column, Γ_{ij}^{nm}, is given by

$$\Gamma_{ij}^{nm} = \begin{cases} C_{n-2(i-1)}^m(j-i), & j \geqslant i \\ 0, & j < i \end{cases}. \tag{11}$$

Since $\mathbf{\Gamma^{nm}}$ is a lower-triangular matrix, its inverse will also be lower triangular. The elements of $\mathbf{p}$, p_j, are given by

$$p_j = p^{n-2(j-1)}C_n^m(j-1). \tag{12}$$

The first few elements of $\mathbf{a}$ can be read from Eq. (10) by inspection:

$$a_{nm} = p^n, \tag{13}$$

$$a_{n-2,m} = \frac{C_n^m(1)}{C_{n-2}^m(0)}p^{n-2}(1-p^2) = -p^n\frac{C_n^m(1)}{C_{n-2}^m(0)} + p^{n-2}\frac{C_n^m(1)}{C_{n-2}^m(0)}, \tag{14}$$

$$a_{n-4,m} = \frac{C_n^m(2)}{C_{n-4}^m(0)}p^{n-4}(1-p^4) - \frac{C_{n-2}^m(1)C_n^m(1)}{C_{n-4}^m(0)C_{n-2}^m(0)}p^{n-2}(1-p^2) = p^n\frac{1}{C_{n-4}^m(0)}\left[\frac{C_{n-2}^m(1)C_n^m(1)}{C_{n-2}^m(0)} - C_n^m(2)\right] - p^{n-2}\left[\frac{C_{n-2}^m(1)C_n^m(1)}{C_{n-4}^m(0)C_{n-2}^m(0)}\right] + p^{n-4}\left[\frac{C_n^m(2)}{C_{n-4}^m(0)}\right]. \tag{15}$$

Because $\mathbf{\Gamma^{nm}}$ is lower triangular, the solution for each coefficient a_{nm} relies only on the coefficients above it—that is, the coefficients of higher radial order. Table 1 contains explicit solutions for arbitrary p, up to the ninth Zernike polynomial term. This matrix will be represented by $\mathbf{H}(p)$ and is called the shrink matrix. The parameter p is included with the matrix notation to emphasize the functional dependence. The columns of $\mathbf{H}(p)$ are composed of the coefficient vectors $\mathbf{a}$, evaluated for each allowed pair of m and n values.

For values of p near 1, good approximations for the first 9×9 elements of the shrink matrix are shown in Table 2. Here the small quantity q is defined as $1 - p$. Table 3 evaluates the shrink matrix coefficients for the $p = 99\%$ and 98% cases. Table 4 contains an approximation of the nonzero elements of the shrink matrix for terms with n values up to 10 and all allowed m values. Terms of second order and higher in q are dropped in the approximation.

Table 1. Matrix of Fit Coefficients for Single Input Zernike Polynomial Terms of Unit Magnitude[a]

Fit Zernike Terms on a Subdomain			Input Zernike Terms of Unit Magnitude								
j			0	1	2	3	4	5	6	7	8
	n		0	1	1	2	2	2	3	3	4
		m	0	1	−1	0	2	−2	1	−1	0
0	0	0	1	0	0	$-(1-p^2)$	0	0	0	0	$-(1-p)(2p$ -
1	1	1		p	0	0	0	0	$-2p(1-p^2)$	0	0
2	1	−1			p	0	0	0	0	$-2p(1-p^2)$	0
3	2	0				p^2	0	0	0	0	$-3p^2(1-p$
4	2	2					p^2	0	0	0	0
5	2	−2						p^2	0	0	0
6	3	1							p^3	0	0
7	3	−1								p^3	0
8	4	0									p^4

[a] Note that the diagonal elements equal p^n. This matrix, $\mathbf{H}(p)$, is called the shrink matrix. Identically zero terms of higher than nth order are om from each column.

