Highly Accurate Wavefront Reconstruction Algorithms Over Broad Spatial-Frequency Bandwidth

Introduction

Frequency-domain wavefront reconstruction methods are as old as the very early wavefront reconstructors.[1,2] Freischlad placed this subject on solid ground.[3] The rectangular map constraint of the conventional Fourier method has been removed in an iterative Gerchberg-type algorithm dealing with an arbitrary boundary shape.[4] A series of recent papers by Poyneer discuss improvements on handling boundary conditions and applications in extreme adaptive optics.[5,6] Similar principles have been applied in shearing interferometers.[7] More serious attention has been paid to the accuracy of the reconstruction methods in Refs. 8–10. The works of Campos and Yaroslavsky presented a solution based on a band-limited integration technique in frequency domain. The two-dimensional (2-D) extension of the same method was not discussed. Complementary to their works, Bahk introduced a full 2-D wavefront reconstructor based on the band-limited derivative calculation.[11] Both approaches emphasize the frequency response of the reconstructed signals. The frequency response of wavefront reconstruction has been discussed earlier in the analysis of lateral-shearing interferometry.[12] Frequency-response characteristics of a reconstruction is important in focal-spot diagnostics for high-power lasers, where the focal spot is indirectly characterized using wavefront information reconstructed from Shack–Hartmann slopes data.[13]

This article develops a set of encompassing mathematical tools for wavefront reconstruction problems, where many additional benefits naturally arise, interconnecting the results of previous works. The benefits are exemplified by the development of two new wavefront reconstructors and the analytical derivation of noise-propagation coefficients of several well-known wavefront reconstructors.

This article is organized as follows: (1) The mathematical tools and symbols regarding band-limited derivative operations, which are needed for the analyses in the subsequent sections, are introduced. (2) A way to improve the accuracy of the finite-difference method is discussed in connection with wavefront reconstruction. The Simpson rule is adopted for developing a new spatial-domain iterative reconstruction algorithm. The exact details of the algorithm and its frequency-domain property are described. (3) A band-limited reconstruction algorithm is extended to hexagonal geometry, which greatly enhances the flexibility of band-limited reconstructors. (4) Finally, the noise-propagation curve is analytically derived and compared with numerical simulations.

Band-Limited Derivative

The main results of band-limited derivative techniques in the context of wavefront reconstruction were summarized in Ref. 11. The full derivation of the results will be presented here for the sake of completeness. Additional new notations are introduced that will simplify the expressions in **Hexagonal Band-Limited Reconstructor** (p. 136).

The motivation for band-limited derivatives, especially for discrete samples, lies in the fact that it provides an analytical tool for converting back and forth between slope measurements and wavefront signal. We start by asking what the exact interpolation formula is for derivatives in discrete samples. According to sampling theorem, a band-limited signal can be exactly reconstructed at any point by convolving a sinc function with discrete samples. The derivative of a band-limted signal is obtained by directly differentiating the sinc function's convolution kernel that becomes a spherical Bessel function (j_1) (Ref. 14). The derivative interpolation expression at discrete points is

$$\frac{d\varphi}{dx}\bigg|(x = m\Delta x) \equiv \varphi_x(m)$$

$$= \frac{1}{\Delta x}\sum_{n=1}^{\infty}\pi\,j_1(n\pi)\big\{\varphi\big[(m+n)\Delta x\big] - \varphi\big[(m-n)\Delta x\big]\big\}, \quad (1)$$

where the spherical Bessel function evaluated at integer multiples of π is equivalent to

$$\pi\,j_1(n\pi) = \frac{(-1)^{n+1}}{n}, \quad n = 1, 2, \ldots \quad (2)$$

The summation of the left-hand side of Eq. (1) for all sample points can be shown to be equal to zero by taking advantage of the expression on the right-hand side and using the periodicity condition of discrete samples:

$$\sum_{m=0}^{N-1} \frac{\mathrm{d}\varphi}{\mathrm{d}x}\bigg|(x = m\Delta x) = 0. \tag{3}$$

Equation (1) is easier to handle in frequency domain. Discrete Fourier transform (DFT) and Fourier series analysis lead to the following equivalent expression:

$$\widetilde{\varphi_x}(k) = i\frac{2\pi}{\Delta x}\mathbb{S}(k)\tilde{\varphi}(k), \tag{4}$$

where the tilde notation means DFT of the symbol beneath it and $\mathbb{S}(k)$ (sawtooth wave) is defined as

$$\mathbb{S}(k) = \begin{cases} \text{for even } N\begin{cases} \dfrac{k}{N}, & k = 0, \ldots, \ N/2 - 1 \\ 0, & k = N/2 \\ \dfrac{k}{N} - 1, & k = \dfrac{N}{2} + 1, \ldots, \ N - 1 \end{cases} \\ \text{for odd } N\begin{cases} \dfrac{k}{N}, & k = 0, \ldots, \ (N-1)/2 \\ \dfrac{k}{N} - 1, & k = (N+1)/2, \ldots, \ N-1 \end{cases} \end{cases} \tag{5}$$

Equation (4) provides a convenient way of calculating exact derivatives from band-limited signals. When the sampling points of a derivative signal are offset by a half-sampling space from the sampling points of the original signal, a slightly different form should be used:

$$\frac{\mathrm{d}\varphi}{\mathrm{d}x}\bigg|\left(x = m\Delta x + \frac{1}{2}\Delta x\right) \equiv \varphi_{x,1/2}(m)$$

$$= \frac{1}{\Delta x}\sum_{n=1}^{\infty}\pi\, j_1\left(n\pi - \frac{1}{2}\pi\right)$$

$$\times \left\{\varphi\big[(m+n)\Delta x\big] - \varphi\big[(m-n+1)\Delta x\big]\right\}, \tag{6}$$

where the Bessel coefficients can be replaced again with an integer expression

$$\pi\, j_1\left(n\pi - \frac{1}{2}\pi\right) = \frac{4}{\pi}\frac{(-1)^{n+1}}{(2n-1)^2}, \quad n = 1, 2, \ldots \tag{7}$$

Employing a similar Fourier series analysis that leads to Eq. (4), the frequency-domain expression of Eq. (6) is reduced to

$$\widetilde{\varphi_{x,1/2}}(k) = i\frac{2\pi}{\Delta x}\exp\left(i\frac{\pi}{N}k\right)\mathbb{T}(k)\tilde{\varphi}(k), \tag{8}$$

where $\mathbb{T}(k)$ (triangular wave) is defined as

$$\mathbb{T}(k) = \begin{cases} k/N, & k = 0, \ldots, \left|\dfrac{N-1}{2}\right| \\ 1 - k/N, & k = \left|\dfrac{N-1}{2}\right| + 1, \ldots, N-1 \end{cases} \tag{9}$$

We also need an interpolation formula for creating a signal shifted by half-sample spacing for Fried geometry:

$$\varphi(x)\bigg|\left(x = m\Delta x + \frac{1}{2}\Delta x\right) \equiv \varphi_{1/2}(m)$$

$$= \sum_{n=1}^{\infty}\text{sinc}\left(n - \frac{1}{2}\right)\left\{\varphi\big[(m+n)\Delta x\big] - \varphi\big[(m-n+1)\Delta x\big]\right\}. \tag{10}$$

