Design and assessment of symmetrical phase-shifting algorithms

K. G. Larkin and B. F. Oreb

Division of Applied Physics, Commonwealth Scientific and Industrial Research Organization, Lindfield 2070, Australia

Received April 2, 1992; accepted May 7, 1992

Conventional phase-shifting algorithms based on a least-squares estimate use N samples over an incomplete period of the sampled waveform. We introduce a class of phase-shifting algorithms having $N + 1$ samples symmetrically disposed over one full period of the sampled waveform. Fourier analysis techniques are used to derive these algorithms and modify them to improve their performance in the presence of phase-shift errors. The algorithms can be used in phase measurement systems having periodic, but not necessarily sinusoidal, waveforms.

1. INTRODUCTION

Measurement of wave-front phase in an interference pattern by phase-shifting techniques has been applied to many areas of precision metrology in recent years.[1-3] Many phase-shifting algorithms that perform well with sinusoidal or near-sinusoidal periodic waveforms have been developed.[4,5] However, in measurement systems that result in nonsinusoidal periodic waveforms (either by choice or by imperfections such as phase-shift errors, multiple interference beams, nonlinearities in detector), the performance of existing algorithms is often inadequate.

We deal with the design and assessment of a class of phase-shifting algorithms that can be used to evaluate the phase of the wave front in systems with systematic errors such as nonsinusoidal periodic waveforms or phase-shift errors. The algorithms can be tailored to result in small errors in systems having specific nonsinusoidal characteristics. These algorithms, which we have called symmetrical phase-shifting algorithms (SPSA's), can be implemented by either phase-stepping or integrating-bucket techniques to give the phase of the wave front at an array of points. The SPSA is characterized by having $N + 1$ samples separated by N equal intervals over one full period. There is a close correspondence with the well-known N-sample, least-squares estimate algorithms[6-8] that do not extend over the full period of the sampled waveform. Associated with the SPSA's is a design and assessment technique based on Fourier representation introduced by other researchers[9] for analysis and performance assessment of existing algorithms. We have extended these principles and derived a strategy for the design of the SPSA's.

The well-known five-sample algorithm is shown to be an example of the SPSA's. By way of an additional example we present in detail an algorithm having seven samples.

Throughout this paper we adopt the convention of calling the SPSA's $N + 1$ algorithms to distinguish them clearly from conventional N-sample algorithms.

2. FOURIER REPRESENTATION OF PHASE-SHIFTING ALGORITHMS

Consider a periodic nonsinusoidal interference pattern or waveform as shown in Fig. 1(a). The phase of this pattern needs to be evaluated at an array of points—generally a one- or two-dimensional grid. The nonsinusoidal pattern can be created by interferometric techniques or by other means such as the projection of a transmission grating.

Let $g(t)$ be the intensity distribution of the pattern at a designated point (x, y) on a two-dimensional array as a function of t, the shift parameter that can represent a temporal or spatial shift. In general g is a real function of the array point position (x, y) and the parameter t. For simplicity our analysis is mostly confined to a single point, although it is applicable to all other points in the array.

It is important to note that g is a periodic function of t but not necessarily a periodic function of x or y, i.e., $g(t + T_g) = g(t)$, where T_g is the period. In general, $g(t)$ can be written as a Fourier series:

$$g(t) = \sum_{n=0}^{\infty} a_n \cos(2\pi n \nu_g t + \Phi_n), \tag{1}$$

where $\Phi_n = \Phi_n(x, y)$ is the phase of the nth harmonic of the fundamental frequency ν_g in the pattern at a designated position (x, y) in the array, n is a positive integer, a_n is the weight of the nth harmonic (a real number), and $\nu_g = 1/T_g$ is the fundamental frequency.

Since $g(t)$ is a periodic function, its spectrum $G(\nu)$, the Fourier transform[10] (FT) of $g(t)$, is a discrete function with components at frequency $\nu = \pm n\nu_g$. Figure 1(b) shows a typical $|G(\nu)|$. Generally it is the phase Φ_1 of the fundamental frequency ($n = 1$) that needs to be measured, but to keep the analysis general we consider an arbitrary (but fixed) harmonic, $n = m$.

In order to measure Φ_m with phase-shifting techniques, we record a number of samples of the shifted intensity pattern $g(t)$ by the use of phase-stepping or integrating-bucket methods.

0740-3232/92/101740-09$05.00

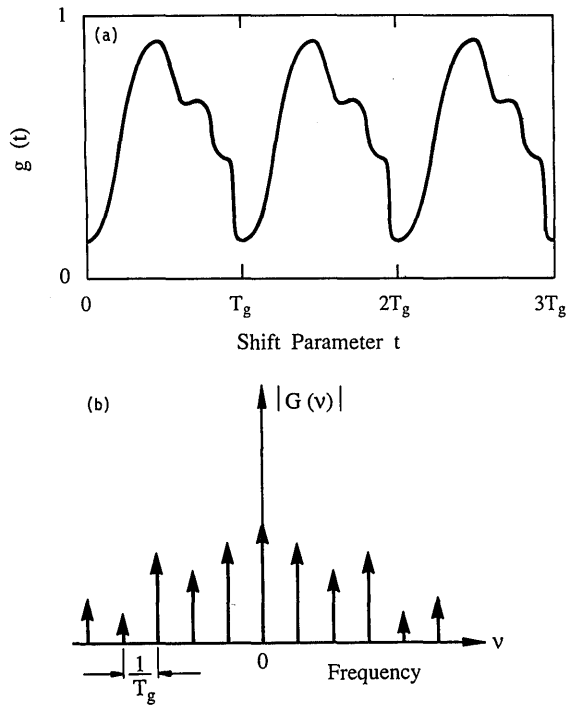

Fig. 1. (a) Periodic interference or grating pattern $g(t)$. (b) $|G(\nu)|$ the modulus of the Fourier spectrum of $g(t)$.

As outlined by Freischlad and Koliopolous,[9] the phase Φ_m can be evaluated by using the principle of quadrature phase detection in heterodyne phase measurements. Here the phase Φ_m is obtained from an arctangent of a function. This function is the ratio of the imaginary part to the real part of the sampled signal's mth Fourier component. The relevant results of the Fourier description of Freischlad and Koliopolous[9] are now presented as the basis for later analysis.

