Determination of phase mode components in terms of local wave-front slopes: an analytical approach

E. Acosta, S. Bará, M. A. Rama, and S. Ríos

Area de Optica, Departamento de Física Aplicada, Facultade de Física, Universidade de Santiago de Compostela, Campus Sur, 15706 Santiago de Compostela, Galicia, Spain

Received January 10, 1995

An analytical formulation that relates the modal expansion coefficients of a given wave front to its local transverse phase derivatives is proposed. The modal coefficients are calculated as a weighted integral over the wave-front slopes. The weighting functions for each mode are the components of a two-dimensional vector whose divergence equals the corresponding mode function. This approach is useful for analytical phase reconstruction from the input data provided by shearing interferometers or Hartmann–Shack wave-front sensors. Numerical results for a simulated experiment in terms of a set of Zernike polynomials are given.

The problem of reconstructing the phase of a wave front from its measured local slopes is a subject of interest in the fields of optical testing and adaptive optics and in some postdetection techniques for image restoration. Shearing interferometers and Hartmann–Shack wave-front sensors can provide direct estimations of the local transverse phase gradient components in a discrete set of points. These data have to be processed for the unknown phase distribution over the entrance pupil of the system to be calculated. Several strategies have been developed to cope with this problem. Southwell has pointed out the advantages of modal approaches (in which the phase is expanded as a linear combination of a given set of orthogonal functions and the problem is then reduced to the determination of the expansion coefficients) over the classical zonal schemes (which attempt to reconstruct the phase locally).

The usual least-squares fitting methods for modal phase reconstruction are based on the solution of a set of linear algebraic equations in which the phase slopes are directly introduced or represented as phase finite differences. Recently Aksenov and Isaev proposed a novel scheme establishing an analytical integral relationship between the expansion coefficients and the phase gradient components. Their approach, based on the properties of the Radon transform, overcomes some drawbacks associated with the standard least-squares techniques and permits greater flexibility in the design of the wave-front sensor devices.

In this Letter we propose an integral formulation that permits us to obtain an analytical solution to this problem. This approach is based on the definition of an auxiliary set of two-dimensional vector functions, each of them associated with a particular phase mode, and the application of the divergence theorem to their domain of definition.

Let $S(r)$ be the phase of the wave front, and let

$$S(r) = \sum a_k \psi_k(r)$$  \hspace{1cm} (1)

be the expansion of $S$ as a linear combination of a set of orthogonal functions $\{\psi_k\}$ defined in a two-dimensional domain $\sigma$ (Fig. 1), which will usually coincide with the input pupil of the system or with any domain conjugate to it. $\{a_k\}$ are the mode coefficients to be determined. Because $\{\psi_k\}$ are orthogonal, the value of the coefficient $a_n$ for a given $n$ is

$$a_n = \langle \psi_n, S \rangle / \langle \psi_n, \psi_n \rangle,$$  \hspace{1cm} (2)

with the inner product of two functions, $f$ and $g$, defined as

$$\langle f, g \rangle = \iint_\sigma f^*(r) g(r) \, d\sigma,$$  \hspace{1cm} (3)

where * denotes the complex conjugate. To obtain $a_n$ as a function of the wave-front slopes $(\nabla S)$, we introduce a set of auxiliary two-dimensional vector functions $\{F_k\}$ related to the basis functions through the condition

$$\nabla \cdot F_k(r) = \psi_k^*(r),$$  \hspace{1cm} (4)

where $\nabla$ is the two-dimensional nabla operator defined in Cartesian coordinates as $\nabla = (\partial/\partial x), (\partial/\partial y)$.

