Phase-shifting algorithms for nonlinear and spatially nonuniform phase shifts

Kenichi Hibino

Mechanical Engineering Laboratory, Namiki, Tsukuba 305, Japan

Bob F. Oreb and David I. Farrant

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

Kieran G. Larkin

School of Physics, University of Sydney, NSW 2006, Australia

Received September 24, 1996; accepted November 4, 1996

In phase-shifting interferometry spatial nonuniformity of the phase shift gives a significant error in the evaluated phase when the phase shift is nonlinear. However, current error-compensating algorithms can counteract the spatial nonuniformity only in linear miscalibrations of the phase shift. We describe an error-expansion method to construct phase-shifting algorithms that can compensate for nonlinear and spatially nonuniform phase shifts. The condition for eliminating the effect of nonlinear and spatially nonuniform phase shifts is given as a set of linear equations of the sampling amplitudes. As examples, three new algorithms (six-sample, eight-sample, and nine-sample algorithms) are given to show the method of compensation for a quadratic and spatially nonuniform phase shift.
[S0740-3232(97)01604-9]

Key words: phase measurement, interferometry, phase-shift algorithm, nonlinear phase shift, spatially nonuniform phase shift.

1. INTRODUCTION

In phase-shifting interferometry[1] the phase difference between an object beam and a reference beam is varied by phase shifting, and the signal irradiance is sampled at equal intervals of the phase difference. The phase distribution of a fringe pattern can be calculated by phase-shifting algorithms.

Systematic phase-shift errors, as well as the nonsinusoidal waveform of the signal, are the most common sources of systematic errors in the evaluated phase.[2] Systematic phase-shift errors are caused by linear miscalibration, nonlinear sensitivity, or spatial nonuniformity of the phase shifter. For some phase shifters, such as a liquid-crystal modulator, the phase shift is not only nonlinear but also spatially nonuniform. Distortion in a lead zirconium titanate crystal (which is widely used in phase-shift devices) also gives a nonlinear and spatially nonuniform phase modulation. The spatial nonuniformity of the phase shift is also an important problem in applying phase-shifting techniques to a scanning white-light interferometer,[3] since the variation of fringe visibility along the optical path is mathematically equivalent to the variation of phase-shift errors (depending on the position).

Many studies have been reported on error-compensating algorithms[4–13] that can eliminate the effect of systematic phase-shift errors. Table 1 lists the phase-shifting algorithms and their immunity to linear miscalibration, nonlinear sensitivity, and spatial nonuniformity of the phase shift. Parameter p represents the maximum order of nonlinearity of the phase shift, and parameter j represents the maximum order of the harmonic signals that the algorithm can compensate for.

In Table 1 the algorithms can be classified into three groups by their ability to compensate for phase-shift errors. Group I, discrete Fourier algorithms (or the synchronous detection algorithms[1]), can compensate for the effect of harmonic signals only and have no immunity to phase-shift errors. Group II, error-compensating algorithms, compensate for linear miscalibration ($p = 1$) or nonlinear sensitivity ($p > 1$) of the phase shift as well as other errors. Group III, error-compensating algorithms, compensate for not only phase-shift errors but also errors caused by a coupling between the phase-shift errors and the harmonic components of the signal.

For linear phase-shift miscalibrations ($p = 1$), the spatial nonuniformity of the phase shift is not a significant problem, since error-compensating algorithms can easily acquire an immunity to the spatial nonuniformity. Subsection 2.C explains that an error-compensating algorithm for linear phase-shift miscalibrations becomes insensitive to the spatial nonuniformity if the algorithm has symmetric and antisymmetric sampling amplitudes in the denominator and the numerator of the arctangential variable. Many algorithms for linear phase-shift miscalibrations have these symmetries in their sampling amplitudes and thus are immune to the spatial nonuniformity of the phase shift (see Table 1 for $p = 1$).

A higher-order immunity to linear miscalibrations of the phase shift was also investigated by several authors,[10,11,13] and several algorithms were obtained[10,11] that give the residual phase errors of the orders $o(\epsilon_1{}^3)$ and $o(\epsilon_1{}^4)$, where ϵ_1 is the error coefficient of the phase shift. The stability of these algorithms against low-frequency mechanical vibrations of the interferometer was pointed out and verified experimentally.[14]

For nonlinear phase shift, however, there have not been any algorithms reported that can compensate for the spatial nonuniformity of the phase shift. The error-compensating algorithms[10,11] with higher-order immunity to linear miscalibration can compensate for the nonlinear phase shift (see Table 1). However, they are not insensitive to the effect of spatial nonuniformity.