The heterodyne process can be described by correlations[10] of the signal $g(t)$ with two sampling functions $f_1(t)$ and $f_2(t)$ that define the phase-shifting algorithms. Typically $f_1(t)$ and $f_2(t)$ represent a small number of equispaced samples having sine and cosine weighting, respectively. The two correlations $p(t)$ and $q(t)$ are given by

$$p(t) = g \oplus f_1 = \int_{-\infty}^{\infty} g(\tau) f_1(\tau + t) \mathrm{d}\tau, \tag{2}$$

$$q(t) = g \oplus f_2 = \int_{-\infty}^{\infty} g(\tau) f_2(\tau + t) \mathrm{d}\tau. \tag{3}$$

It can be shown that

$$p(0) = \sum_{n=0}^{\infty} \frac{a_n}{2} [\exp(i\Phi_n) F_1^*(n\nu_g) + \exp(-i\Phi_n) F_1^*(-n\nu_g)], \tag{4}$$

$$q(0) = \sum_{n=0}^{\infty} \frac{a_n}{2} [\exp(i\Phi_n) F_2^*(n\nu_g) + \exp(-i\Phi_n) F_2^*(-n\nu_g)], \tag{5}$$

where $i = \sqrt{-1}$, $F_1^*(\nu)$ is the complex conjugate of $F_1(\nu)$, the FT of $f_1(t)$, and $F_2^*(\nu)$ is the complex conjugate of the FT of $f_2(t)$. From Eqs. (4) and (5) it can be seen that both $p(0)$ and $q(0)$ are series related to the values of $F_1(\nu)$ and $F_2(\nu)$ at discrete values of the frequency $\nu = n\nu_g$.

The phase Φ_m can be obtained from the ratio r of $p(0)$ and $q(0)$, i.e.,

$$r = \frac{p(0)}{q(0)}. \tag{6}$$

The above relation can be used to determine Φ_m if the following conditions are satisfied:

$$a_n F_1^*(n\nu_g) = -iA\delta(n, m), \tag{7a}$$

$$a_n F_2^*(n\nu_g) = A\delta(n, m), \tag{7b}$$

where A is a constant (generally complex), $\delta(n, m)$ is the Kronecker delta function, and m and n are integers, i.e.,

$$\delta(n, m) = 1 \quad \text{for } m = n,$$

$$\delta(n, m) = 0 \quad \text{for } m \neq n.$$

In other words it is necessary for the spectra $G(\nu)$ and $F_1(\nu)$ to have only one common frequency component. A similar relation also holds between $G(\nu)$ and $F_2(\nu)$. It follows from Eqs. (7) that the components of $F_1(\nu)$ and $F_2(\nu)$ at $\nu = m\nu_g$ should be 90° out of phase and of equal magnitude,

$$F_1(m\nu_g) = iF_2(m\nu_g). \tag{8}$$

Under these conditions Eq. (6) leads to

$$r = \tan(\Phi_m). \tag{9}$$

3. DEVELOPMENT OF SYMMETRICAL PHASE-SHIFTING ALGORITHMS

Based on the concepts outlined in Section 2, we now develop the characteristics of a class of phase-shifting algorithms that, as far as we know, have not been reported previously.

Consider sampling functions with the following properties:

$$f_1(t) = -f_1(-t), \quad \text{i.e., } f_1(t) \text{ is real and odd}, \tag{10}$$

$$f_2(t) = f_2(-t), \quad \text{i.e., } f_2(t) \text{ is real and even}. \tag{11}$$

Hence $F_1(\nu)$ is imaginary and odd, i.e., $F_1(\nu) = -F_1(-\nu)$, $F_1(0) = 0$; $F_2(\nu)$ is real and even, i.e., $F_2(\nu) = F_2(-\nu)$. It therefore follows that $F_1(\nu)$ and $F_2(\nu)$ can potentially satisfy Eq. (8).

Let us now consider the SPSA's that are based on $N + 1$ samples, evenly spaced over T_f, the total duration of the sampling functions. Ideally, $T_f = T_g$, but the analysis will be kept general at this stage. This is different from the algorithms considered by others[6-8] in that the present algorithms have an additional sample at the end of the period T_f (see Section 5 for examples). Therefore $f_1(t)$ can be written as follows for the phase-stepping case (and similar equations can be written for the integrating-bucket case):

$$f_1(t) = \sum_{n=0}^{N} \alpha_n \delta(t - t_n), \tag{12}$$

where α_n is the (real) coefficient of the nth discrete sample, δ designates the delta function, also known as the impulse response, and

$$t_n = \frac{nT_f}{N} - \frac{T_f}{2} \tag{13}$$

gives the sample positions.

Similarly for $f_2(t)$,

$$f_2(t) = \sum_{n=0}^{N} \beta_n \delta(t - t_n), \tag{14}$$

where β_n is the (real) coefficient of the nth discrete sample. The summation range can be reduced by half when the symmetry properties are used. If we define an integer M such that $M = N/2$ for N even, $M = (N - 1)/2$ for N odd, then, for the $f_1(t)$ given in Eq. (10),

$$\alpha_n = -\alpha_{N-n} \quad \text{for } n = 0\text{–}M, \tag{15}$$

$$\sum_{n=0}^{N} \alpha_n = 0. \tag{16}$$

One should note that $\alpha_M = 0$ when N is even. The FT of $f_1(t)$ is then

$$F_1(\nu) = -2i \sum_{n=0}^{M} \alpha_n \sin(2\pi\nu t_n). \tag{17}$$

Similarly, since $f_2(t)$ is even, as shown in Eq. (11), then

$$\beta_{N-n} = \beta_n \quad \text{for } n = 0\text{–}M. \tag{18}$$

The FT of $f_2(t)$ for N even is then

$$F_2(\nu) = \beta_M \exp(-2\pi i \nu t_M) + 2 \sum_{n=0}^{M-1} \beta_n \cos(2\pi\nu t_n), \tag{19a}$$

or for N odd,

$$F_2(\nu) = 2 \sum_{n=0}^{M} \beta_n \cos(2\pi\nu t_n). \tag{19b}$$

For condition (7b) to be satisfied, in general $F_2^*(0) = 0$, and this leads to

$$F_2(0) = 0. \tag{20}$$

Therefore, from Eqs. (20) and (19) for N even,

$$\beta_M + 2 \sum_{n=0}^{M-1} \beta_n = 0, \tag{21a}$$

or for N odd,

$$2 \sum_{n=0}^{M} \beta_n = 0. \tag{21b}$$

That is to say that the sums of the sample coefficients of both $f_1(t)$ and $f_2(t)$ are zero.