The DFT of Eq. (10) is

$$\widetilde{\varphi_{1/2}}(k) = \exp\left(i\frac{\pi}{N}k\right)\mathbb{R}(k)\tilde{\varphi}(k), \tag{11}$$

where $\mathbb{R}(k)$ (rectangular wave) is

$$\mathbb{R}(k) = \begin{cases} \text{for even } N\begin{cases} 1, & k = 0, \ldots, \ N/2 - 1 \\ 0, & k = N/2 \\ -1, & k = N/2 + 1, \ldots, \ N-1 \end{cases} \\ \text{for odd } N\begin{cases} 1, & k = 0, \ldots, \ (N-1)/2 \\ -1, & k = (N+1)/2, \ldots, \ N-1 \end{cases} \end{cases} \tag{12}$$

Therefore, the partial derivative in the x direction for Fried geometry in frequency domain is

$$\widetilde{\varphi_{x,1/2,1/2}}(p,q)$$

$$= i\frac{2\pi}{\Delta x}\exp\left[i\frac{\pi}{N}(p+q)\right]\mathbb{T}(q)\mathbb{R}(p)\tilde{\varphi}(p,q). \tag{13}$$

Equation (13) has an additional degree of freedom (index p for the y direction) because the reconstructed sample point in the Fried geometry must first be shifted in the y direction by a half-sample size before applying the half-sample shifted-derivative operation in the x direction.

For consistency, we can verify that the sequential operations of half-pixel shift and derivative operation using $\mathbb{R}(k)$ and $\mathbb{S}(k)$ produce the same result as single operation of $\mathbb{T}(k)$, i.e., $\mathbb{T}(k) = \mathbb{R}(k)\mathbb{S}(k)$. This relation, however, does not hold for the value at $N/2$ for even N, where the left-hand side is 0.5, whereas the right-hand side is 0. To remove this paradox for an even number of samples, we choose to use $\mathbb{S}(N/2) = 0.5$ and $\mathbb{R}(N/2) = 1$ or $\mathbb{S}(N/2) = -0.5$ and $\mathbb{R}(N/2) = -1$. A similar choice was made in Ref. 9 for band-limited integration operators from a different perspective. *The redefinition of $\mathbb{S}$ and $\mathbb{R}$ at the midpoint is implied from hereon.* Using the new definition, the half-pixel operator used in the right-hand side of Eq. (11) can be alternatively expressed as

$$\exp\left(\pm\frac{\pi i}{N}k\right)\mathbb{R}(k) = \exp\left[\pm\pi i\mathbb{S}(k)\right]. \qquad (14)$$

We can establish the connection from discrete to continuous variable derivative as follows: $\mathbb{S}$ can be considered as a discrete angular frequency vector circularly shifted by $\lfloor N/2 \rfloor$. If we define $k_x(p) = (2\pi/\Delta x)p$ for $p = 0, \ldots, (N-1)$, then

$$\frac{2\pi}{\Delta x}\mathbb{S} = \overline{k}_x, \qquad (15)$$

where the bar over k_x denotes a circular shift by $\lfloor N/2 \rfloor$. Using Eq. (14), Eqs. (4), (8), and (11) can be alternatively expressed as

$$\widetilde{\varphi_x} = i\overline{k}_x\tilde{\varphi}, \qquad (16)$$

$$\widetilde{\varphi_{x,1/2}} = i\overline{k}_x\exp\left(i\frac{\Delta x}{2}\overline{k}_x\right)\tilde{\varphi}, \qquad (17)$$

$$\widetilde{\varphi_{1/2}} = \exp\left(i\frac{\Delta x}{2}\overline{k}_x\right)\tilde{\varphi}, \qquad (18)$$

The $\overline{k}_x$ notation establishes the formal connection with continuous variable derivatives.

In many practical situations, the band-limited calculations may not produce exact results, depending on the nature of signals. The magnitudes of the Fourier coefficients of a linear function, for example, decrease as 1/(spatial frequency), whereas Eq. (16) indicates that the coefficients of the derivative are multiplied by the spatial-frequency vector. Therefore, the highest spatial-frequency coefficient does not vanish, even if N approaches ∞. Therefore, the linear terms are not band limited and need to be treated separately. Equations (16)–(18) form the basis of the following analysis.

Simpson Reconstructor

The analysis in the previous section suggests that an accurate derivative calculation at discrete samples requires the superposition sum of the whole set of samples. This can be done more conveniently in the frequency domain, which results in a band-limited reconstructor with unity frequency response.[11] On the other hand, it is still worthwhile to investigate an improved finite-difference scheme for purely spatial-domain operation. In finite-difference methods involving only a few points, a high degree of accuracy is preserved by distributing the finite difference over both the measured derivative samples (i.e., slopes) and the integrated samples. Denoting the wavefront estimate as $\hat{\varphi}$ and the measured slope as S, one can start from a general finite-difference expression such as

$$\sum_j a(j)\hat{\varphi}(i+j) = \sum_k b(k)S(i+k) \qquad (19)$$

for 1-D problems. Coefficients a and b belong to a specific finite-difference scheme. For example, Southwell[14] showed a reconstructor based on

$$\hat{\varphi}(i+1) - \hat{\varphi}(i) = \frac{\Delta x}{2}\left[S(i) + S(i+1)\right]. \qquad (20)$$

The frequency response of the Southwell reconstructor is low at high spatial frequency. We enhance the frequency response using the Simpson rule, which is

$$\hat{\varphi}(i+1) - \hat{\varphi}(i-1) = \frac{\Delta x}{3}\left[S(i-1) + 4S(i) + S(i+1)\right]. \qquad (21)$$

An iterative wavefront reconstruction based on this scheme will be developed in the next section.

1. Simpson Iterator

Casting the local 1-D Eq. (21) into a least squares form in 2-D, we obtain an error metric (ε) as

$$\varepsilon = \sum_{i,j}\left\{\frac{1}{2\Delta x}\left[\hat{\varphi}(i,j+1) - \hat{\varphi}(i,j-1)\right]\right.$$
$$\left. -\frac{1}{6}\left[S_x(i,j+1) + 4S_x(i,j) + S_x(i,j-1)\right]\right\}^2$$
$$+\left\{\frac{1}{2\Delta y}\left[\hat{\varphi}(i+1,j) - \hat{\varphi}(i-1,j)\right]\right.$$
$$\left. -\frac{1}{6}\left[S_y(i+1,j) + 4S_y(i,j) + S_y(i-1,j)\right]\right\}^2. \qquad (22)$$

Δx and Δy are moved around to the $\hat{\varphi}$ side so that squared terms are in units of slopes. This provides equal weight to the differences in x and y directions on the assumption that the magnitude of slopes is comparable in either direction.