For example, de Groot calculated the residual phase error $\Delta\phi$ of his seven-frame algorithm,[11]

$$\phi = \arctan\left[\frac{7(I_2 - I_4) - (I_0 - I_6)}{-4(I_1 + I_5) + 8I_3}\right], \qquad (1)$$

Table 1. Phase-Shifting Algorithms and Their Immunity to Phase-Shift Errors

Group	Immunity to Phase-Shift Errors	Immunity to the Spatial Uniformity of the Phase Shift[a]: Spatially Uniform	Spatially Nonuniform
I	None	Bruning[1]	
II	Linear miscalibration ($p = 1$)		Carre (4)[4] Schwider (5)[5] Hariharan (5)[6] Schwider[b] (4)[7] Larkin ($j + 3$)[8] Surrel ($j + 3$)[9]
	Quadratic nonlinearity ($p = 2$)	Schmit (5, 6)[10] de Groot (7)[11] Present work ($j + 4$)	Present work (6, 8)
III	Linear miscalibration ($p = 1$)		Hibino (7, 11)[12] Surrel[b] (6, 10),[13] Surrel ($2j + 3$)[13]
	Quadratic nonlinearity ($p = 2$)		Present work (9)

[a] Each algorithm has its reference number as a superscript. The number listed in the parentheses after each (first) author's name represents the number of samples of the algorithm.

[b] The symmetry in the sampling amplitudes is not apparent, since the amplitudes take different notations for offset bias in the phase-shift values.

for quadratic phase shifts

$$\alpha_r = \frac{\pi}{2}(r - 3)\left[1 + \epsilon_1 + \frac{\epsilon_2}{2}(r - 3)\right] \quad \text{for } r = 0, \ldots, 6, \qquad (2)$$

which results in

$$\Delta\phi = \frac{\pi\epsilon_2}{2} - \frac{\pi^2\epsilon_2{}^2}{64}\sin(2\phi) + \frac{\pi^4\epsilon_1{}^4}{256}\sin(2\phi) + \ldots . \qquad (3)$$

The residual error contains a dc component $\pi\epsilon_2/2$ that is linear with error coefficient ϵ_2. For a spatially nonuniform phase shift, coefficient ϵ_2 is no longer uniform across the observed aperture and causes a significant phase error. This type of dc component in the residual error is common in the algorithms[10,11] that have higher-order immunity to the linear miscalibration.

Systematic approaches for deriving error-compensating algorithms have been proposed by several authors based on the averaging method of successive samples,[5,7,10,15] a Fourier description of the sampling functions,[8,11,13,16] and an analytical expansion of the phase error.[12] However, none of these approaches has been developed so that it can provide immunity to the spatial nonuniformity of the nonlinear phase shift.

In this paper we develop the analytical expansion method[12] so that it can obtain an immunity to a nonlinear and spatially nonuniform phase shift, i.e., a most general type of systematic phase-shift error. The residual phase errors are expanded into a Taylor series of both nonlinear coefficients of the phase shift and harmonic amplitudes of the signal. The conditions for eliminating both effects will be given as a set of linear equations for the sampling amplitudes of an algorithm.

We demonstrate that in order to acquire an immunity to a spatially nonuniform phase shift, the sampling amplitudes of an error-compensating algorithm need to satisfy additional $p/2$ linear equations when order p is even. Correction of the number of samples in our previous work[12] is also necessary, which will be discussed in Subsection 2.F.

As examples, new six-sample (group II with $p = 2$ and $j = 1$), new eight-sample (group II with $p = 2$ and $j = 2$), and new nine-sample (group III with $p = 2$ and $j = 2$) algorithms are derived that compensate for a quadratic and spatially nonuniform phase shift.

In our previous work[12] on the analytical expansion method, the derivation is limited to group III algorithms only, which compensate for a linear miscalibration of the phase shift and harmonic components of the signal. With the present extended formulation, all types of algorithm classified in Table 1, except for Carre's strict solution,[4] can be derived. For example, a generic ($j + 4$)-sample algorithm of group II is derived by which the effects of a quadratic phase-shift error and harmonic components of the signal up to the jth order are eliminated.

Another advantage of the present analytical approach is that since all the conditions for suppressing systematic errors are represented by a set of linear equations for the sampling amplitudes, the algorithm designs for balancing random and systematic phase errors can easily be formulated.[17] The variance of random phase errors caused by random uncorrelated irradiance noise is proportional to the total square of the sampling amplitudes of an algorithm.[17,18] Therefore one strategy for algorithm design can be reduced to a problem of minimizing the total square of the sampling amplitudes for a given number of samples and under the constraint of a set of linear equations to suppress systematic errors.