Applying the least-squares fit method of Greivenkamp[7] and Morgan[8] to our $N + 1$ samples over a full period T_f, we can write the sampling coefficients α_n and β_n as follows:

$$\alpha_n = -\sin\left(\frac{2\pi t_n}{T_f}\right) = \sin\left(\frac{2\pi n}{N}\right) \quad \text{for } n = 0\text{–}N, \tag{22}$$

$$\beta_n = \cos\left(\frac{2\pi t_n}{T_f}\right) = -\cos\left(\frac{2\pi n}{N}\right) \quad \text{for } n = 1\text{–}(N - 1), \tag{23}$$

with $\beta_N = \beta_0 = -\frac{1}{2}$.

It is important to note that the results in Eqs. (22) and (23) apply only if a weighted least-squares method is used. In this case the end samples ($n = 0$ and $n = N$) are given weights of 1/2 and all the other samples weights of 1. With this selection of weights the 3×3 matrix of Greivenkamp[7] is diagonal and yields these simple results.

The symmetry properties of α_n and β_n expressed in Eqs. (15) and (16) and (18) and (21), respectively, are encapsulated in Eqs. (22) and (23). Equations (22) and (23) fully define the coefficients for the samples in the SPSA's. The sample positions are given by Eq. (13).

4. FREQUENCY CHARACTERISTICS OF $F_1(\nu)$ and $F_2(\nu)$

The frequency spectra of $f_1(t)$ and $f_2(t)$ are important in determining the performance of a phase-shifting algorithm in the presence of measurement errors such as phase-shift errors. Stetson and Brohinsky[2] have analyzed the spectra of the conventional N-step algorithms. We now discuss some relevant properties of $F_1(\nu)$ and $F_2(\nu)$ for the SPSA's and use these in later sections.

From Eqs. (17), (13), and (22),

$$F_1(\nu) = -2i \sum_{n=1}^{M} \sin\left(\frac{2\pi n}{N}\right) \sin\left[2\pi\nu\left(\frac{nT_f}{N} - \frac{T_f}{2}\right)\right]. \tag{24}$$

This is a continuous, periodic function with a period of N/T_f when N is even, or a period of $2N/T_f$ when N is odd. At the harmonics of the sampling waveform $\nu = m\nu_f$, where $\nu_f = 1/T_f$ and m is an integer, it can be shown from Eq. (24) that

$$\left.\begin{aligned} F_1(m\nu_f) &= iN/2 && \text{if } m = kN + 1 \\ F_1(m\nu_f) &= -iN/2 && \text{if } m = kN - 1 \\ F_1(m\nu_f) &= 0 && \text{otherwise} \end{aligned}\right\} \tag{25a}$$

for any integer k when N is even. When N is odd, the following applies:

$$\left.\begin{aligned} F_1(m\nu_f) &= iN/2 && \text{if } m = 2kN + 1 \\ &= -iN/2 && \text{if } m = 2kN - 1 \\ &= iN/2 && \text{if } m = (2k + 1)N - 1 \\ &= -iN/2 && \text{if } m = (2k + 1)N + 1 \\ &= 0 && \text{otherwise} \end{aligned}\right\}. \tag{25b}$$

From Eqs. (19), (13), and (23),

$$F_2(\nu) = 1 - \cos(\pi\nu T_f) - 2 \sum_{n=1}^{M-1} \cos\left(\frac{2\pi n}{N}\right) \cos\left[2\pi\nu\left(\frac{nT_f}{N} - \frac{T_f}{2}\right)\right] \tag{26a}$$

for N even, or

$$F_2(\nu) = -\cos(\pi\nu T_f) - 2 \sum_{n=1}^{M} \cos\left(\frac{2\pi n}{N}\right) \cos\left[2\pi\nu\left(\frac{nT_f}{N} - \frac{T_f}{2}\right)\right] \tag{26b}$$

for N odd. From Eq. (26a) it can be shown that

$$\left.\begin{aligned} F_2(m\nu_f) &= N/2 && \text{if } m = kN \pm 1 \\ F_2(m\nu_f) &= 0 && \text{otherwise} \end{aligned}\right\}, \tag{27a}$$

where k is any integer and N is even. Likewise from Eq. (26b), where N is odd, the following applies:

$$\left.\begin{aligned} F_2(m\nu_f) &= N/2 && \text{if } m = 2kN \pm 1 \\ &= -N/2 && \text{if } m = (2k+1)N \pm 1 \\ &= 0 && \text{otherwise} \end{aligned}\right\}. \quad (27b)$$

Note that Eq. (26a) is a function with period N/T_f (except when $N = 4$ and the period in this special case is $2/T_f$). Similarly, Eq. (26b) is a function with period $2N/T_f$. From Eqs. (25) it follows that $F_1(m\nu_f) = 0$, and, from Eqs. (27), $F_2(m\nu_f) = 0$ only for $m \neq kN \pm 1$; it is therefore possible to obtain an increasing number of harmonics with a zero response [i.e., $F_1(m\nu_f) = F_2(m\nu_f) = 0$] as the number of steps N in the algorithm is increased. As is seen in later sections (e.g., Subsection 5.D), this property is important in determining the algorithm's susceptibility to errors such as measurement nonlinearities.

If an integrating-bucket technique is used instead of a phase-stepping technique to implement the SPSA, then the $f_1(t)$ and $f_2(t)$ functions from Eqs. (12) and (14), respectively, are convolved with a rectangular function. The corresponding $F_1(\nu)$ and $F_2(\nu)$ functions are therefore multiplied by a sinc function that tends to decrease the response at higher harmonics with respect to the phase-stepping case.

5. EXAMPLES OF SYMMETRICAL PHASE-SHIFTING ALGORITHMS

A general expression for the generation of any SPSA according to the analysis in the previous sections is given in Section 8, Eq. (66). Although the algorithms in the following examples could be generated directly from Eq. (66), instead we derive these with reference to the steps outlined in the previous analysis. This approach is intended to give an insight into the nature and characteristics of the SPSA's.

The first nontrivial symmetrical algorithm occurs for $N = 3$. The resulting four-sample algorithm is not considered here since its properties will be investigated in detail in a separate publication at a later date. The next symmetrical algorithm has $N = 4$ and uses five samples. In fact this is the well-known five-sample algorithm that has been analyzed in the literature[11,12] but has not been previously associated with the class of symmetrical algorithms. If we consider this algorithm a member in the class of SPSA's, it will later become evident why its performance is so good in the presence of measurement errors.