The condition $\partial \varepsilon / \partial \hat{\varphi}(i,j) = 0$ leads to an equation that can be used for the iterative algorithm. It is assumed that phase and slope points are embedded in an arbitrary region. The differentiation of the error metric results in four groups, which are indicated by different colors in Fig. 127.22. Each group can be used in the equation only when all of its elements exist. This strategy is realized by using g parameters as shown in the following:

SIMPSON:

$$g_L\left[\hat{\varphi}(i,j) - \hat{\varphi}(i,j-2)\right] + g_R\left[\hat{\varphi}(i,j) - \hat{\varphi}(i,j+2)\right]$$

$$+ g_U\left(\frac{\Delta x}{\Delta y}\right)^2\left[\hat{\varphi}(i,j) - \hat{\varphi}(i-2,j)\right]$$

$$+ g_D\left(\frac{\Delta x}{\Delta y}\right)^2\left[\hat{\varphi}(i,j) - \hat{\varphi}(i+2,j)\right]$$

$$= g_L\left[S_x(i,j-2) + 4S_x(i,j-1) + S_x(i,j)\right]\frac{\Delta x}{3}$$

$$- g_R\left[S_x(i,j) + 4S_x(i,j+1) + S_x(i,j+2)\right]\frac{\Delta x}{3}$$

$$+ g_U\left(\frac{\Delta x}{\Delta y}\right)^2\left[S_y(i-2,j) + 4S_y(i-1,j) + S_y(i,j)\right]\frac{\Delta y}{3}$$

$$- g_D\left(\frac{\Delta x}{\Delta y}\right)^2\left[S_y(i,j) + 4S_y(i+1,j) + S_y(i+2,j)\right]\frac{\Delta y}{3}$$

$$\equiv \Delta S(i,j). \tag{23}$$

g_L, g_R, g_U, and g_D are flags with values 0 or 1, where L, R, U, and D indicate left, right, up, and down directions, respectively. They are 0's if the quantities in the parentheses next to them are incalculable or 1's otherwise. For example, g_L at the point (i,j) does not vanish only when the slopes' measurements exist at the additional points at $(i,j-1)$ and $(i,j-2)$. The scope of each flag is graphically indicated in Fig. 127.22. For comparison, the iterative equation for the Southwell reconstructor is written here using the same format.
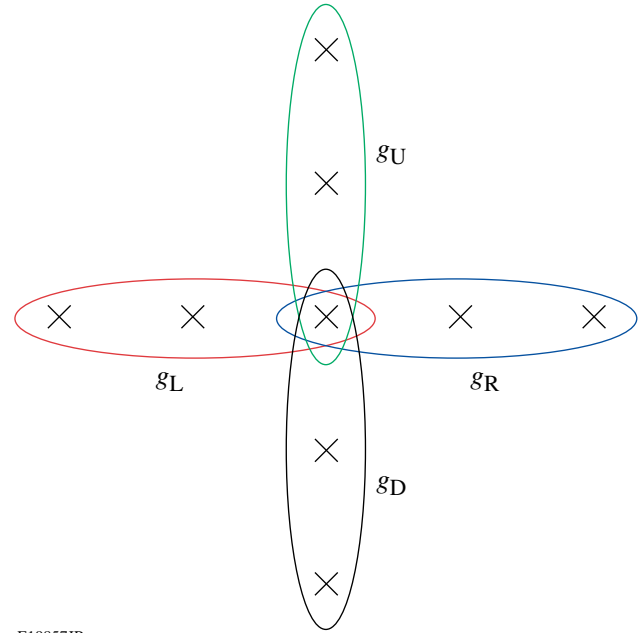
E19857JR

Figure 127.22
Simpson iterator geometry.

SOUTHWELL:

$$g_L\left[\hat{\varphi}(i,j) - \hat{\varphi}(i,j-1)\right] + g_R\left[\hat{\varphi}(i,j) - \hat{\varphi}(i,j+1)\right]$$

$$+ g_U\left(\frac{\Delta x}{\Delta y}\right)^2\left[\hat{\varphi}(i,j) - \hat{\varphi}(i-1,j)\right]$$

$$+ g_D\left(\frac{\Delta x}{\Delta y}\right)^2\left[\hat{\varphi}(i,j) - \hat{\varphi}(i+1,j)\right]$$

$$= g_L\left[S_x(i,j-1) + S_x(i,j)\right]\frac{\Delta x}{2}$$

$$- g_R\left[S_x(i,j) + S_x(i,j+1)\right]\frac{\Delta x}{2}$$

$$+ g_U\left(\frac{\Delta x}{\Delta y}\right)^2\left[S_y(i-1,j) + S_y(i,j)\right]\frac{\Delta y}{2}$$

$$- g_D\left(\frac{\Delta x}{\Delta y}\right)^2\left[S_y(i,j) + S_y(i+1,j)\right]\frac{\Delta y}{2}. \tag{24}$$

The same successive-over-relaxation technique[15] can be applied to the Simpson iterative reconstructor:

$$\hat{\varphi}^{(m+1)}(i,j) = \hat{\varphi}^{(m)}(i,j) + \omega\left[\overline{\hat{\varphi}^{(m)}}(i,j) - \hat{\varphi}^{(m)}(i,j)\right], \tag{25}$$

where

$$\overline{\hat{\varphi}^{(m)}}(i,j)$$

$$= \left[\overline{\hat{\varphi}_0^{(m)}}(i,j) + \Delta S(i,j)\right] \Big/ \left(g_L + g_R + g'_U + g'_D\right), \quad (26)$$

$$\overline{\hat{\varphi}_0^{(m)}}(i,j) = g_L \hat{\varphi}^{(m)}(i,j-2)$$

$$+ g_R \hat{\varphi}^{(m)}(i,j+2) + g'_U \hat{\varphi}^{(m)}(i-2,j)$$

$$+ g'_D \hat{\varphi}^{(m)}(i+2,j). \quad (27)$$

Here, ω is the over-relaxation parameter; g' is $(\Delta x/\Delta y)^2 \cdot g$.

2. Frequency Response and Regularization

The frequency response of the Simpson reconstructor will be calculated following the method presented in Ref. 11. The sum of the squared error in the spatial domain in Eq. (22) is equivalent to the sum of the squared error of the Fourier-transformed component by the Parseval theorem:

$$\varepsilon = \frac{1}{N^2}\sum\left\{\left|D_x\tilde{\varphi} - A_x\widetilde{S_x}\right|^2 + \left|D_y\tilde{\varphi} - A_y\widetilde{S_y}\right|^2\right\}, \quad (28)$$

where

$$D_x = \frac{1}{2\Delta x}\left[\exp(i\overline{k_x}\Delta x) - \exp(-i\overline{k_x}\Delta x)\right] \equiv D_{x,\text{Simpson}} \quad (29)$$

and

$$A_x = \frac{1}{6}\left[\exp(i\overline{k_x}\Delta x) + 4 + \exp(-i\overline{k_x}\Delta x)\right] \equiv A_{x,\text{Simpson}}. \quad (30)$$

The solution for $\tilde{\varphi}$ in Eq. (28) is

$$\tilde{\varphi} = \frac{\left(D_x^* A_x \tilde{S}_x + D_y^* A_y \tilde{S}_y\right)}{|D_x|^2 + |D_y|^2}. \quad (31)$$

As $\tilde{S}_x = i\overline{k_x}\tilde{\varphi}$ and $\tilde{S}_y = i\overline{k_y}\tilde{\varphi}$, the frequency response H defined as the ratio of the reconstructed wavefront amplitude to the true wavefront amplitude associated with the measured slopes at a given spatial frequency point is

$$H = \frac{\tilde{\varphi}}{\tilde{\varphi}} = \frac{\left(D_x^* A_x D_{x,0} + D_y^* A_y D_{y,0}\right)}{|D_x|^2 + |D_y|^2}, \quad (32)$$

where new notations $D_{x,0}$ and $D_{y,0}$ were introduced in place of $i\overline{k_x}$ and $i\overline{k_y}$, respectively.