2. DERIVATION OF PHASE-SHIFTING ALGORITHMS

A. Phase Error Expansions

Consider the irradiance $I(x, y, \alpha)$ at a designated point (x, y) with a nonsinusoidal periodic waveform as a function of a phase-shift parameter α:

$$I(x, y, \alpha) = s_0(x, y) + \sum_{k=1}^{j} s_k(x, y) \times \cos[k\alpha - \phi_k(x, y)], \tag{4}$$

where s_k and ϕ_k are the amplitude and the phase of the kth harmonic component of the signal, j is the maximum order of the harmonic components, and ϕ_1 is the object phase (at the fundamental frequency) to be measured. To simplify the mathematical notation, we will confine our analysis to a single point, although it is equally applicable to all other points of interest. Therefore the (x, y) dependence of I, s_k, and ϕ_k will not be highlighted.

Consider an m-sample phase-shifting algorithm, where the reference phases are separated by $m - 1$ equal intervals of $2\pi/n$ rad, where n is an integer. A general expression for the calculated phase in m-sample algorithms is

$$\phi = \arctan\left(\frac{\sum_{r=1}^{m} b_r I_r}{\sum_{r=1}^{m} a_r I_r}\right), \tag{5}$$

where a_r and b_r are the rth sampling amplitudes and $I_r = I(\alpha_r)$ is the rth sampled irradiance.

When the phase shift is nonlinear, each phase-shift value α_r is a function of the phase-shift parameter. The phase-shift value for the rth sample can be given by a polynomial of the unperturbed phase-shift value α_{0r} as

$$\begin{aligned}\alpha_r &= \alpha_{0r}[1 + \epsilon(\alpha_{0r})] \\ &= \alpha_{0r}\left[1 + \epsilon_1 + \epsilon_2 \frac{\alpha_{0r}}{\pi} + \epsilon_3\left(\frac{\alpha_{0r}}{\pi}\right)^2 + \ldots \right. \\ &\quad \left. + \epsilon_p\left(\frac{\alpha_{0r}}{\pi}\right)^{p-1}\right] \qquad \text{for } r = 1, 2, \ldots, m,\end{aligned} \tag{6}$$

where p $(p \leq m - 1)$ is the maximum order of the nonlinearity, ϵ_q $(1 \leq q \leq p)$ are the error coefficients, which can be spatially nonuniform, and $\alpha_{0r} = 2\pi[r - (m + 1)/2]/n$ is the unperturbed phase shift. An offset value $(m + 1)/2$ for the reference phase is introduced for convenience of notation and gives only a spatially uniform constant bias to the calculated phase.

In the following we assume that nonlinear coefficients ϵ_q (for $q = 1, \ldots, p$) and amplitude ratios s_k/s_1 (for $k = 2, \ldots, j$) are smaller than unity.

The error $\Delta\phi$ in the calculated phase is a function of amplitudes s_k/s_1, phases ϕ_k (for $k = 1, 2, \ldots, j$), and error coefficients ϵ_q (for $q = 1, 2, \ldots, p$) and can be expanded into a Taylor series by

$$\begin{aligned}\Delta\phi &= \phi - \phi_1 \\ &= o(s_k) + o(\epsilon_q) + o(\epsilon_q s_k) + o(s^2) + o(\epsilon^2),\end{aligned} \tag{7}$$

where $o(\epsilon_q s_k)$, $o(s^2)$, and $o(\epsilon^2)$ denote the terms involving $\epsilon_q s_k/s_1$, $s_k s_h/{s_1}^2$, and $\epsilon_q \epsilon_t$ for $k, h = 2, \ldots, j$, and $q, t = 1, \ldots, p$. Since the numerator and the denominator of the algorithm of Eq. (5) are linear with amplitudes s_k, the numerator and the denominator do not contain errors $o(s^2)$ after the error $o(s_k)$ is compensated. Therefore, when the error $o(s_k)$ is compensated by an algorithm, the errors $o(s^2)$ in $\Delta\phi$ disappear. In the following we discuss the suppression of errors $o(s_k)$, $o(\epsilon_q)$, and $o(\epsilon_q s_k)$ and neglect the higher-order errors $o(\epsilon^2)$.

Suppressing the harmonic components of the signal has already been discussed by many authors.[1,2,5,6,8–13,15,16,19] When the algorithm is insensitive to harmonic components of the signal up to the jth order $(j \geq 1)$, the sampling amplitudes $\{a_r\}$ and $\{b_r\}$ satisfy the following $4j + 2$ equations[8,12,16]:

$$\sum_{r=1}^{m} a_r \sin(k\alpha_{0r}) = 0 \qquad \text{for } k = 1, \ldots, j, \tag{8}$$

$$\sum_{r=1}^{m} a_r \cos(k\alpha_{0r}) = \delta(k, 1) \qquad \text{for } k = 0, 1, \ldots, j, \tag{9}$$

$$\sum_{r1}^{m} b_r \sin(k\alpha_{0r}) = \delta(k, 1) \qquad \text{for } k = 1, \ldots, j, \tag{10}$$

$$\sum_{r=1}^{m} b_r \cos(k\alpha_{0r}) = 0 \qquad \text{for } k = 0, 1, \ldots, j, \tag{11}$$

where $\delta(k, 1)$ is the Kronecker delta function.