Table 2. Approximate Shrink Matrix $\mathbf{H}(p)$[a]

Fit Zernike Terms on a Subdomain			Input Zernike Terms of Unit Magnitude								
j			0	1	2	3	4	5	6	7	8
	n		0	1	1	2	2	2	3	3	4
		m	0	1	−1	0	2	−2	1	−1	0
0	0	0	1	0	0	$-2q$	0	0	0	0	$-2q$
1	1	1		p	0	0	0	0	$-4q$	0	0
2	1	−1			p	0	0	0	0	$-4p$	0
3	2	0				p^2	0	0	0	0	$-6p$
4	2	2					p^2	0	0	0	0
5	2	−2						p^2	0	0	0
6	3	1							p^3	0	0
7	3	−1								p^3	0
8	4	0									p^4

[a] Here q is equal to $1-p$, where p is the maximum radius of the subdomain on the full unit circle. Only the lowest-ordered q-dependent components are shown.

B. Shrink Matrix Elements

The elements of $\mathbf{H}(p)$ were derived in Subsection 3.A from the column vectors $\mathbf{a}$, calculated by matrix inversion or back-substitution from Eq. (10). It is useful, however, to have an explicit definition of those individual elements. The individual elements can be found from the solution for $\mathbf{a}$, $\mathbf{a} = \mathbf{\Gamma}^{-1}\mathbf{p}$, with the appropriate n and m input and output values. Consider an element of $\mathbf{a}$ that corresponds to the column of $\mathbf{H}(p)$ with input parameters n_0 and m_0. We can define this column by $a^{n_0 m_0}$. Then the element of $a^{n_0 m_0}$ that becomes a coefficient of U_n^m in the wave-front representation is $a_{nm}^{n_0 m_0}$.

Some confusion may arise from the fact that we are treating the index pair nm as a single index. This step is necessary in the matrix formalism, where rows and columns are described by separate, single indices. Since only some nm pairs are allowed, a relationship between nm and the single index i or j must be defined. This relationship can be defined in any self-consistent way.[7–10]

We have already discussed the fact that orthogonality forces all $a_{nm}^{n_0 m_0}$ terms with m not equal to m_0 to be 0. From $\mathbf{a} = \mathbf{\Gamma}^{-1}\mathbf{p}$, we can write an explicit solution for $a_{nm}^{n_0 m_0}$, considering only (n_0, m_0) and allowed values of (n, m):

$$a_{nm}^{n_0 m_0} = \delta_{m m_0}[n \leqslant n_0] \sum_{j=1}^{L} [(\Gamma^{n_0 m_0})^{-1}]_{ij} p_j. \qquad (16)$$

In Eq. (16), the row index i represents the pair of indices nm. The bracketed expression represents a binary operator, behaving like the Heaviside unit function[11] (also called the step function): When the expression expr is true, [expr] equals 1; otherwise, [expr] equals 0. By Eq. (12), we can write the p_j terms explicitly in terms of powers of the scalar p and C_n^m:

$$a_{nm}^{n_0 m_0} = \delta_{m m_0}[n \leqslant n_0] \times \sum_{j=1}^{L} [(\Gamma^{n_0 m_0})^{-1}]_{ij} p^{n_0 - 2(j-1)} C_{n_0}^{m_0}(j-1). \quad (17)$$

These are the matrix elements of $\mathbf{H}(p)$.

C. Shrink Matrix Formulation

With the elements of the shrink matrix now defined, the results can be generalized for an arbitrary input wave front. As in Subsection 3.B, we represent the allowed n and m indices in terms of a single parameter j. The input wave front W can be represented by its Zernike coefficients as a vector $\mathbf{w}$ defined on the basis of Zernike polynomials, $\mathbf{U}$. W may be written as

$$W = \sum_j w_j U_j = \mathbf{U} \cdot \mathbf{w} = \mathbf{U}^T \mathbf{w}. \quad (18)$$

With use of the shrink matrix $\mathbf{H}(p)$, the fit wave-front coefficient vector on the centered subdomain $\mathbf{w}_p$ is

Table 3. Matrix of Fit Coefficients for Single Input Zernike Polynomial Terms of Unit Magnitude with $p = 99\%$ and $p = 98\%$ Subdomain Radii[a]