A. 4 + 1-Sample Algorithm

The 4 + 1-sample algorithm is the second member in the class of SPSA's. For the phase-stepping case of this algorithm, from Eq. (13) the phase steps are given by $t_n = (nT_f)/4$, i.e., each step is $(T_f/4) = (\pi/2)$ rad. From Eq. (22) the coefficients of the sampling function $f_1(t)$ are therefore

$$\alpha_0 = 0, \quad \alpha_1 = 1, \quad \alpha_2 = 0, \quad \alpha_3 = -1, \quad \alpha_4 = 0. \quad (28)$$

Likewise from Eq. (23) the coefficients of the sampling function $f_2(t)$ are

$$\beta_0 = -\tfrac{1}{2}, \quad \beta_1 = 0, \quad \beta_2 = 1, \quad \beta_3 = 0, \quad \beta_4 = -\tfrac{1}{2}. \quad (29)$$

Figure 2 shows a graphic representation of these sample coefficients. Samples with zero coefficients are represented by filled circles; other samples are represented by an arrow, indicating an impulse function.[10] From Eqs. (2) and (12) the correlation $p(t)$ for $t = 0$ can therefore be written as

$$p(0) = g\left(\frac{-T_f}{4}\right) - g\left(\frac{T_f}{4}\right). \quad (30)$$

Likewise from Eqs. (3) and (14)

$$q(0) = \frac{-g\left(\frac{-T_f}{2}\right)}{2} + g(0)\frac{-g\left(\frac{T_f}{2}\right)}{2}. \quad (31)$$

Now let

$$I_1 = g\left(\frac{-T_f}{2}\right), \quad I_2 = g\left(\frac{-T_f}{4}\right), \quad I_3 = g(0),$$
$$I_4 = g\left(\frac{T_f}{4}\right), \quad I_5 = g\left(\frac{T_f}{2}\right), \quad (32)$$

where I_1 to I_5 designate the intensities of the sampled pattern $g(t)$. Then, from Eqs. (6) and (30)–(32),

$$r = \frac{p(0)}{q(0)} = \frac{2(I_2 - I_4)}{2I_3 - I_1 - I_5}. \quad (33)$$

Now if $T_g = T_f$, i.e., the grating period is equal to the sampling period, then $\nu_g = \nu_f$. From Eqs. (25), $F_1(\nu_g) = 2i$

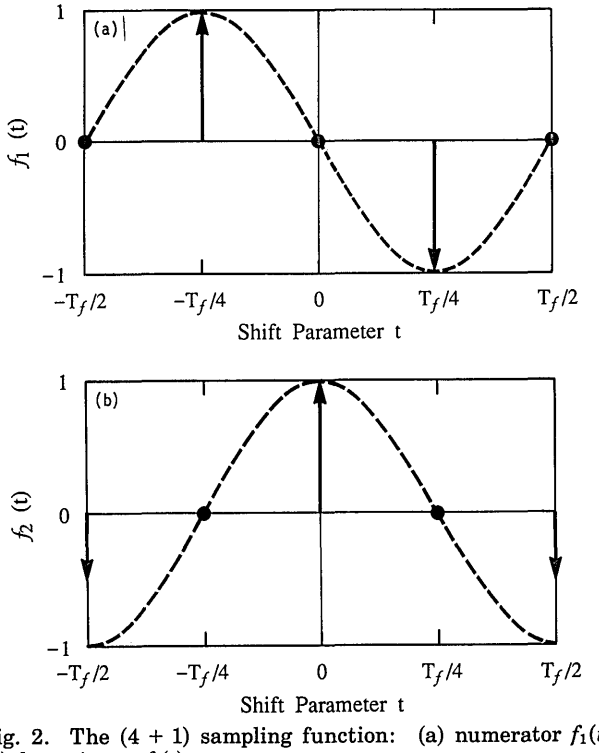

Fig. 2. The (4 + 1) sampling function: (a) numerator $f_1(t)$, (b) denominator $f_2(t)$.

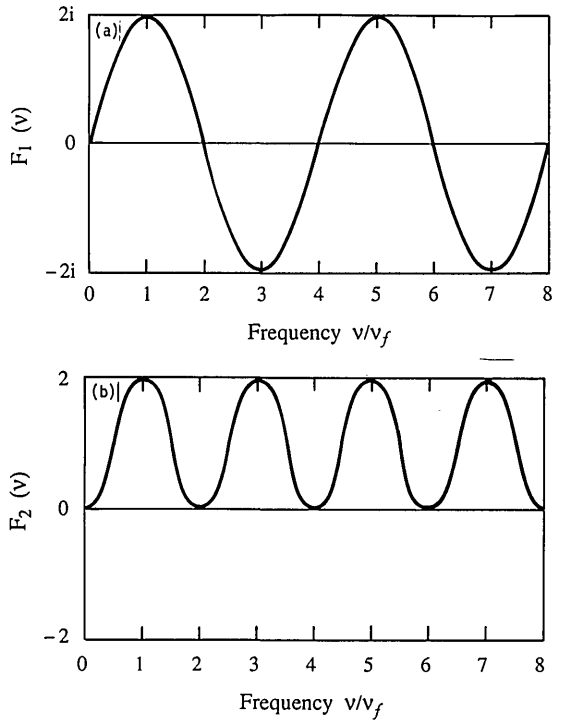

Fig. 3. The (4 + 1) sampling function, FT: (a) numerator $F_1(\nu)$, (b) denominator $F_2(\nu)$.

for $m = 1$, i.e., the fundamental frequency; therefore, from Eqs. (4) and (25),

$$p(0) = 2a_1 \sin(\Phi_1). \tag{34}$$

Likewise, from Eqs. (5) and (27),

$$q(0) = 2a_1 \cos(\Phi_1). \tag{35}$$

Equations (34) and (35) are valid only if the condition in Eqs. (7) applies, that is to say, only if there are no odd-order harmonics in $g(t)$. From Eqs. (34) and (35) one finds

$$r = p(0)/q(0) = \tan(\Phi_1), \tag{36}$$

and from Eqs. (33) and (36),

$$\tan(\Phi_1) = \frac{2(I_2 - I_4)}{2I_3 - I_1 - I_5}. \tag{37}$$

This is the well-known expression for the five-sample phase-shifting algorithm.[11,12]

Figures 2(a) and 2(b) show the sampling functions $f_1(t)$ and $f_2(t)$ for the 4 + 1-sample symmetric algorithm. Figures 3(a) and 3(b) show the Fourier transforms $F_1(\nu)$ and $F_2(\nu)$ for the 4 + 1-sample symmetrical algorithm. Note that both $F_1(\nu)$ and $F_2(\nu)$ have zeros at all even harmonics and $F_1(\nu)$ has stationary points at all odd harmonics, while $F_2(\nu)$ has stationary points at all harmonics including the first (or fundamental). $F_1(\nu)$ and $F_2(\nu)$ for the five-sample case were derived previously.[9]

B. 6 + 1-Sample Algorithm

The third member of the class of symmetrical phase-shifting algorithms is the 5 + 1-sample algorithm. By using the techniques outlined so far, it is a simple exercise to derive the algorithm and hence the frequency response characteristics. However, it is the fourth member of the class—the 6 + 1-sample algorithm—that is analyzed in more detail here. The 6 + 1-sample algorithm has properties at some harmonics that make it particularly useful. To our knowledge the 3 + 1, 5 + 1, and 6 + 1 algorithms have not been reported previously in the literature. For the 6 + 1 algorithm the step size is $T_f/6 = \pi/3$ rad.