Now we calculate the residual phase errors and expand them into a series of error coefficients of the phase shift and the amplitudes of the harmonic signals. Substituting Eqs. (4) and (6) into Eq. (5) and expanding both the denominator and the numerator into powers of errors ϵ_q, and using the relations of Eqs. (8)–(11), we can calculate the phase as

$$
\begin{aligned}
\phi &= \arctan\left[\frac{\displaystyle\sum_{r=1}^{m} b_r\left\{s_0 + \sum_{k=1}^{j} s_k \cos\left(k\alpha_{0r}\left[1 + \sum_{q=1}^{p}\epsilon_q\left(\frac{\alpha_{0r}}{\pi}\right)^{q-1}\right] - \phi_k\right)\right\}}{\displaystyle\sum_{r=1}^{m} a_r\left\{s_0 + \sum_{k=1}^{j} s_k \cos\left(k\alpha_{0r}\left[1 + \sum_{q=1}^{p}\epsilon_q\left(\frac{\alpha_{0r}}{\pi}\right)^{q-1}\right] - \phi_k\right)\right\}}\right] \\
&= \arctan\left[\frac{\displaystyle\sin\phi_1 - \pi\sum_{r=1}^{m}\sum_{k=1}^{j}\sum_{q=1}^{p}\frac{s_k}{s_1}k\epsilon_q\left(\frac{\alpha_{0r}}{\pi}\right)^{q} b_r\sin(k\alpha_{0r} - \phi_k)}{\displaystyle\cos\phi_1 - \pi\sum_{r=1}^{m}\sum_{k=1}^{j}\sum_{q=1}^{p}\frac{s_k}{s_1}k\epsilon_q\left(\frac{\alpha_{0r}}{\pi}\right)^{q} a_r\sin(k\alpha_{0r} - \phi_k)}\right] \\
&= \arctan\left[\frac{\displaystyle\sin\phi_1 - \sum_{q=1}^{p}\epsilon_q B_q(\phi_1) - \sum_{q=1}^{p}\sum_{k=2}^{j}\epsilon_q\frac{s_k}{s_1}D_{k,q}(\phi_k)}{\displaystyle\cos\phi_1 - \sum_{q=1}^{p}\epsilon_q A_q(\phi_1) - \sum_{q=1}^{p}\sum_{k=2}^{j}\epsilon_q\frac{s_k}{s_1}C_{k,q}(\phi_k)}\right] \\
&\simeq \arctan\left(\left\{1 + \sum_{q=1}^{p}\epsilon_q\left[\frac{A_q(\phi_1)}{\cos\phi_1} - \frac{B_q(\phi_1)}{\sin\phi_1}\right] + \sum_{q=1}^{p}\sum_{k=2}^{j}\epsilon_q\frac{s_k}{s_1}\left[\frac{C_{k,q}(\phi_k)}{\cos\phi_1} - \frac{D_{k,q}(\phi_k)}{\sin\phi_1}\right]\right\}\tan\phi_1\right) \\
&\simeq \phi_1 + \frac{1}{2}\sum_{q=1}^{p}\epsilon_q\left[\frac{A_q(\phi_1)}{\cos\phi_1} - \frac{B_q(\phi_1)}{\sin\phi_1}\right]\sin(2\phi_1) + \frac{1}{2}\sum_{q=1}^{p}\sum_{k=2}^{j}\epsilon_q\frac{s_k}{s_1}\left[\frac{C_{k,q}(\phi_k)}{\cos\phi_1} - \frac{D_{k,q}(\phi_k)}{\sin\phi_1}\right]\sin(2\phi_1), \qquad (12)
\end{aligned}
$$

where

$$A_q(\phi_1) = \pi\sum_{r=1}^{m} a_r\left(\frac{\alpha_{0r}}{\pi}\right)^q \sin(\alpha_{0r} - \phi_1), \qquad (13)$$

$$B_q(\phi_1) = \pi\sum_{r=1}^{m} b_r\left(\frac{\alpha_{0r}}{\pi}\right)^q \sin(\alpha_{0r} - \phi_1), \qquad (14)$$

$$C_{k,q}(\phi_k) = \pi k\sum_{r=1}^{m} a_r\left(\frac{\alpha_{0r}}{\pi}\right)^q \sin(k\alpha_{0r} - \phi_k), \qquad (15)$$

$$D_{k,q}(\phi_k) = \pi k\sum_{r=1}^{m} b_r\left(\frac{\alpha_{0r}}{\pi}\right)^q \sin(k\alpha_{0r} - \phi_k), \qquad (16)$$

and an approximation, $\arctan[(1 + \delta)\tan\phi_r] \simeq \phi_1 + (\delta/2)\sin(2\phi_1)$, for $|\delta| \ll 1$ is used.