Fit Zernike Terms			Input Zernike Terms of Unit Magnitude																
			$p = 99\%$ Subdomain Radii									$p = 98\%$ Subdomain Radii								
j			0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
	n		0	1	1	2	2	2	3	3	4	0	1	1	2	2	2	3	3	4
		m	0	1	−1	0	2	−2	1	−1	0	0	1	−1	0	2	−2	1	−1	0
0	0	0	1	0	0	−0.020	0	0	0	0	−0.019	1	0	0	−0.040	0	0	0	0	−0.036
1	1	1		0.99	0	0	0	0	−0.039	0	0		0.98	0	0	0	0	−0.078	0	0
2	1	−1			0.99	0	0	0	0	−0.039	0			0.98	0	0	0	0	−0.078	0
3	2	0				0.98	0	0	0	0	−0.059				0.96	0	0	0	0	−0.11
4	2	2					0.98	0	0	0	0					0.96	0	0	0	0
5	2	−2						0.98	0	0	0						0.96	0	0	0
6	3	1							0.97	0	0							0.94	0	0
7	3	−1								0.97	0								0.94	0
8	4	0									0.96									0.92

[a] Terms are rounded to two significant figures.

Table 4. Approximate Nonzero Elements of the Shrink Matrix on a Subdomain of Radius p with $q = 1 - p$[a]

$U_0^0 \to U_0^0$
$U_1^{\pm 1} \to p U_1^{\pm 1}$
$U_2^0 \to p^2 U_2^0 - 2q U_0^0$
$U_2^2 \to p^2 U_2^2$
$U_3^{\pm 1} \to p^3 U_3^{\pm 1} - 4q U_1^{\pm 1}$
$U_3^{\pm 3} \to p^3 U_3^{\pm 3}$
$U_4^0 \to p^4 U_4^0 - 6U_2^0 - 2q U_0^0$
$U_4^{\pm 2} \to p^4 U_4^{\pm 2} - 6U_2^{\pm 2}$
$U_4^{\pm 4} \to p^4 U_4^{\pm 4}$
$U_5^{\pm 1} \to p^5 U_5^{\pm 1} - 8q U_3^{\pm 1} - 4q U_1^{\pm 1}$
$U_5^{\pm 3} \to p^5 U_5^{\pm 3} - 8q U_3^{\pm 3}$
$U_5^{\pm 5} \to p^5 U_2^{\pm 5}$
$U_6^0 \to p^6 U_6^0 - 10q U_4^0 - 6q U_2^0 - 2q U_0^0$
$U_6^{\pm 2} \to p^6 U_6^{\pm 2} - 10q U_4^{\pm 2} - 6q U_2^{\pm 2}$
$U_6^{\pm 4} \to p^6 U_6^{\pm 4} - 10q U_4^{\pm 4}$
$U_6^{\pm 6} \to p^6 U_6^{\pm 6}$
$U_7^{\pm 1} \to p^7 U_7^{\pm 1} - 12q U_5^{\pm 1} - 8q U_3^{\pm 1} - 4q U_1^{\pm 1}$
$U_7^{\pm 3} \to p^7 U_7^{\pm 3} - 12q U_5^{\pm 3} - 8q U_3^{\pm 3}$
$U_7^{\pm 5} \to p^7 U_7^{\pm 5} - 12q U_5^{\pm 5}$
$U_7^{\pm 7} \to p^7 U_7^{\pm 7}$
$U_8^0 \to p^8 U_8^0 - 14q U_6^0 - 10q U_4^0 - 6q U_2^0 - 2q U_0^0$
$U_8^{\pm 2} \to p^8 U_8^{\pm 2} - 14q U_6^{\pm 2} - 10q U_4^{\pm 2} - 6q U_2^{\pm 2}$
$U_8^{\pm 4} \to p^8 U_8^{\pm 4} - 14q U_6^{\pm 4} - 10q U_4^{\pm 4}$
$U_8^{\pm 6} \to p^8 U_8^{\pm 6} - 14q U_6^{\pm 6}$
$U_8^{\pm 8} \to p^8 U_8^{\pm 8}$
$U_9^{\pm 1} \to p^9 U_9^{\pm 1} - 16q U_7^{\pm 1} - 12q U_5^{\pm 1} - 8q U_3^{\pm 1} - 4q U_1^{\pm 1}$
$U_9^{\pm 3} \to p^9 U_9^{\pm 3} - 16q U_7^{\pm 3} - 12q U_5^{\pm 3} - 8q U_3^{\pm 3}$
$U_9^{\pm 5} \to p^9 U_9^{\pm 5} - 16q U_7^{\pm 5} - 12q U_5^{\pm 5}$
$U_9^{\pm 7} \to p^9 U_9^{\pm 7} - 16q U_7^{\pm 7}$
$U_9^{\pm 9} \to p^9 U_9^{\pm 9}$
$U_{10}^0 \to p^{10} U_{10}^0 - 18q U_8^0 - 14q U_6^0 - 10q U_4^0 - 6q U_2^0 - 2q U_0^0$
$U_{10}^{\pm 2} \to p^{10} U_{10}^{\pm 2} - 18q U_8^{\pm 2} - 14q U_6^{\pm 2} - 10q U_4^{\pm 2} - 6q U_2^{\pm 2}$
$U_{10}^{\pm 4} \to p^{10} U_{10}^{\pm 4} - 18q U_8^{\pm 4} - 14q U_6^{\pm 4} - 10q U_4^{\pm 4}$
$U_{10}^{\pm 6} \to p^{10} U_{10}^{\pm 6} - 18q U_8^{\pm 6} - 14q U_6^{\pm 6}$
$U_{10}^{\pm 8} \to p^{10} U_{10}^{\pm 8} - 18q U_8^{\pm 8}$
$U_{10}^{\pm 10} \to p^{10} U_{10}^{\pm 10}$