From Eq. (22), the coefficients of the sampling function $f_1(t)$ are given by

$$\alpha_0 = 0, \quad \alpha_1 = \frac{\sqrt{3}}{2}, \quad \alpha_2 = \frac{\sqrt{3}}{2}, \quad \alpha_3 = 0,$$
$$\alpha_4 = -\frac{\sqrt{3}}{2}, \quad \alpha_5 = -\frac{\sqrt{3}}{2}, \quad \alpha_6 = 0. \tag{38}$$

Likewise, from Eq. (23), the coefficients of the sampling function $f_2(t)$ are

$$\beta_0 = -\tfrac{1}{2}, \quad \beta_1 = -\tfrac{1}{2}, \quad \beta_2 = \tfrac{1}{2}, \quad \beta_3 = 1,$$
$$\beta_4 = \tfrac{1}{2}, \quad \beta_5 = -\tfrac{1}{2}, \quad \beta_6 = -\tfrac{1}{2}. \tag{39}$$

Figure 4 shows a graphic representation of α_n and β_n sampling coefficients.

In a way similar to the aforementioned 4 + 1-sample algorithm, it can be shown that

$$\tan(\Phi_1) = \sqrt{3}\,\frac{(I_2 + I_3 - I_5 - I_6)}{(-I_1 - I_2 + I_3 + 2I_4 + I_5 - I_6 - I_7)}, \tag{40}$$

where

$$I_1 = g\left(\frac{-Tg}{2}\right), \quad I_2 = g\left(\frac{-Tg}{3}\right),$$
$$I_3 = g\left(\frac{-Tg}{6}\right), \quad I_4 = g(0), \quad I_5 = g\left(\frac{Tg}{6}\right),$$
$$I_6 = g\left(\frac{Tg}{3}\right), \quad I_7 = g\left(\frac{Tg}{2}\right). \tag{41}$$

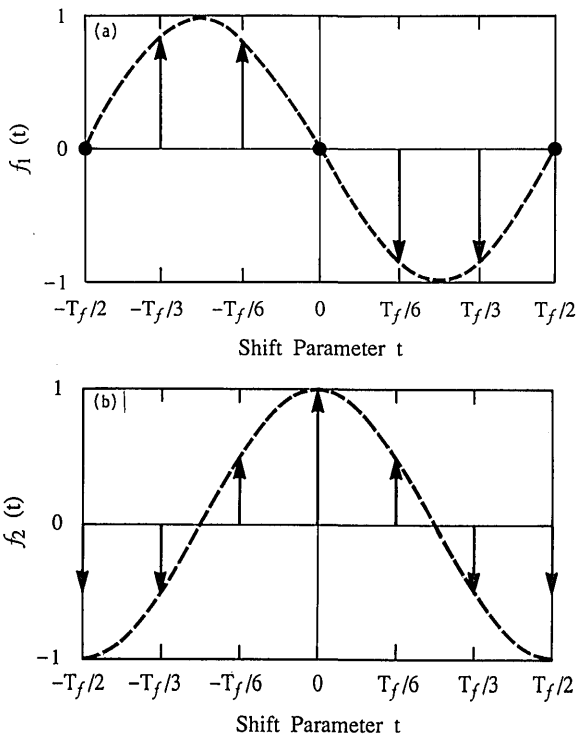

Fig. 4. The (6 + 1) sampling function: (a) numerator $f_1(t)$, (b) denominator $f_2(t)$.

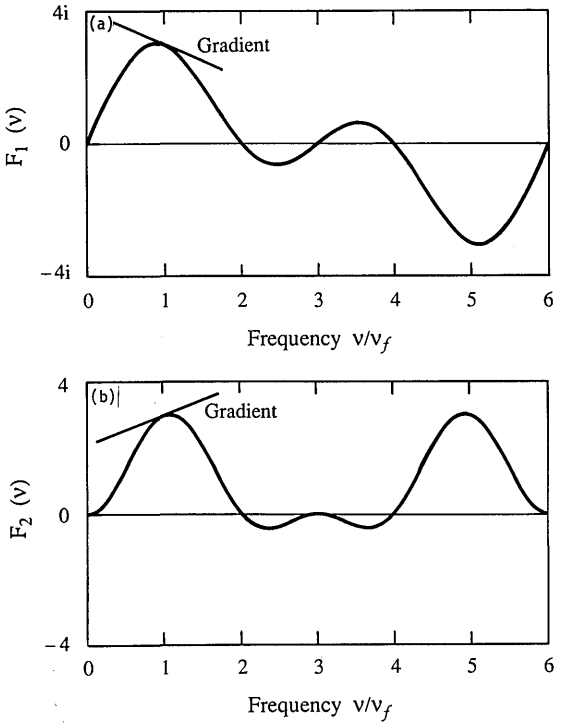

Fig. 5. The (6 + 1) sampling function, FT: (a) numerator $F_1(\nu)$, (b) denominator $F_2(\nu)$.

C. Frequency Spectra of $F_1(\nu)$ and $F_2(\nu)$ for the 6 + 1-Sample Algorithm

From Eq. (24) with $N = 6$,

$$F_1(\nu) = \sqrt{3}\,i\left[\sin\left(\frac{\pi\nu}{3\nu_f}\right) + \sin\left(\frac{2\pi\nu}{3\nu_f}\right)\right]. \tag{42}$$

Likewise, from Eq. (26a),

$$F_2(\nu) = 1 - \cos\left(\frac{\pi\nu}{\nu_f}\right) + \cos\left(\frac{\pi\nu}{3\nu_f}\right) - \cos\left(\frac{2\pi\nu}{3\nu_f}\right). \tag{43}$$

These equations are plotted as a function of ν/ν_f in Fig. 5. Although these functions extend indefinitely for both positive and negative frequencies, they have been plotted only for $0 \le \nu/\nu_f \le 6$. As we mentioned after Eq. (24), the period of these functions is $6\nu_f$ in this case.