In Eq. (12) the errors $o(s_k)$ are eliminated by using the conditions of Eqs. (8)–(11), and the specific forms of the residual errors $o(\epsilon_q)$ and $o(\epsilon_q s_k)$ are obtained as the second and third terms of the right-hand side of the equation.

B. Symmetries of the Sampling Amplitudes of the Algorithm

An important result can be derived from Eqs. (8)–(11) indicating that, in order to construct an algorithm with the fewest number of samples, the sampling amplitudes need to have symmetries: $a_r = a_{m+1-r}$ and $b_r = -b_{m+1-r}$ for $r = 1, 2, \ldots, m$. We verify this conclusion here.

When we replace the sampling amplitudes a_r and b_r by a_{m+1-r} and $-b_{m+1-r}$, respectively, in Eqs. (8)–(11), the new equations are still satisfied, since the rth phase shift satisfies $\alpha_{0r} = -\alpha_{0,m+1-r}$ from its definition and gives the relations as $\cos(k\alpha_{0r}) = \cos(k\alpha_{0,m+1-r})$ and $\sin(k\alpha_{0r}) = -\sin(k\alpha_{0,m+1-r})$. Therefore, if a set of sampling amplitudes $\{a_r\}$ and $\{b_r\}$ are the solutions for Eqs. (8)–(11), another set of sampling amplitudes $\{a_{m+1-r}\}$ and $\{-b_{m+1-r}\}$ are also the solutions for them. This result is also true for Eqs. (19)–(21) and (24)–(27), which will be derived in following subsections.

Since all the conditions necessary for eliminating the systematic errors are given by a set of linear equations for the sampling amplitudes, the solutions for the equations generally become unique when the number of sampling amplitudes is minimum. Then, when the algorithm is constructed with the fewest number of samples, the two solutions of amplitudes $\{a_r\}$ and $\{b_r\}$ and amplitudes $\{a_{m+1-r}\}$ and $\{-b_{m+1-r}\}$ need to be identical, since the solutions are unique.

Therefore the symmetry of the sampling amplitudes, $a_r = a_{m+1-r}$ and $b_r = -b_{m+1-r}$, is a necessary condition to construct algorithms with the fewest number of samples. Note that this conclusion is derived under the assumption that the phase shift has an appropriate offset so that it satisfies $\alpha_{0r} = -\alpha_{0,m+1-r}$.

C. Compensation for Spatially Nonuniform Phase Shift in Group II

Group II algorithms compensate for errors $o(\epsilon_q)$ and $o(s_k)$. In order to eliminate phase error $o(\epsilon_q)$ (for $q = 1, \ldots, p$) caused by the pth-order nonlinearity of the phase shift, the corresponding second term on the right-hand side of Eq. (12) needs to be constant or zero for arbitrary values of error coefficients ϵ_q and phase ϕ_1. For spatially nonuniform phase shifts, this term needs to be zero, since error coefficients ϵ_q are not uniform but are a function of the position. Eliminating this term requires that

$$\left[\frac{A_q(\phi_1)}{\cos\phi_1} - \frac{B_q(\phi_1)}{\sin\phi_1}\right]\sin(2\phi_1) = 0 \qquad \text{for } q = 1, 2, \ldots, p. \tag{17}$$

Substituting expressions (13) and (14) for A_q and B_q into Eq. (17) and expanding the trigonometric function into powers of cos ϕ_1 and sin ϕ_1, we can rewrite Eq. (17) as

$$\sum_{r=1}^{m} \alpha_{0r}{}^q(a_r \sin\alpha_{0r} + b_r \cos\alpha_{0r})\sin(2\phi_1) + \sum_{r=1}^{m} \alpha_{0r}{}^q(a_r \cos\alpha_{0r} - b_r \sin\alpha_{0r})\cos(2\phi_1) + \sum_{r=1}^{m} \alpha_{0r}{}^q(a_r \cos\alpha_{0r} + b_r \sin\alpha_{0r}) = 0. \tag{18}$$