[a] For a given input aberration with unit magnitude of a single Zernike polynomial term U_n^m, the components of the fit wave front are shown. The results shown assume that $p \approx 1$ and are thus valid for small values of q only. Only the lowest-ordered q-dependent components are shown.

$$\mathbf{w}_p = \mathbf{H}(p)\mathbf{w}. \tag{19}$$

And the fit wave front may be written as

$$W' = \mathbf{U} \cdot \mathbf{w}_p = \mathbf{U} \cdot [\mathbf{H}(p)\mathbf{w}]. \tag{20}$$

4. INTERPRETATION OF THE RESULTS

In the interpretation, we must be careful to avoid the temptation to extrapolate beyond the restricted subdomain into the area where the wave front is not measured. We cannot assume measured values for the input wave front across the full pupil when only the subdomain fit wave front has been measured. However, we can say that given a well-behaved wave front across the full pupil, the error magnitudes that we can expect to observe in the subdomain wave front are as shown in the tables presented here. To ensure high accuracy, we must choose the size of the subdomain large enough to reduce the relative measurement errors to below acceptable limits. Likewise, the measurement uncertainties can be predicted in part by the measurement errors revealed by the calculations presented here. It is important to note that the most difficult errors to account for are those introduced by the presence of higher-ordered aberrations, which generally couple strongly into the measurement lower-ordered aberrations.

The coefficients of the shrink matrix contain the relative systematic error magnitude information that can be used to set error tolerances or bounds on the minimum size of the measurement subdomain. In this section, two examples are given: The first is specific, and the second is general.