D. Error Susceptibility of the 6 + 1-Sample Algorithm

From Fig. 5 it can be seen that because of the zero response in both $F_1(m\nu_f)$ and $F_2(m\nu_f)$ for $m = 0, 2, 3, 4, 6, \ldots$, this algorithm will have low susceptibility to errors in calculated phase because of measurement nonlinearities (e.g., detector and illumination) that result in $g(t)$ having components at these harmonics. This algorithm is, however, not immune to errors produced by the presence of the fifth harmonic in $g(t)$. If the 6 + 1-sample algorithm is implemented with an integrating-bucket method, the fifth harmonic will be less of a problem. If the fifth harmonic causes unacceptable errors, then an algorithm with more than seven steps will cure this problem.

6. IMPORTANCE OF GRADIENTS ALONG $F_1(\nu)$ and $F_2(\nu)$

In a practical phase measurement system, Eq. (8) may not be satisfied, that is, the sampling functions may not be equal in magnitude and in quadrature. This could, for example, be due to phase-shift calibration errors where $\nu_g \neq \nu_f$. Since Φ_1 is evaluated from the ratio of $p(0)$ and $q(0)$, i.e., Eq. (6), in the presence of phase-shift errors, the change in $p(0)$ should ideally be matched by that in $q(0)$ so that the ratio remains unchanged. For a sinusoidal intensity $g(t)$, this ratio defined in Eqs. (4)–(6) may be written as

$$r = \tan(\Phi_1') = \frac{\mathrm{Re}[F_1(\nu_g)]}{\mathrm{Re}[F_2(\nu_g)]}\tan(\Phi_1), \tag{44}$$

where Φ_1' is the calculated phase and can be expressed as $\Phi_1' = \Phi_1 + \Delta\Phi_1$, where $\Delta\Phi_1$ is the error in the calculated phase. $F_1(\nu_g)$ and $F_2(\nu_g)$ can be approximated by the first-order Taylor series expansions at ν_f:

$$F_1(\nu_g) \simeq F_1(\nu_f) + \Delta\nu\frac{dF_1(\nu_f)}{d\nu}. \tag{45}$$

Similarly, for $F_2(\nu_g)$,

$$F_2(\nu_g) \simeq F_2(\nu_f) + \Delta\nu\frac{dF_2(\nu_f)}{d\nu}, \tag{46}$$

where $\Delta\nu = \nu_g - \nu_f$ is small. $\Delta\Phi_1$ can be determined from the relative gradients of $F_1(\nu)$ and $F_2(\nu)$ in the vicinity of $\nu = \nu_f$. Ideally, in order to minimize errors in Φ_1, one should match these gradients or slopes, i.e.,

$$\frac{dF_1(\nu_f)}{d\nu} = i\frac{dF_2(\nu_f)}{d\nu}. \tag{47}$$

To have stationary points at $F_1(\nu_f)$ and $F_2(\nu_f)$ is the ideal situation. To show how these gradients affect the error in the calculated phase, by way of an example we now consider the performance of the 6 + 1 algorithm presented in Subsection 5.B. From Eq. (42),

$$\frac{dF_1(\nu_f)}{d\nu} = -\frac{\pi i}{2\sqrt{3}\,\nu_f}. \tag{48}$$

Likewise, from Eq. (43),

$$\frac{dF_2(\nu_f)}{d\nu} = \frac{\pi}{2\sqrt{3}\,\nu_f}. \tag{49}$$

These gradients are shown in Fig. 5 by tangents to $F_1(\nu)$ and $F_2(\nu)$ at $\nu = \nu_f$.

From Eqs. (48) and (49),

$$\frac{dF_1(\nu_f)}{d\nu} - i\frac{dF_2(\nu_f)}{d\nu} = -\frac{i\pi}{\sqrt{3}\,\nu_f}, \tag{50}$$

which shows clearly that condition (47) is not satisfied.

If e is the step-to-step phase-shift error, then the error $\Delta\Phi_1$ in the calculated phase Φ_1 of the fundamental can be shown from expansion of Eq. (40) or directly from Eq. (44) to be

$$\Delta\Phi_1 \simeq (e/2\sqrt{3})\sin(2\Phi_1) \tag{51}$$

for $e \ll 1$ rad. The error is at twice the frequency of the fundamental.

The preceding analysis considered a system with sinusoidal intensity distribution $g(t)$. In practical systems,

where $g(t)$ has higher-order harmonics and phase-shift errors occur, the error in the calculated phase Φ_1' will also depend on the gradients at those harmonics. This dependence can be formulated by expressing $p(0)$ and $q(0)$ from Eqs. (4) and (5), respectively, as first-order Taylor series expansions.

7. MODIFIED SYMMETRICAL PHASE-SHIFTING ALGORITHMS THAT ELIMINATE PHASE-SHIFT ERRORS

The error in Φ_1' that is due to the phase-shift error can be eliminated by modifying $F_1(\nu)$ so that Eq. (47) holds. This can be achieved by adding to $F_1(\nu)$ another function $C(\nu)$ that has the following properties:

(i) Zero value at all harmonics ($\nu = m\nu_f$);
(ii) Nonzero imaginary slope at $\nu = \nu_f$; and
(iii) Its inverse Fourier transform should be a discrete function.

Condition (i) ensures that the modified algorithm has exactly the same effect as the unmodified algorithm when the phase-shift error is zero. Condition (ii) allows a change to be made to the gradient of the spectrum of the sample function $F_1(\nu)$. Condition (iii) ensures that the modified algorithm can be implemented by using discrete samples.