Equation (18) shows that the necessary and sufficient conditions to satisfy condition (17) are the following $3p$ equations:

$$\sum_{r=1}^{m} \alpha_{0r}{}^q(a_r \cos\alpha_{0r} + b_r \sin\alpha_{0r}) = 0 \qquad \text{for } q = 1, 2, \ldots, p, \tag{19}$$

$$\sum_{r=1}^{m} \alpha_{0r}{}^q(a_r \cos\alpha_{0r} - b_r \sin\alpha_{0r}) = 0 \qquad \text{for } q = 1, 2, \ldots, p, \tag{20}$$

$$\sum_{r=1}^{m} \alpha_{0r}{}^q(a_r \sin\alpha_{0r} + b_r \cos\alpha_{0r}) = 0 \qquad \text{for } q = 1, 2, .., p. \tag{21}$$

Equations (8)–(11), together with Eqs. (19)–(21), are the necessary and sufficient conditions for the algorithms to compensate for the effects of a nonlinear and spatially nonuniform phase shift of the pth order and harmonic components of the signal up to the jth order.

In contrast, when the phase shift is spatially uniform and the system is not concerned with a fixed dc component, consisting of coefficients ϵ_q in the measured phase, it is sufficient to require that

$$\left[\frac{A_q(\phi_1)}{\cos\phi_1} - \frac{B_q(\phi_1)}{\sin\phi_1}\right]\sin(2\phi_1) = \text{const.} \qquad \text{for } q = 1, 2, \ldots, p \tag{22}$$

for arbitrary phase ϕ_1.

In a similar discussion to that for Eq. (17), the necessary and sufficient conditions to satisfy condition (22) lead to the same equations as Eqs. (20) and (21). Therefore Eq. (19), consisting of p linear equations, is related to the immunity to nonuniform phase shifts.

We shall now show that the error-compensating algorithms for linear miscalibrations of the phase shift can acquire an immunity to a spatially nonuniform phase shift without increasing the number of samples. As mentioned in Subsection 2.B, under the symmetric phase shift $\alpha_{0r} = -\alpha_{0,m+1\text{-}r}$, constructing an algorithm with a minimum number of samples requires that the sampling amplitudes have the symmetries $a_r = a_{m+1-r}$ and $b_r = -b_{m+1-r}$. If an error-compensating algorithm has these symmetries, Eq. (19) for $p = 1$ (or for odd order q) are satisfied automatically, since we note that $\alpha_{0r}{}^q \cos\alpha_{0r} = -\alpha_{0,m+1-r}{}^q \cos\alpha_{0,m+1-r}$ and $\alpha_{0r}{}^q \sin\alpha_{0r} = \alpha_{0,m+1-r}{}^q \sin\alpha_{0,m+1-r}$ for odd order q. The algorithm also satisfies Eqs. (20) and (21), since we have assumed that the algorithm is insensitive to linear miscalibrations.

Therefore we concluded that group II algorithms for linear phase-shift miscalibrations become insensitive to the spatial nonuniformity of the phase shift if the sampling amplitudes have the symmetries $a_r = a_{m+1-r}$ and $b_r = -b_{m+1-r}$.

For a nonlinear phase shift ($p \geqslant 2$), however, it is clear that these symmetries are not sufficient conditions to satisfy Eq. (19) for even order p. In Subsection 3.A we will derive a new algorithm that satisfies these equations for a quadratic phase shift ($p = 2$).

D. Compensation for Spatially Nonuniform Phase Shift in Group III

Group III algorithms compensate for errors $o(\epsilon_q s_k)$ caused by the coupling between the phase-shift errors and the harmonic signals, as well as other errors $o(s_k)$ and $o(\epsilon_q)$. Here we consider the elimination of the errors $o(\epsilon_q s_k)$ when the phase shift is not spatially uniform.

The specific form of the errors $o(\epsilon_q s_k)$ has already been derived, as in the third term on the right-hand side of Eq. (12). With a discussion similar to that in Subsection 2.C, eliminating the third term requires that

$$\left[\frac{C_{k,q}(\phi_k)}{\cos\phi_1} - \frac{D_{k,q}(\phi_k)}{\sin\phi_1}\right]\sin(2\phi_1) = 0 \qquad \text{for } k = 2, \ldots, j \text{ and } q = 1, .., p \tag{23}$$

for arbitrary values of phases ϕ_k.