Therefore Eq. (2) can be rewritten as

$$a_n = c_n \int_\sigma \nabla \cdot (S F_n) \, d\sigma - c_n \int_\sigma (\nabla S) \cdot F_n \, d\sigma,$$  \hspace{1cm} (5)

where $c_n = \langle \psi_n, \psi_n \rangle^{-1}$, the explicit dependence of $F_n$ and $S$ on $r$ has been omitted, and it has been taken into account that

$$(\nabla \cdot F_n) S = \nabla \cdot (S F_n) - (\nabla S) \cdot F_n.$$  \hspace{1cm} (6)

Fig. 1. Definition domain of the wave-front function.
Now, if $SF_{\mathbf{F}}$ is continuously differentiable in $\sigma$ and has continuous components in the contour $C$, the divergence theorem$^9$ can be applied to the first integral on the right-hand side of Eq. (5), yielding

$$a_n = c_n \int_C SF_{\mathbf{F}} \cdot d\mathbf{c} - c_n \int_0^\sigma (\nabla S) \cdot F_{\mathbf{d}} d\sigma,$$

(7)

where $d\mathbf{c}$ is a differential contour element of length $d\mathbf{c}$ and direction given by the unitary vector normal to $C$ pointing out of the domain.

Finally, note that for each $\psi_k$ the definition of $F_{\mathbf{F}}$ is no longer unique. There is a wide family of functions that verify Eq. (4). As far possible we choose $F_{\mathbf{F}}(r)$ in such way that the integral along $C$ in Eq. (7) vanishes (the most obvious choices are to require $F$ to be either orthogonal to $d\mathbf{c}$ or zero everywhere in $C$); then Eq. (7) becomes

$$a_n = -c_n \int_0^\sigma (\nabla S) \cdot F_{\mathbf{d}} d\sigma,$$

(8)

which gives the desired value of $a_n$ as a weighted integral of the wave-front slopes represented by the components of $\nabla S$.

As an example application, consider the case when the basis functions are the real Zernike polynomials $\{\psi_n\}$ defined in the unit radius circle as$^{10}$

$$\psi_{nl}(\rho, \theta) = R_{nl}^{|l|}(\rho) \sin(|l|\theta) \quad \text{for } l > 0,$$

$$\psi_{nl}(\rho, \theta) = R_{nl}^{|l|}(\rho) \cos(|l|\theta) \quad \text{for } l \leq 0,$$

(9)

where $\rho$ is the normalized radial distance from the origin, $\theta$ is the angle measured from the $Y$ axis, and $R_{nl}^{|l|}(\rho)$ are the standard radial functions defined in Ref. 10. Equation (4) can be explicitly written as

$$\frac{(1/\rho) \partial}{\partial \rho}(\rho F_{nl,\rho}) + \frac{(1/\rho) \partial}{\partial \theta} F_{nl,\theta} = \psi_{nl}(\rho, \theta),$$

(10)

where $F_{nl,\rho}$ and $F_{nl,\theta}$, respectively, are the radial and angular components of $F_{\mathbf{F}}(\rho, \theta)$.

A possible choice for $F_{nl}(\rho, \theta)$ is given by

(a) Mode coefficients $a_{nl}$ with $l \neq 0$. We obtain a simple and useful set of solutions by imposing $F_{nl,\rho}(\rho, \theta) = 0$ in $\sigma + C$, which ensures the orthogonality of $F_{nl}(\rho, \theta)$ and $d\mathbf{c}$ in $C$. Then, integrating Eq. (10), we obtain

$$F_{nl,\theta}(\rho, \theta) = -(\rho/|l|)R_{nl}^{|l|}(\rho) \cos(|l|\theta) + u(\rho) \quad \text{for } l > 0,$$

$$F_{nl,\theta}(\rho, \theta) = (\rho/|l|)R_{nl}^{|l|}(\rho) \sin(|l|\theta) + w(\rho) \quad \text{for } l < 0,$$

(11)

where $u(\rho)$ and $w(\rho)$ are any functions of $\rho$ at least continuously differentiable in $\sigma$ and continuous in $C$.