A. Example: Spherical Aberration

As discussed above, when a hypothetical input wave front consists of a single Zernike polynomial of unit magnitude, the fit wave front measured on the restricted subdomain will not match the input exactly. Consider, for example, the case of an input wave front with only spherical aberration—an aberration term of fourth order in r. The form of spherical aberration in the Zernike polynomial description is

$$U_4^0(r, \theta) = 6r^4 - 6r^2 + 1. \tag{21}$$

When the input wave front has unit magnitude of spherical aberration and the restricted subdomain has a relative radius p, then the fit wave front will have a spherical aberration magnitude of p^4: a relative error of $1 - p^4$, underestimating the input wave-front error magnitude. In addition, because of the presence of the second-order and constant components, spurious defocus and piston will appear in the fit wave front with magnitudes $-3p^2(1 - p^2)$ and $(1 - p)(2p - 1)$, respectively, even though no net defocus or piston was present in the input wave front. (These magnitudes were read directly from the $j = 8$ column of Table 1.) If we define q equal to $1 - p$, as in Subsection 3.A, then the two error magnitudes may be written approximately as $-6q$ for defocus and $-2q$ for piston.

In practice, reasonable values for p and q may be 99% and 1%, respectively. In this case, a single wave of spherical aberration in the input wave front will appear as 0.96 wave of spherical aberration with -0.059 wave of defocus and -0.019 wave of piston in the fit wave front.

B. Example: Arbitrary Input Wave Front

To further illustrate the operation of the shrink matrix, we consider how it affects an arbitrary input wave front. On the full-pupil domain, let the input wave front have aberrations defined by the coefficient column vector $\mathbf{w}$, defined in the Zernike polynomial convention used throughout this paper.[8] Suppose that only the first nine elements of $\mathbf{w}$ (piston through spherical aberration) contain nonzero elements.

In a hypothetical interferometric measurement, a concentric subdomain of the full pupil is selected, and a unit-circle coordinate system is defined based on this radius. In the coordinates of the full pupil, the subdomain radius is p. q is defined as $1 - p$. The approximate shrink matrix, given in Table 2, enables us to quickly determine the observed fit wave front on the restricted subdomain. The approximate fit wave front $\mathbf{w}'$ is given by

$$\mathbf{w}' = \begin{pmatrix} w_0'^0 \\ w_1'^1 \\ w_1'^{-1} \\ w_2'^0 \\ w_2'^2 \\ w_2'^{-2} \\ w_3'^1 \\ w_3'^{-1} \\ w_4'^0 \end{pmatrix} = \mathbf{H}(p)\mathbf{w} = \begin{pmatrix} w_0^0 - 2qw_2^0 - 2qw_4^0 \\ pw_1^1 - 4qw_3^1 \\ pw_1^{-1} - 4qw_3^{-1} \\ p^2w_2^0 - 6qw_4^0 \\ p^2w_2^2 \\ p^2w_2^{-2} \\ p^3w_3^1 \\ p^3w_3^{-1} \\ p^4w_4^0 \end{pmatrix}. \tag{22}$$

The measurement error is contained in the differences between the full wave-front elements w_n^m and the fit wave-front elements, which can be read from Eq. (22). Again, we must acknowledge that the truncation of the series at spherical aberration, the ninth term, has limited contributions from higher-ordered aberrations that may be present experimentally.

C. General Error Estimation

In practice, the subdomain size may be selected based on the relative or the absolute fit fidelity required in specific aberration terms. The requirements become increasingly more restrictive for terms of higher radial order.

Consider a single Zernike aberration polynomial term of absolute magnitude a and radial order n in the input wave front. With a restricted subdomain radius p, the systematic error ϵ in the measurement of that Zernike term's magnitude is given approximately by

$$\epsilon = a(1 - p^n) \approx anq. \tag{23}$$

Barring all other systematic error sources, to achieve an accuracy of ϵ, we must choose the subdomain radius to satisfy

$$p > 1 - \frac{\epsilon}{an}. \tag{24}$$

This follows from Eq. (23). If the error tolerance ϵ were set at 0.05 nm and the magnitude of spherical aberration (fourth order) in the input wave front, a, were 0.25 nm,

then p would be greater than 95%. This represents a relative error tolerance of 20%. If ϵ were 0.01 nm and a were 0.25 nm, then p would be greater than 99%—a relative error tolerance of 4%.

Note that in Eq. (23) the error magnitude increases linearly with the radial order of the Zernike term under consideration, with the absolute magnitude of the input wave-front aberration, and with the fraction by which the subdomain is smaller than the full domain.