Substituting expressions (15) and (16) for $C_{k,q}$ and $D_{k,q}$ into Eq. (23) and expanding the trigonometric function into powers of cos ϕ_k and sin ϕ_k, the necessary and sufficient conditions to satisfy condition (23) are the following $4p(j-1)$ equations:

$$\sum_{r=1}^{m} \alpha_{0r}{}^q a_r \sin(k\alpha_{0r}) = 0, \tag{24}$$

$$\sum_{r=1}^{m} \alpha_{0r}{}^q a_r \cos(k\alpha_{0r}) = 0, \tag{25}$$

$$\sum_{r=1}^{m} \alpha_{0r}{}^q b_r \sin(k\alpha_{0r}) = 0, \tag{26}$$

$$\sum_{r=1}^{m} \alpha_{0r}{}^q b_r \cos(k\alpha_{0r}) = 0, \tag{27}$$

for $k = 2, \ldots, j$ and $q = 1, 2, \ldots, p$.

Equations (8)–(11), (19)–(21), and (24)–(27) are the conditions for group III algorithms to eliminate phase errors $o(s_j)$, $o(\epsilon_p)$, and $o(\epsilon_p s_j)$ caused by the spatially nonuniform phase shift of the pth order and harmonic signals up to the jth order.

In contrast, for a spatially uniform phase shift, the condition of Eq. (23) can be replaced by

$$\left[\frac{C_{k,q}(\phi_k)}{\cos\,\phi_1} - \frac{D_{k,q}(\phi_k)}{\sin\,\phi_1}\right]\sin(2\phi_1) = \text{const.}$$

$$\text{for}\quad k = 2,\, \ldots,\, j \quad \text{and} \quad q = 1,\, \ldots,\, p \tag{28}$$

for arbitrary values of phases ϕ_k. However, the resulting conditions for satisfying Eq. (28) reduce to the same equations as Eqs. (24)–(27). Therefore, for the elimination of the additional error $o(\epsilon_p s_j)$, we do not need any additional conditions, even when the phase shift becomes spatially nonuniform.

As for the suppression of linear and nonuniform phase-shift miscalibrations, the same discussion as that for group II in Subsection 2.C can be applied to group III algorithms. If a group III algorithm for linear phase-shift miscalibrations has the symmetries $a_r = a_{m+1-r}$ and $b_r = -b_{m+1-r}$ in their sampling amplitudes, it always satisfies Eq. (19) and thus becomes insensitive to the nonuniformity of the phase shift.

As a brief summary of Subsections 2.C and 2.D, Table 2 shows the necessary equations for deriving group II and group III algorithms with different immunities to the nonlinearity and the spatial uniformity of the phase shift. The present analytical method, with the linear equations, can derive all the classes of error-compensating algorithms with immunities, as shown in Table 1.

E. Number of Samples Necessary for Constructing Group II Algorithms

Here we discuss the number of samples necessary for constructing the algorithms of group II for a spatially nonuniform phase shift. In our present analytical expression, the number of samples is obtained by estimating the number of independent equations among linear equations (8)–(11) and (19)–(21).

Equations (8)–(11) are common conditions for group I, discrete Fourier algorithms. For discrete Fourier algorithms, the number of samples necessary to eliminate the harmonic signals up to the jth order has already been investigated by several authors.[12,13,19] These results show that for a sinusoidal signal ($j = 1$) the minimum number of samples is 3 for any phase-shift interval and that for $j \geqslant 2$ the minimum number of samples depends on the phase-shift interval and is equal to its minimum of $j + 2$ when the interval equals $2\pi/(j + 2)$ rad. In our present expression, these results are equivalent to those that state that $j + 2$ samples are at least necessary to satisfy Eqs. (8)–(11), which was shown in our previous work.[12] The number of samples for a smaller phase-shift interval than $2\pi(j + 2)$ rad has also been thoroughly investigated, using characteristic polynomials, by Surrel.[13]

It was shown in Subsection 2.B that in order to minimize the number of samples, the sampling amplitudes need to satisfy the symmetric relations $a_r = a_{m+1-r}$ and $b_r = -b_{m+1-r}$ for $r = 1\ 2,\, \ldots,\, m$, where m is the number of samples. Note that this conclusion is valid when the phase shift is defined to satisfy $\alpha_{0r} = -\alpha_{0,m+1-r}$. In the remainder of this subsection we assume these symmetries for the sampling amplitudes:

$$\{a_r\} = (a_1,\, a_2,\, \ldots,\, a_2,\, a_1),$$
$$\{b_r\} = (b_1,\, b_2,\, \ldots,\, -b_2,\, -b_1). \tag{29}$$

The number of independent amplitudes among a_r and b_r is, from Eqs. (29), equal to the total number of samples m.