Applying the Simpson derivative and average operators [Eqs. (29) and (30)], we obtain

$$H_{\text{Simpson}}$$

$$= \frac{1}{3}\frac{\omega_x \sin\omega_x(2 + \cos\omega_x) + \omega_y \sin\omega_y(2 + \cos\omega_y)}{\sin^2\omega_x + \sin^2\omega_y}, \quad (33)$$

where $\overline{k_x}\Delta x \equiv \omega_x = 2\pi f_x$ and $\overline{k_y}\Delta x \equiv \omega_y = 2\pi f_y$, f_x and f_y are normalized frequencies ranging from -0.5 to 0.5. It is assumed $\Delta x = \Delta y$. The frequency response of H_{Simpson} has eight singularities on the four corners and side centers. Except for the region near the poles, the frequency response is nearly unity everywhere, which proves higher accuracy of the Simpson rule than the traditional reconstructors over all spectrums in wavefront reconstruction (refer to Fig. 1 of Ref. 11).

The singularities can be removed by introducing the following Phillips regularization term[16] to the right-hand side of Eq. (22):

$$\varepsilon_{\text{reg}} = \lambda \sum_{i,j}\left\{\frac{1}{(2\Delta x)^2}\left[\hat{\varphi}(i,j+1) - 2\hat{\varphi}(i,j) + \hat{\varphi}(i,j-1)\right]^2\right.$$

$$\left. + \frac{1}{(2\Delta y)^2}\left[\hat{\varphi}(i+1,j) - 2\hat{\varphi}(i,j) + \hat{\varphi}(i-1,j)\right]^2\right\}. \quad (34)$$

In the frequency domain, this transforms into

$$\varepsilon_{\text{reg}} = \lambda\frac{1}{N^2}\sum\left\{\left|D_{x,\text{reg}}\tilde{\varphi}\right|^2 + \left|D_{y,\text{reg}}\tilde{\varphi}\right|^2\right\}, \quad (35)$$

where

$$D_{x,\text{reg}} = \frac{2}{\Delta x}\sin^2\left(\frac{\overline{k_x}\Delta x}{2}\right), \quad (36)$$

and similarly for $D_{y,\text{reg}}$.

The denominator of the Simpson frequency response will have an additional term of $\lambda\left(|D_{x,\text{reg}}|^2 + |D_{y,\text{reg}}|^2\right)$ that removes the singularity. The regularized frequency response is

$$H_{\text{Simpson}}(\lambda)$$

$$= \frac{1}{3}\frac{\omega_x\sin\omega_x(2+\cos\omega_x) + \omega_y\sin\omega_y(2+\cos\omega_y)}{\sin^2\omega_x + \sin^2\omega_y + 4\lambda\left(\sin^4\frac{1}{2}\omega_x + \sin^4\frac{1}{2}\omega_y\right)}. \quad (37)$$

The 1-D frequency response with the regularization term has a second peak near high spatial frequency for sufficiently small λ (<0.08) [Fig. 127.23(b)]. The free parameter λ can be fixed to a value such that the second peak is 1. The numerically determined value of λ for such a condition is 0.07489. This choice of λ gives only a 3% error in wavefront amplitude over 80% of the frequency range. Another choice can be $\lambda = 0.07026$, which balances the local maximum and minimum around 1. The second option reduces the maximum deviation below 2.2% within 85% of the spectral range. Figure 127.23(a) shows a 3-D view of the frequency response of the Simpson-rule reconstructor with $\lambda = 0.07489$. The 1-D response is shown in Fig. 127.23(b). The solid line was calculated from an analytic expression [Eq. (36)], whereas the circles are from numerical simulations. The numerical simulation consists of steps of generating slopes from sinusoid wavefronts at a given spatial frequency and of reconstructing the wavefront and comparing the ratio between the original and the reconstructed wavefront amplitude at that frequency. The reconstruction algorithm used in the simulation will be explained in detail in the following section. The result shows good agreement with the analytic curve.

3. Iterative Algorithm with Regularization Terms

The frequency-domain analysis does not give a detailed picture of how the successive-over-relaxation method can be applied in spatial-domain iteration, especially around the measurement boundary. Resolving the stationary condition with the regularization term gives additional terms on the left-hand side of Eq. (23). These are fully written out using g flags:

$$
\begin{aligned}
(2\Delta x)^2 & \frac{\partial \varepsilon_{\text{reg}}}{\partial \hat{\varphi}(i,j)} \\
= & \ \lambda g_{\text{L}}\left[\hat{\varphi}(i,j) - 2\hat{\varphi}(i,j-1) + \hat{\varphi}(i,j-2)\right] \\
& + \lambda g_{\text{R}}\left[\hat{\varphi}(i,j+2) - 2\hat{\varphi}(i,j+1) + \hat{\varphi}(i,j)\right] \\
& + \lambda g'_{\text{U}}\left[\hat{\varphi}(i,j) - 2\hat{\varphi}(i-1,j) + \hat{\varphi}(i-2,j)\right] \\
& + \lambda g'_{\text{D}}\left[\hat{\varphi}(i+2,j) - 2\hat{\varphi}(i+1,j) + \hat{\varphi}(i,j)\right] \\
& - 2\lambda g'_{\text{LR}}\left[\hat{\varphi}(i,j+1) - 2\hat{\varphi}(i,j) + \hat{\varphi}(i,j-1)\right] \\
& - 2\lambda g'_{\text{UD}}\left[\hat{\varphi}(i+1,j) - 2\hat{\varphi}(i,j) + \hat{\varphi}(i-1,j)\right]. \quad (38)
\end{aligned}
$$

g_{LR} or g_{UD} is 1 only if two points exist to the right and left or up and down, respectively, and zero otherwise. The iteration formula (26) will be modified to

$$
\overline{\hat{\varphi}}(i,j) = \frac{\left[\overline{\hat{\varphi}}_0(i,j) + \Delta S(i,j) + \lambda \Delta \hat{\varphi}(i,j)\right]}{\left[(1+\lambda)(g_{\text{L}} + g_{\text{R}} + g'_{\text{U}} + g'_{\text{D}}) + 4\lambda(g_{\text{LR}} + g'_{\text{UD}})\right]}, \quad (39)
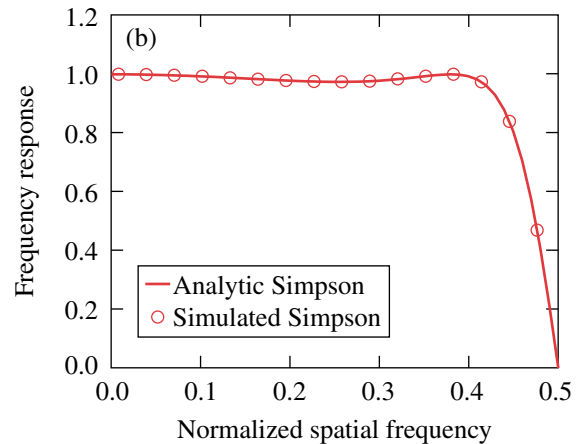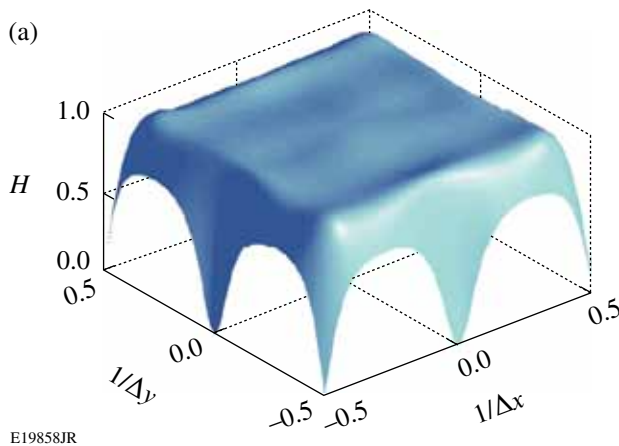$$