It is worth noting that the modified algorithm may no longer satisfy the weighted least-squares fit conditions that determine Eqs. (22) and (23). However, it can be shown that condition (i) maintains the algorithm's performance at all harmonics when the stepping is correct. A simple sine wave turns out to be the function $C(\nu)$ that fulfills the aforementioned conditions:

$$C(\nu) = 2ic_0 \sin\left(\frac{\pi\nu}{\nu_f}\right), \tag{52}$$

where c_0 is a real number. The discrete sampling function corresponding to $C(\nu)$ is its inverse Fourier transform:

$$c(t) = c_0\left\{\delta\left(t - \frac{T_f}{2}\right) - \delta\left(t + \frac{T_f}{2}\right)\right\}. \tag{53}$$

Figure 6 shows both $C(\nu)$ and $c(t)$. Now the slope of $C(\nu)$ is

$$\frac{dC(\nu)}{d\nu} = \frac{2\pi ic_0}{\nu_f}\cos\left(\frac{\pi\nu}{\nu_f}\right); \tag{54}$$

therefore,

$$\frac{dC(\nu_f)}{d\nu} = -\frac{ic_0 2\pi}{\nu_f}. \tag{55}$$

Now if we define a new sampling function

$$f_1'(t) = f_1(t) + c(t) \tag{56}$$

and its FT

$$F_1'(\nu) = F_1(\nu) + C(\nu), \tag{57}$$

we have

$$\frac{dF_1'(\nu_f)}{d\nu} = \frac{dF_1(\nu)}{d\nu} + \frac{dC(\nu)}{d\nu}. \tag{58}$$

From Eq. (24), which defines the spectrum $F_1(\nu)$,

$$\frac{dF_1(\nu_f)}{d\nu} = \frac{2i\pi T_f}{N}\sum_{n=1}^{M}(2n - N)\sin\left(\frac{2\pi n}{N}\right)\cos\left(\frac{2\pi n}{N}\right). \tag{59}$$

Likewise from Eqs. (26), which define the spectrum $F_2(\nu)$, for N even

$$\frac{dF_2(\nu_f)}{d\nu} = -\frac{2\pi T_f}{N}\sum_{n=1}^{M-1}(2n - N)\cos\left(\frac{2\pi n}{N}\right)\sin\left(\frac{2\pi n}{N}\right), \tag{60a}$$

and for N odd

$$\frac{dF_2(\nu_f)}{d\nu} = -\frac{2\pi T_f}{N}\sum_{n=1}^{M}(2n - N)\cos\left(\frac{2\pi n}{N}\right)\sin\left(\frac{2\pi n}{N}\right). \tag{60b}$$

Equations (59) and (60) hold for any SPSA.

To correct for phase-shift errors in systems with sinusoidal intensity distributions, we try to make

$$\frac{dF_1'(\nu_f)}{d\nu} = i\frac{dF_2(\nu_f)}{d\nu}. \tag{61}$$

This equation is analogous to Eq. (47) in Section 6.

Combining Eqs. (61) and (58) gives

$$\frac{dC(\nu_f)}{d\nu} = i\frac{dF_2(\nu_f)}{d\nu} - \frac{dF_1(\nu_f)}{d\nu}. \tag{62}$$

Combining Eqs. (62), (60a), (59), and (55) results in Eq. (63a), which is valid for N even:

$$c_0 = \frac{1}{N}\sum_{n=1}^{M-1}(N - 2n)\sin\left(\frac{4\pi n}{N}\right). \tag{63a}$$

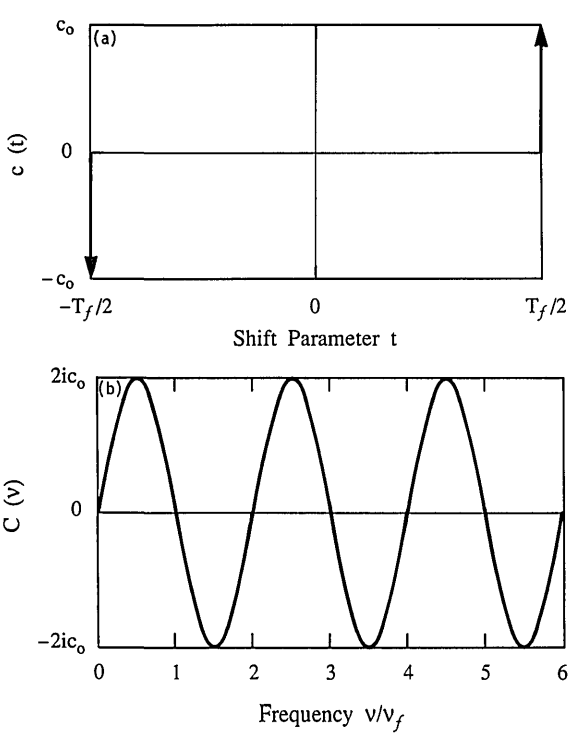

Fig. 6. (a) Modifying sampling function $c(t)$. (b) The FT $C(\nu)$ of the modifying sampling function.

Similarly, when N is odd,

$$c_0 = \frac{1}{N}\sum_{n=1}^{M}(N - 2n)\sin\left(\frac{4\pi n}{N}\right). \tag{63b}$$

Equations (63) define how the symmetrical phase algorithms discussed in Section 3 can be modified to produce new algorithms with well-defined Fourier components (specifically, the gradients of the FT at the fundamental frequency) that result in error-correcting characteristics.

It is interesting to note that for the 4 + 1-sample algorithm ($N = 4$) $c_0 = 0$, so it is not necessary to modify the $F_1(\nu)$ function in this case since Eq. (61) is inherently satisfied. In fact, for the 4 + 1-sample algorithm, the gradients of $F_1(\nu)$ and $F_2(\nu)$ at $\nu = \nu_f$ are equal to zero since there are turning points at this fundamental frequency. This explains the good performance of this algorithm in the presence of phase-shift errors.

For the 6 + 1-sample algorithm $N = 6$ and from Eq. (63a),

$$c_0 = \frac{1}{2\sqrt{3}}. \tag{64}$$

Equation (40) therefore changes to the following:

$$\tan(\Phi_1) = \frac{\sqrt{3}(I_2 + I_3 - I_5 - I_6) + (I_7 - I_1)/\sqrt{3}}{(-I_1 - I_2 + I_3 + 2I_4 + I_5 - I_6 - I_7)}. \tag{65}$$

This is the modified seven-sample algorithm that compensates for phase-shift errors. With reference to Fig. 4(a), the above algorithm introduces two additional samples for α_0 (i.e., $-I_1$) and α_6 (i.e., I_7) at either end of the sample period T_f. If there are no phase-shift errors these two samples make no contribution to the sampled signal. However, with phase-shift errors, the sampled waveform moves with respect to the sample positions, and these additional samples ensure that the change in the numerator of Eq. (65) compensates for the change in the denominator. This will cancel the effect of the phase-shift error.

Figure 7(a) shows the modified sampling function $f_1'(t)$ for the 6 + 1-sample algorithm, and Fig. 7(b) shows its Fourier transform $F_1'(\nu)$. A general expression for generating any SPSA with error-correcting properties is given in Section 8, Eq. (68).