As stated above, in this analysis one must be aware of and not neglect the contributions of higher-ordered terms than the term being considered. If they are present with significant absolute magnitude, they may make the subdomain minimum size requirements stricter. Because of the case-dependent nature of those considerations, the contributions of higher-ordered terms are difficult to generalize.

5. CONCLUSION

Wave-front analysis with Zernike polynomials requires the definition of a unit-circle domain on which a wave-front fit is performed and the coefficients of aberration polynomials are determined. On the discrete domain of measurement, the definition of the unit circle is often made on a restricted subdomain of the full wave-front data in order to remove unreliable edge points from the analysis. The subdomain selection can be a significant and yet often a neglected systematic error and uncertainty source in wave-front measurement. These errors include a reduction in the apparent magnitude of some wave-front aberrations and spurious contributions from aberrations that may or may not be present in the full wave front.

By studying the way in which aberrations on a full-pupil domain map to aberrations on a restricted subdomain, we have developed a method for estimating measurement errors. Dependent on the size of the restricted subdomain and on the magnitude and the composition of the wave-front errors, the method described can be used to set tolerances on the minimum size of the restricted subdomain that is required to achieve arbitrary fit fidelity.

Using linear superposition of the aberration components and a vector–matrix notation to represent the polynomial coefficients, we can apply the method to arbitrary wave-front aberrations defined by orthogonal aberration polynomials. The analysis described here is applied to circular pupils, for which the Zernike polynomials form an appropriate basis to describe the aberrations. Specific examples are given for all of the terms that appear in the first 36 conventional Zernike polynomial terms.

Calculations show that the error magnitudes, and hence the accuracy level, always scale with the magnitude of the aberrations in the input wave front and with the radial order of constituent Zernike polynomial terms. Thus the systematic errors are reduced as the wave-front aberration magnitude is reduced; furthermore, the errors are most significant for the measurement of the aberration terms with the highest radial order and in the presence of higher-order aberrations.

Address correspond to Kenneth A. Goldberg at the location on the title page or by e-mail, kagoldberg@lbl.gov.

REFERENCES

1. D. J. Fischer, J. T. O'Bryan, R. Lopez, and H. P. Stahl, "Vector formulation for interferogram surface fitting," Appl. Opt. **32**, 4738–4743 (1993).
2. J. Y. Wang and D. E. Silva, "Wave-front interpretation with Zernike polynomials," Appl. Opt. **19**, 1510–1518 (1980).
3. M. Carpio and D. Malacara, "Closed Cartesian representation of the Zernike polynomials," Opt. Commun. **110**, 514–516 (1994).
4. S. D. Conte and C. de Boor, *Elementary Numerical Analysis: An Algorithmic Approach*, 3rd ed. (McGraw-Hill, New York, 1980), Chap. 6.
5. P. R. Bevington and D. K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences*, 2nd ed. (McGraw-Hill, Boston, 1992), p. 126.
6. C. R. Hayslett and W. H. Swantner, "Wave-front derivation by three computer programs," Appl. Opt. **19**, 161–163 (1980).
7. V. N. Mahajan, "Zernike circle polynomials and optical aberrations of systems with circular pupils," Appl. Opt. **33**, 8121–8124 (1994).
8. J. S. Loomis, *FRINGE User's Manual* (Optical Sciences Center, University of Arizona, Tucson, Ariz., 1976).
9. "FRINGE Zernikes," in *Code V Reference Manual*, R. C. Juergens, ed. (Optical Research Associates, Pasadena, Calif., 1998), Vol. 1, version 8.30, pp. 2A-595–2A-596.
10. "Fitting interferograms using Zernike polynomials," in *FAST! V/AI Operations Manual* (Phase Shift Technology, Tucson, Ariz., 1990), pp. 2-3–2-5.
11. J. Spanier and K. B. Oldham, "The unit-step and related functions," in *An Atlas of Functions* (Hemisphere, Washington, D.C., 1987), Chap. 8, pp. 63–69.