(b) Mode coefficients $a_{nl}$ with $l = 0$. In this case the choice $F_{n0,\rho}(\rho, \theta) = 0$ would lead to an $SF_{\mathbf{F}}$ function nondifferentiable in $\sigma$. An alternative possibility is to set $F_{n0,\theta} = 0$, which gives

$$F_{n0,\rho}(\rho, \theta) = (1/\rho) \int_0^\rho R_{n0}(|l|\rho') d\rho'.$$

(12)

This kind of function vanishes at any point in the unit radius circle ($\rho = 1$) because of the orthogonality properties of the $R_{nl}^{|l|}$ polynomials,$^{11}$ thus permitting the application of Eq. (8).

The application of Eq. (8) requires a choice of the numerical procedure for computation of the value of the integral from the wave-front slope measurement data. Simpson’s method is a standard one, which, although it has some advantages, requires a relatively high number of sampling points for obtaining precise results. In some situations, more efficient algorithms can be used. If the wave-front phase in Eq. (1) can be represented accurately by a truncated series of Zernike polynomials $\{\psi_n\}$, whose radial orders $n$ are smaller than or equal to a given number $N$, with no modeling error (that is, with contributions of $n > N$ negligible), then Eq. (8) can be evaluated by a precise numerical integration with a rather small grid. To make this evaluation we first recall that in polar coordinates ($\rho, \theta$) the integrand in Eq. (8) will have a radial part consisting of powers of $\rho$ of order $2N + 1$ or smaller and an angular part formed by different combinations of trigonometric functions that have the highest angular dependence determined by $\sin(2N\theta)$ or $\cos(2N\theta)$. The radial contribution can be exactly evaluated by a standard Gaussian quadrature method$^{12}$ that samples the wave-front slopes at a set of $N + 1$ radial locations determined by $\rho[i] = (1/2)(z[i] - 1)$, $i = 0 \ldots N$, where $z[i]$ are the zeros of the Legendre polynomial of order $N + 1$. We can also precisely evaluate the angular part by sampling the wave-front slopes at a set of $2N + 1$ equally spaced angular locations$^{13}$ given by $\theta[j] = j2\pi/(2N + 1)$. This procedure requires the use of a polar grid with $(2N + 1)(N + 1)$ sampling points for an exact evaluation of Eq. (8) under the assumptions of no modeling error and no slope measurement noise. Recall that, to order $N$, there are $(1/2)(N + 1)(N + 2)$ linearly independent Zernike polynomials whose modal coefficients can be accurately determined in this way.

As an example application we generated a simulated wave front by adding the contributions of the first 14 nontrivial Zernike polynomials, with relative weights $c_{nl}$ equal to 1 or 0, as indicated in Table 1. The resulting phase distribution is depicted in Fig. 2. The wave-front slopes were analytically calculated from the wave-front phase function at a set of 45 points (polar grid of $5 \times 9$ angular positions). The sampling locations are determined by the formulas given above. The $a_{nl}$ coefficients can be recovered within the precision of the numerical routines used to compute the integral (see Table 1). Provided that no modeling error or noise is present in this simplified problem, the same results can essentially be obtained by a standard least-squares procedure.

In comparison with Simpson’s and related methods of numerical evaluation of Eq. (8), Gaussian quadrature permits higher precision with a considerably smaller number of sampling points, at least as far as a low-order expansion of the wave-front phase is useful. Nevertheless, it seems to be more sensitive to wave-front modeling errors. Compared with the standard least-squares method, this approach shares the advantages pointed out by Wang and Silva when they analyzed the recovery of the modal coefficients through
As a conclusion, the modal expansion coefficients of a wave front have been calculated in terms of a weighted integral of the wave-front slopes. The method is valid for any set of functions forming an orthogonal basis on the aperture. As in the Aksenov–Isaev approach, this formulation overcomes some problems that arise in standard least-squares techniques. It is also highly compatible with a wide range of wave-front sensor sampling geometries.

References