E19858JR

Figure 127.23
Frequency response of Simpson reconstructor with $\lambda = 0.07489$. (a) A 3-D view of the frequency response; (b) cross section along f_x axis. The solid line was calculated from the analytic expression; the circles are from simulations of Simpson iterator geometry.

where

$$\Delta\hat{\varphi}(i,j) = g_L\left[2\hat{\varphi}(i,j-1) - \hat{\varphi}(i,j-2)\right]$$
$$+ g_R\left[2\hat{\varphi}(i,j+1) - \hat{\varphi}(i,j+2)\right]$$
$$+ g'_U\left[2\hat{\varphi}(i-1,j) - \hat{\varphi}(i-2,j)\right]$$
$$+ g'_D\left[2\hat{\varphi}(i+1,j) - \hat{\varphi}(i+2,j)\right]$$
$$+ 2g_{LR}\left[\hat{\varphi}(i,j+1) + \hat{\varphi}(i,j-1)\right]$$
$$+ 2g'_{UD}\left[\hat{\varphi}(i+1,j) + \hat{\varphi}(i-1,j)\right]. \quad (40)$$

Hexagonal Band-Limited Reconstructor

Band-limited reconstruction provides a unity frequency response over all spatial bandwidths. The band-limited reconstructor for the Southwell geometry was presented in Ref. 11. It was shown that band-limited derivative operators are also available for Hudgin and Fried geometries. Table 127.II summarizes the three operators depending on the geometry.

Table 127.II: Summary of band-limited derivative operators ($D_{x,0}$).

Geometry	$D_{x,0}$
Southwell	$i\overline{k_x}$
Hudgin	$i\overline{k_x}\exp\left[i\overline{k_x}(\Delta x/2)\right]$
Fried	$i\overline{k_x}\exp\left[i\overline{k_x}(\Delta x/2) + i\overline{k_y}(\Delta y/2)\right]$

Here we present band-limited reconstructors for hexagonal arrays. Hexagonal geometry may be well suited for adaptive optic systems for large telescopes with hexagonal mirror arrays (e.g., James Webb).[17] Large deformable mirrors used in some laser fusion facilities (National Ignition Facility)[18] also have hexagonal actuator patterns. The number density of lenslets is slightly higher in hexagonal geometry than square. Figure 127.24 shows two possible hexagonal arrays. In Fig. 127.24(a) the unit hexagon is lying on its facet, whereas in Fig. 127.24(b) the unit hexagon is standing on the apex. The circles indicate the measurement points and the ×'s are reconstruction points. In Fig. 127.24(a) geometry, the band-limited derivative calculation for the indicated square array involves first grouping the slopes measured at red- and black-circled positions. Marking them as index 1 and 2, respectively, the DFT's of slopes at the reconstruction points are

$$\widetilde{S_x} = \begin{bmatrix} \widetilde{S_{1,x}} + T_{hex} \circ \widetilde{S_{2,x}} \\ \widetilde{S_{1,x}} - T_{hex} \circ \widetilde{S_{2,x}} \end{bmatrix} \equiv \begin{bmatrix} \widetilde{S_{x,I}} \\ \widetilde{S_{x,II}} \end{bmatrix}. \quad (41)$$
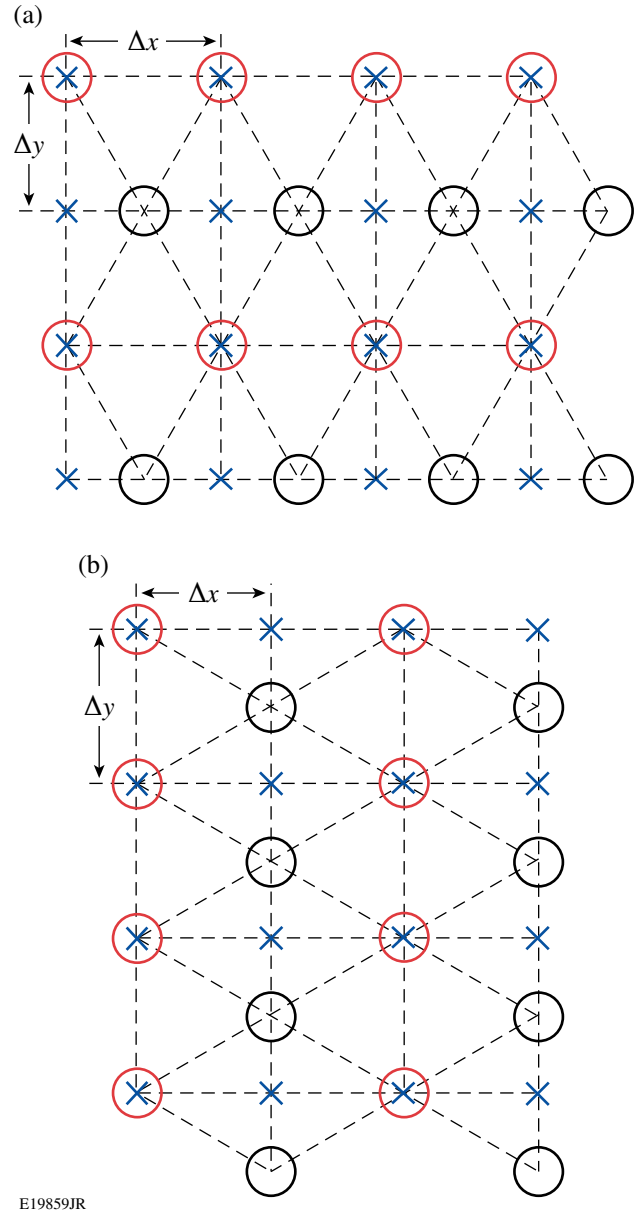
E19859JR

Figure 127.24
(a) Prostrate hexagon array; (b) standing hexagon array.

T_{hex} is a matrix whose size is M by N (i.e., the size of either array 1 or array 2) and "∘" denotes entry-wise matrix multiplication. The pth row and qth column element of T_{hex} is

$$T_{hex}(p,q) = \exp\left[-\frac{\pi i}{M}p - \pi i\, \mathbb{S}(q)\right]. \quad (42)$$

The combined total array is therefore a vertical concatenation of the two matrices. On the other hand, the resulting total matrix for Fig. 127.24(b) geometry is a horizontal concatenation:

$$\widetilde{S_x} = \left[\widetilde{S_{1,x}} + T_{\text{hex}} \circ \widetilde{S_{2,x}} \quad \widetilde{S_{1,x}} - T_{\text{hex}} \circ \widetilde{S_{2,x}}\right]$$

$$\equiv \left[\widetilde{S_{x,\text{I}}} \quad \widetilde{S_{x,\text{II}}}\right]. \tag{43}$$

$$T_{\text{hex}}(p,q) = \exp\left[-\pi i \mathbb{S}(p) - \frac{\pi i}{N} q\right]. \tag{44}$$

The same combination rule applies to *y*-slope measurements.