As we mentioned above, the modified sampling function $f_1'(t)$ may no longer be derivable from the least-squares fitting outlined in Section 3. However, it may be the case that $f_1'(t)$ is derivable from a more general least-squares fitting technique that encompasses the effects of phase-shift errors. This conjecture has not been investigated so far. The value of c_0 determined by Eqs. (63) produces functions F_1' and F_2 with matched gradients at $\nu = \nu_f$. Interestingly, criteria other than gradient matching may be used to determine a value for c_0. For example, the phase error $\Delta\Phi$ can be decomposed into a series of factors related to the harmonics present in $g(t)$. Hence c_0 may be defined to constrain these harmonic components in some way (for example, a minimum rms error).

8. SUMMARY

The Fourier analysis used here to derive a new class of phase-shifting algorithms also adds insight into the mechanism of the phase-shifting process. The symmetrical phase-shifting algorithms offer considerable flexibility in tailoring the most suitable algorithm for specific measurements. For example, if the frequency spectrum of the waveform to be measured is known, then the most suitable phase-shifting algorithm can be easily derived.

The outlined procedure for modifying the SPSA's to minimize the effect of phase-shift error adds to the phase measurement accuracy achievable in practical systems while relaxing the tolerance requirements on the phase-shifting mechanism. The modified algorithms require the same number of steps and computer storage facilities as their unmodified counterparts. SPSA's can be easily implemented in modern microcomputers for rapid and accurate phase measurements over a large grid of points.

The N + 1-sample SPSA's developed in this paper can be concisely represented by the following generic equations. First, the unmodified SPSA is

$$\tan(\Phi_1) = \frac{\sum_{n=1}^{N-1}\sin\left(\frac{2\pi n}{N}\right)I_{n+1}}{-\frac{(I_1 + I_{N+1})}{2} - \sum_{n=1}^{N-1}\cos\left(\frac{2\pi n}{N}\right)I_{n+1}}. \tag{66}$$

The intensity values are defined by

$$I_{n+1} = g(t_n) \quad \text{for } n = 0\text{–}N, \tag{67}$$

where t_n is given by Eq. (13) and the phase-step size is T_f/N or, equivalently, (360°/N).

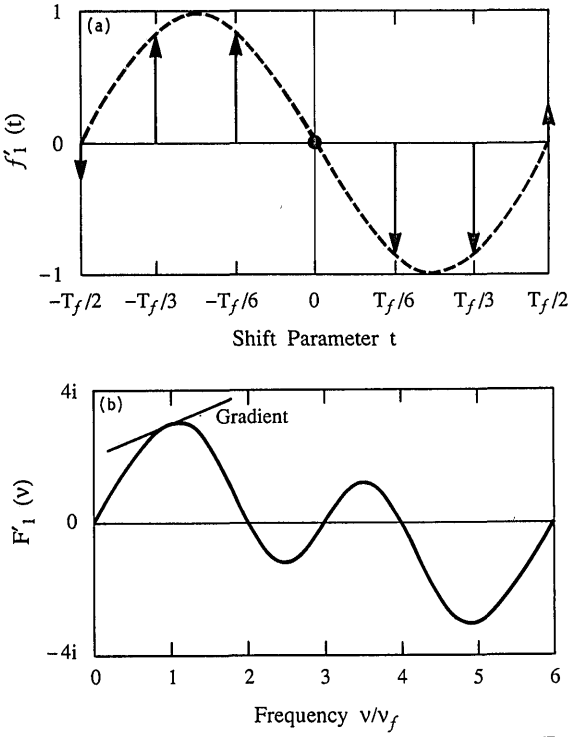

Fig. 7. (a) Modified (6 + 1) sampling function $f_1'(t)$ with error-correcting properties. (b) The modified (6 + 1) sampling function FT $F_1'(\nu)$.

Second, the equation for the modified (error-correcting) SPSA is

$$\tan(\Phi_1) = \frac{c_0(I_{N+1} - I_1) + \sum_{n=1}^{N-1} \sin\left(\frac{2\pi n}{N}\right) I_{n+1}}{-\frac{(I_1 + I_{N+1})}{2} - \sum_{n=1}^{N-1} \cos\left(\frac{2\pi n}{N}\right) I_{n+1}}, \tag{68}$$

and c_0 is given by Eqs. (63). Equations (66) and (68) can be used to generate SPSA's for any value of $N > 2$.

REFERENCES

1. K. Creath, "Phase-measurement interferometry techniques," in *Progress in Optics*, E. Wolf, ed. (Elsevier, New York, 1988), Vol. 26, Chap. 5, pp. 349–383.
2. K. A. Stetson and W. R. Brohinsky, "Electrooptic holography and its application to hologram interferometry," Appl. Opt. **24,** 3631—3637 (1985).
3. P. Harihan, B. F. Oreb, and N. Brown, "Real-time holographic interferometry: a microcomputer system for the measurement of vector displacements," Appl. Opt. **22,** 876–880 (1983).
4. P. Carre, "Installation et utilisation du comparateur photoelectrique et interferentiel du Bureau International des Poids et Mésures," Metrologia **2,** 13–16 (1966).
5. J. C. Wyant, "Interferometric optical metrology: basic principles and new systems," Laser Focus (May, 1982), pp. 65–71.
6. J. H. Bruning, D. R. Herriott, J. E. Gallagher, D. P. Rosenfeld, A. D. White, and D. J. Brangaccio, "Digital wavefront measuring interferometer for testing optical surfaces and lenses," Appl. Opt. **13,** 2693–2703 (1974).
7. J. E. Greivenkamp, "Generalized data reduction for heterodyne interferometry," Opt. Eng. **23,** 350–352 (1984).
8. C. J. Morgan, "Least-squares estimation in phase-measurement interferometry," Opt. Lett. **7,** 368–370 (1982).
9. K. Freischlad and C. L. Koliopoulos, "Fourier description of digital phase-measuring interferometry," J. Opt. Soc. Am. A **7,** 542–551 (1990).
10. R. Bracewell, *The Fourier Transform and Its Applications,* USA, Electrical and Electronic Engineering Series (McGraw-Hill, New York, 1965).
11. P. Hariharan, B. F. Oreb, and T. Eiju, "Digital phase-shifting interferometry: a simple error-compensating phase calculation algorithm," Appl. Opt. **26,** 2504–2506 (1987).
12. J. Schwider, R. Burow, K. E. Elssner, J. Grzanna, R. Spolaszyk, and K. Merkel, "Digital wave-front measuring interferometry: some systematic error sources," Appl. Opt. **22,** 3421–3432 (1983).