The above decomposition technique can be inverted such that each subgroup of the hexagonal array can also be expressed as the linear sum of blocks I and II of the rectangular array. This inversion is used only for wavefront points in the algorithm, which is

$$\widetilde{\varphi_1} = \frac{1}{2}\left(\widetilde{\varphi_{\text{I}}} + \widetilde{\varphi_{\text{II}}}\right), \tag{45}$$

$$\widetilde{\varphi_2} = \frac{1}{2} T_{\text{hex}}^* \circ \left(\widetilde{\varphi_{\text{I}}} + \widetilde{\varphi_{\text{II}}}\right). \tag{46}$$

Using the basic results obtained in **Band-Limited Derivative** (p. 130) and the DFT procedures for the hexagonal arrays in this section, the band-limited reconstruction algorithm for hexagonal slope arrays can be implemented as shown in the flowchart in Fig. 127.25.

Step 1 consists of fitting the slopes over low-order polynomials, e.g., third order, which will significantly reduce non-band-limited components of the wavefront. If the regions of interest are disconnected, the fitting must be performed per each region. Owing to the sum requirement [Eq. (3)], a column and row are appended to the edge of the measured slope matrices (groups 1 and 2 separately), which will satisfy the zero-sum conditions in the *x* and *y* directions.

Step 2 initializes the slopes with measured values. Steps 3–8 form a closed loop required for extrapolating slopes outside the non-rectangular region. The iteration is not required if the region is rectangular.

Slopes in groups 1 and 2 are separately Fourier transformed using Eqs. (40)–(43) in Step 3. In Step 4, wavefront matrices corresponding to each block (I or II) are reconstructed in the Fourier domain using the band-limited filter function, which is
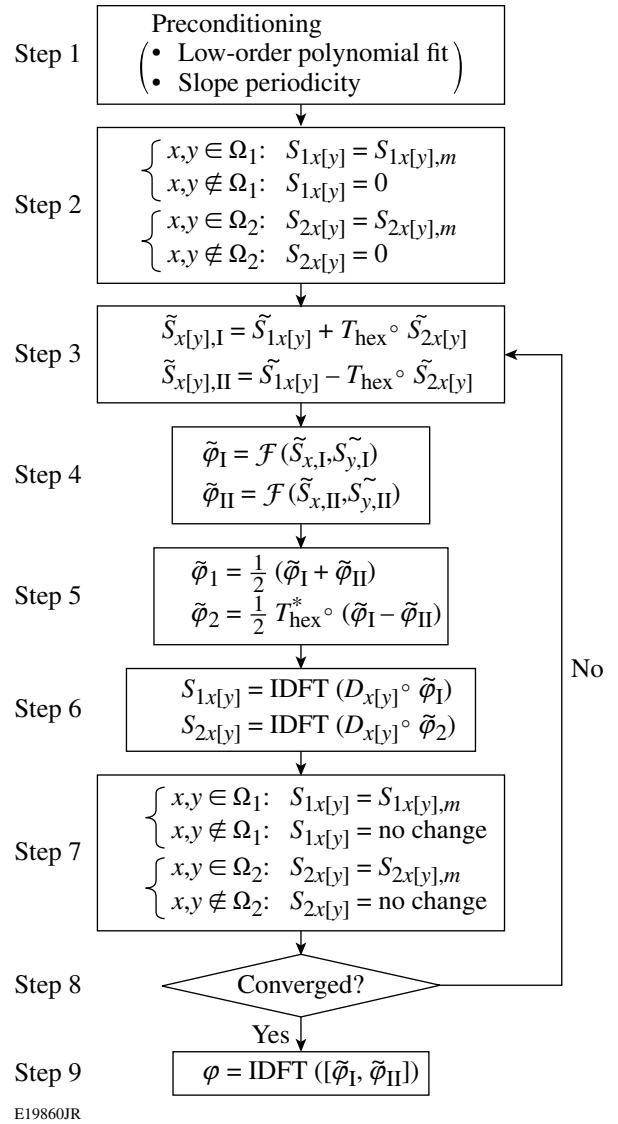

Step 1 — Preconditioning
- Low-order polynomial fit
- Slope periodicity

Step 2 —
$$\begin{cases} x,y \in \Omega_1: & S_{1x[y]} = S_{1x[y],m} \\ x,y \notin \Omega_1: & S_{1x[y]} = 0 \end{cases}$$
$$\begin{cases} x,y \in \Omega_2: & S_{2x[y]} = S_{2x[y],m} \\ x,y \notin \Omega_2: & S_{2x[y]} = 0 \end{cases}$$

Step 3 —
$$\tilde{S}_{x[y],\text{I}} = \tilde{S}_{1x[y]} + T_{\text{hex}} \circ \tilde{S}_{2x[y]}$$
$$\tilde{S}_{x[y],\text{II}} = \tilde{S}_{1x[y]} - T_{\text{hex}} \circ \tilde{S}_{2x[y]}$$

Step 4 —
$$\tilde{\varphi}_{\text{I}} = \mathcal{F}(\tilde{S}_{x,\text{I}}, \tilde{S_{y,\text{I}}})$$
$$\tilde{\varphi}_{\text{II}} = \mathcal{F}(\tilde{S}_{x,\text{II}}, \tilde{S_{y,\text{II}}})$$

Step 5 —
$$\tilde{\varphi}_1 = \frac{1}{2}(\tilde{\varphi}_{\text{I}} + \tilde{\varphi}_{\text{II}})$$
$$\tilde{\varphi}_2 = \frac{1}{2} T_{\text{hex}}^* \circ (\tilde{\varphi}_{\text{I}} - \tilde{\varphi}_{\text{II}})$$

Step 6 —
$$S_{1x[y]} = \text{IDFT}(D_{x[y]} \circ \tilde{\varphi}_{\text{I}})$$
$$S_{2x[y]} = \text{IDFT}(D_{x[y]} \circ \tilde{\varphi}_2)$$

Step 7 —
$$\begin{cases} x,y \in \Omega_1: & S_{1x[y]} = S_{1x[y],m} \\ x,y \notin \Omega_1: & S_{1x[y]} = \text{no change} \end{cases}$$
$$\begin{cases} x,y \in \Omega_2: & S_{2x[y]} = S_{2x[y],m} \\ x,y \notin \Omega_2: & S_{2x[y]} = \text{no change} \end{cases}$$

Step 8 — Converged? — No (loop back to Step 3) / Yes

Step 9 — $\varphi = \text{IDFT}([\tilde{\varphi}_{\text{I}}, \tilde{\varphi}_{\text{II}}])$

E19860JR

Figure 127.25
Flowchart of band-limited reconstruction for a hexagonal geometry. $\mathcal{F}$ is the band-limited filter function [Eq. (46)]. The "*m*" subscript denotes the measured slopes. IDFT stands for inverse discrete Fourier transform. Ω_1 and Ω_2 are the regions where the slopes' groups 1 and 2 data exist.

$$\mathcal{F}\left[\widetilde{S_x}(p,q), \widetilde{S_y}(p,q)\right]$$

$$= \frac{D_x^*(p,q)\widetilde{S_x}(p,q) + D_y^*(p,q)\widetilde{S_y}(p,q)}{\left|D_x(p,q)\right|^2 + \left|D_y(p,q)\right|^2}, \tag{47}$$

where the band-limited derivative operators D_x and D_y are defined as

$$D_x(p,q) = \frac{2\pi i}{\Delta x} \mathbb{S}(q), \qquad (48)$$

$$D_y(p,q) = \frac{2\pi i}{(2\Delta y)} \mathbb{S}(p), \qquad (49)$$

for a prostrate hexagon array and

$$D_x(p,q) = \frac{2\pi i}{(2\Delta x)} \mathbb{S}(q), \qquad (50)$$

$$D_y(p,q) = \frac{2\pi i}{\Delta y} \mathbb{S}(p), \qquad (51)$$

for a standing hexagon array.

Step 5 creates wavefront groups 1 and 2 by using Eqs. (51) and (52). In Steps 3 and 5, the correct T_{hex} must be used according to its geometry. In Step 6, derivative operators are applied to these temporary wavefront matrices to obtain slopes in groups 1 and 2, respectively. These new slopes are different from the measured slopes. We leave the values external to the boundary untouched while restoring the internal values to the original measured slopes. The difference between the measured slopes and the calculated slopes decreases over the course of iterations. Step 8 determines whether this difference is within tolerance. Once the convergence criterion is met, the wavefront matrices generated in Step 4 $(\widetilde{\varphi_{\text{I}}}, \widetilde{\varphi_{\text{II}}})$ are combined to form a single matrix by either vertical or horizontal concatenation, depending on the hexagon geometry and inverse Fourier transformed to the spatial domain to produce the final result in Step 9. Small terms in the imaginary part of the solution can be neglected.

The band-limited algorithms shown in Fig. 2 of Ref. 11 and Fig. 127.25 can be used together with a non-band-limited filter function, which enables one to conveniently switch between different algorithms. The reconstruction algorithms proposed here are not limited to a specific boundary shape.

Error Propagation

The wavefront reconstructors have traditionally been characterized with a so-called error propagation curve. This indicates the sensitivity of the noise in the reconstructed phase to the noise in the slopes measurements. Early numerical and theoretical works show that this sensitivity is a logarithmic function of the number of measurement points.[1,2,15] Simulations confirm this. The noise-propagation coefficient will be calculated using discrete samples and frequency-domain filter functions. We limit the scope to the rectangular area.

Let σ_φ be the root mean square of the reconstructed phase φ. According to the Wiener–Khintchin theorem,

$$\sigma_\varphi^2 = \frac{1}{N_t} \sum \langle \Delta \varphi^2 \rangle = \frac{1}{N_t N^2} \sum \langle |\widetilde{\Delta \varphi}|^2 \rangle, \qquad (52)$$

where $\langle \cdot \rangle$ denotes ensemble average of the quantity inside.

According to linear stochastic system theory, the power spectrum of input and output signals is related by the absolute square of the linear system function. In the case of wavefront reconstruction dictated by the linear response

$$\tilde{\varphi} = \frac{\left(D_x^* A_x \widetilde{S_x} + D_y^* A_y \widetilde{S_y} \right)}{|D_x|^2 + |D_y|^2 + \lambda \left(|D_{x,\text{reg}}|^2 + |D_{y,\text{reg}}|^2 \right)}, \qquad (53)$$

the corresponding stochastic response in power spectrum is

$$\langle |\widetilde{\Delta \varphi}|^2 \rangle$$
$$= \frac{\left(|D_x A_x \Delta \widetilde{S_x}|^2 + |D_y A_y \Delta \widetilde{S_y}|^2 \right)}{\left[|D_x|^2 + |D_y|^2 + \lambda \left(|D_{x,\text{reg}}|^2 + |D_{y,\text{reg}}|^2 \right) \right]^2}. \qquad (54)$$

Assuming that $\Delta \widetilde{S_x}$ and $\Delta \widetilde{S_y}$ are uncorrelated white noise with a variance of $\tilde{\sigma}_S^2$ for each, and since $\sigma_S^2 = (1/N)\tilde{\sigma}_S^2$, the noise-propagation coefficient is

$$\eta = \frac{\sigma_\varphi^2}{h^2 \sigma_S^2}$$
$$= \frac{1}{L^2} \sum_{k_x, k_y} \frac{\left(|D_x A_x|^2 + |D_y A_y|^2 \right)}{\left[|D_x|^2 + |D_y|^2 + \lambda \left(|D_{x,\text{reg}}|^2 + |D_{y,\text{reg}}|^2 \right) \right]^2}, \qquad (55)$$

where $\sigma_\varphi^2 = (1/N_t) \langle |\widetilde{\Delta \varphi}|^2 \rangle$, $h = \Delta x = \Delta y$, and L is the aperture size. This result is equivalent to Noll's[19] in the case of band-limited operators.

Table 127.III summarizes finite-difference derivative/averaging operators for four geometries to be used with Eq. (55).

The right-hand side of Eq. (50) is inversely proportional to $|D_{x,0}|^2 + |D_{y,0}|^2$ for band-limited reconstruction and is difficult to visualize in linear scale. We define a "noise-response function (S_N)" with the inverse power dependence removed as follows:

$$S_N = \frac{\left(|D_{x,0}|^2 + |D_{y,0}|^2\right)\left(|D_x A_x|^2 + |D_y A_y|^2\right)}{\left[|D_x|^2 + |D_y|^2 + \lambda\left(|D_{x,\text{reg}}|^2 + |D_{y,\text{reg}}|^2\right)\right]^2}. \quad (56)$$

It can be shown by the Cauchy–Schwartz inequality that the noise response is always larger than or equal to the absolute frequency response squared,

$$|H|^2 \leq S_N. \quad (57)$$

The inequality [Eq. (52)] shows that the error propagation is intimately related to the frequency response of a reconstructor. The lower bound of η is

$$\sum_{k_x, k_y} \frac{|H|^2}{k_x^2 + k_y^2} \leq \sum_{k_x, k_y} \frac{S_N}{k_x^2 + k_y^2} = \eta. \quad (58)$$

From this, one can expect that the Southwell reconstructor will have the lowest lower bound and the Fried reconstructor the highest. It agrees with the result of Zou.[20]

The analytic expression for η can be calculated and fit to a logarithmic curve, although the logarithm dependence is only approximate except for the band-limited reconstructors. The result is summarized in the second column of Table 127.IV. Singularity points were excluded in the summation over spatial-frequency space. The third column shows the simulated η obtained by running actual reconstructors with zero slopes input with Gaussian noise. Two hundred realizations were performed at each N, where N^2 is the number of points. N was varied from 10 to 100 by 10. The logarithm fit over the averaged η is shown in the column. The multiplicative coefficients roughly agree with the analytic ones up to the second decimal point, but the additive constants from simulation are always estimated higher than the calculated ones. The offset is about 0.2771 on average. The discrepancy appears to come from the apparent inconsistency in assuming white noise in the slopes power spectrum and the use of band-limited derivative formalism. For example, the reconstructed wavefront from white spectrum noise always has some amount of low-order polynomial terms, which cannot be represented by Eq. (48). The constant offset 0.2771 therefore can be considered as the ratio of energy conversion from white noise to non-band-limited signals.

The legacy formulas of noise propagation for each reconstructor are also shown in the fourth column of Table 127.IV, quoted from the three authors' original publications.[1,2,15] The quoted Southwell η is estimated only from the graph in the original paper since no explicit formula was given. Noll's calculation essentially corresponds to the band-limited case. Considering the fact that there is some ambiguity in the determination of the constant offset, at least the multiplicative coefficient of the Fried formula comes close to our analytic result; whereas there

Table 127.III: Summary of frequency-domain equivalents of the associated finite-difference schemes.

Finite-difference scheme	D_x	A_x
Southwell[14]	$\frac{1}{\Delta x}\left[\exp(i\overline{k_x}\Delta x) - 1\right]$	$\frac{1}{2}\left[\exp(i\overline{k_x}\Delta x) + 1\right]$
Hudgin[2]	$\frac{1}{\Delta x}\left[\exp(i\overline{k_x}\Delta x) - 1\right]$	1
Fried[1]	$\frac{1}{2\Delta x}\left[\exp(i\overline{k_x}\Delta x) - 1\right]\left[\exp(i\overline{k_y}\Delta y) + 1\right]$	1
Simpson	$\frac{1}{2\Delta x}\left[\exp(i\overline{k_x}\Delta x) - \exp(-i\overline{k_x}\Delta x)\right]$	$\frac{1}{6}\left[\exp(i\overline{k_x}\Delta x) + 4 + \exp(-i\overline{k_x}\Delta x)\right]$

Table 127.IV: Summary of noise propagation.

Numerical Scheme	Calculated η	Simulated η	Quoted η
Southwell difference	$-0.1211 + 0.1591 \ln N$	$0.1356 + 0.1638 \ln N$	$-0.1237 + 0.3164 \ln N$
Hudgin difference	$0.0485 + 0.1592 \ln N$	$0.3456 + 0.1528 \ln N$	$0.561 + 0.103 \ln N$
Fried difference	$-0.0865 + 0.3202 \ln N$	—	$0.6558 + 0.3206 \ln N$
Band limited	$-0.0285 + 0.1594 \ln N$	$0.2449 + 0.1615 \ln N$	$0.1072 + 0.318 \ln N$
Simpson difference	$0.0314 + 0.1610 \ln N$	$0.3125 + 0.1647 \ln N$	—

is about a factor-of-2 difference in the Southwell and Noll's expression compared with ours. On the other hand, Hudgin's formula does not agree with our results. Fried's formula is based on a comparatively large number of N (≤ 39) compared with Southwell and Hudgin's calculations ($N \leq 20$).

Conclusion

We have presented derivations of band-limited derivative operators in the frequency domain. These are important tools for characterizing and improving the frequency response of wavefront reconstructors over broad bandwidth. Two new wavefront reconstructors were proposed utilizing these tools. The reconstructors were designed to be accurate up to high spatial frequency. The first one is based on the Simpson integration rule. The bandwidth of the frequency response of this reconstructor, after being regularized, is excellent up to 85% of the spatial frequency range. A successive-over-relaxation iterative solver was presented in detail, where the outermost samples are elegantly handled using g flags. The frequency-response behavior of the iterative solver agrees well with the predicted frequency-response curve. The second reconstructor is an extension of the band-limited reconstruction algorithm previously developed; the measurement points are on a hexagonal array instead of a rectangular array. A Fourier-domain iterative algorithm was proposed for two types of hexagonal arrays. As was previously pointed out in Ref. 11, the reconstruction process must be preconditioned with the low-order polynomial fit. The Simpson-rule–based algorithm works purely in the spatial domain; therefore, it is computationally less complex than band-limited algorithms, whereas the latter provides flexibility against any geometry change. Fourier-domain algorithms have a potential of boosting reduction speed with the help of digital-signal processors.

The new wavefront reconstructors are compared with the traditional reconstructors in terms of noise-propagation properties through a generalized noise-propagation expression. The analytically calculated noise-propagation coefficients are consistent with the numerical fit deduced from our own simulations. We did not find, however, universal agreement with the published results.

The broad-bandwidth wavefront reconstructors developed here are used in wavefront-reduction software to characterize focal spots of the OMEGA EP laser beams.[13] The importance of the band-limited reconstructor was well illustrated in Ref. 21 for a closed-loop wavefront-shaping application. One may also find applications in the study of metrology and atmospheric turbulence.[22]

REFERENCES

1. D. L. Fried, J. Opt. Soc. Am. **67**, 370 (1977).

2. R. H. Hudgin, J. Opt. Soc. Am. **67**, 375 (1977).

3. K. R. Freischlad and C. L. Koliopoulos, J. Opt. Soc. Am. A **3**, 1852 (1986).

4. F. Roddier and C. Roddier, Appl. Opt. **30**, 1325 (1991).

5. L. A. Poyneer, D. T. Gavel, and J. M. Brase, J. Opt. Soc. Am. A **19**, 2100 (2002).

6. L. A. Poyneer and B. Macintosh, J. Opt. Soc. Am. A **21**, 810 (2004).

7. A. Dubra, C. Paterson, and C. Dainty, Appl. Opt. **43**, 1108 (2004).

8. C. Elster and I. Weingartner, in *X-Ray Mirrors, Crystals, and Multilayers II*, edited by A. K. Freund *et al.* (SPIE, Bellingham, WA, 2002), Vol. 4782, pp. 12–20.

9. J. Campos *et al.*, Opt. Lett. **27**, 1986 (2002).

10. L. Yaroslavsky, A. Moreno, and J. Campos, Opt. Express **13**, 2892 (2005).

11. S.-W. Bahk, Opt. Lett. **33**, 1321 (2008).

12. M. Servin, D. Malacara, and J. L. Marroquin, Appl. Opt. **35**, 4343 (1996).

13. J. Bromage, S.-W. Bahk, D. Irwin, J. Kwiatkowski, A. Pruyne, M. Millecchia, M. Moore, and J. D. Zuegel, Opt. Express **16**, 16,561 (2008).

14. R. Marks, II and M. W. Hall, IEEE Trans. Audio, Speech, Lang. Process. **29**, 872 (1981).

15. W. H. Southwell, J. Opt. Soc. Am. **70**, 998 (1980).

16. D. L. Phillips, J. ACM **9**, 84 (1962).

17. See http://www.jwst.nasa.gov/.

18. S. E. Winters, J. H. Chung, and S. A. Velinsky, J. Dyn. Syst. Meas. Contr. **124**, 297 (2002).

19. R. J. Noll, J. Opt. Soc. Am. **68**, 139 (1978).

20. W. Zou and J. P. Rolland, J. Opt. Soc. Am. A **23**, 2629 (2006).

21. S.-W. Bahk, E. Fess, B. E. Kruschwitz, and J. D. Zuegel, Opt. Express **18**, 9151 (2010).

22. T. W. Nicholls, G. D. Boreman, and J. C. Dainty, Opt. Lett. **20**, 2460 (1995).