For odd nonlinear order q, Eqs. (19) and (20) become trivial equations if we consider Eqs. (29) and note that ${\alpha_{0r}}^q \cos \alpha_{0r} = -{\alpha_{0,m+1-r}}^q \cos \alpha_{0,m+1-r}$ and ${\alpha_{0r}}^q \sin \alpha_{0r} = {\alpha_{0,m+1-r}}^q \sin \alpha_{0,m+1-r}$. For even order q, Eq. (21) is reduced to a trivial equation because of the similar symmetries. Therefore the number of independent equations among the $3p$ equations (19)–(21) is $[3p/2]$, where $[x]$ is the truncation function defined by

$$[x] = n \qquad \text{for}\ x = \text{integer}\ n\ \ \text{or}\ x = n + 1/2. \tag{30}$$

Since the number of independent equations among Eqs. (8)–(11) is at least $j + 2$, the total number of independent equations for group II is not less than $j + [3p/2] + 2$. Except for a few cases shown below in this subsection, the number of samples necessary for constructing the group II algorithm is equal to the number of independent equations. That is,

$$m \geqslant j + [3p/2] + 2, \tag{31}$$

where the equality holds for sinusoidal signals ($j = 1$) and for nonsinusoidal signals ($j \geqslant 2$) with a phase-shift interval of $2\pi/(j + 2)$ rad.

For nonsinusoidal signals with smaller phase-shift intervals, the necessary number of samples increases, since the number of independent equations among Eqs. (8)–(11) generally becomes larger than $j + 2$ with the smaller

Table 2. Number of Independent Equations Necessary for Constructing the Algorithms When the Sampling Amplitudes Have Symmetries[a]

Group	Immunity to Spatially Uniform and Nonuniform Phase-Shift Errors	No. of Independent Equations	Equations for Sampling Amplitudes
I	None	$j + 2$	(8)–(11)
II	Uniform	$j + p + 2$	(8)–(11), (20) and (21)
	Nonuniform	$j + [3p/2] + 2$	(8)–(11) and (19)–(21)
III	Uniform	$(p + 1)j + 2$	(8)–(11), (20), (21) and (24)–(27)
	Nonuniform	$(p + 1)j + [p/2] + 2$	(8)–(11), (19)–(21), and (24)–(27)

[a] For $j > 1$ the phase-shift interval is assumed to be equal to $2\pi/(j + 2)$ rad.

intervals. In contrast, for sinusoidal signals ($j = 1$), the necessary number of samples is $3 + [3p/2]$ for any phase-shift interval.

Note that, for a spatially nonuniform phase shift, the number of independent equations is not always equal to the minimum number of samples, since there is an exception in which the equations do not have nontrivial solutions with the minimum number of samples. In this case the number of samples necessary for constructing the algorithm is larger than the number of independent equations. We will see an example for this case in Subsection 3.B. The authors have not yet found any similar exception for a spatially uniform phase shift.

Also note that there is an unexpected reduction of the number of samples in the case in which Eqs. (8)–(11) and (19)–(21) give zero values for the pair of amplitudes a_i and b_i with one of the number $i (1 \leqslant i \leqslant m)$. In this case the necessary number of samples is less than the number of independent equations. In Subsection 3.D we will see an example of this case.

F. Number of Samples Necessary for Constructing Group III Algorithms

Here we discuss the number of samples necessary for constructing group III algorithms for a nonuniform phase shift. Also, we point out a correction in our previous paper[12] on the group III algorithms.

For group III algorithms with immunity to a nonuniform phase shift, the sampling amplitudes must satisfy Eqs. (8)–(11), (19)–(21), and (24)–(27). With the same discussion as that for group II in Subsection 2.E, the number of independent equations among Eqs. (8)–(11) and (19)–(21) is not less than $j + [3p/2] + 2$.

As for the new equations (24)–(27), the following two properties can be derived if we note the symmetries of the sampling amplitudes and the asymmetry of the phase-shift values:

a1. For odd order q, Eqs. (25) and (26) reduce to trivial equations.

a2. For even order q, Eqs. (24) and (27) reduce to trivial equations.

Moreover, if we note the two identities for arbitrary integer k and the phase-shift interval of $2\pi/(j + 2)$ rad,

$$\sin(j + 2 - k)\alpha_{0r} = (-1)^m \sin(k\alpha_{0r}), \tag{32}$$

$$\cos(j + 2 - k)\alpha_{0r} = (-1)^{m+1} \cos(k\alpha_{0r}), \tag{33}$$

where m is the number of samples and $\alpha_{0r} = [2\pi/(j + 2)]/[r - (m + 1)/2]$, we can show the following properties:

b1. Eqs. (24)–(27) for $k = 2$ are identical to those for $k = j$.

b2. Eqs. (24)–(27) for $k = 3$ are identical to those for $k = j - 1$.

b3. Eqs. (24)–(27) for $k = 4$ are identical to those for $k = j - 2$.

Etc.

c1. For even j and even m, Eqs. (25) and (27) for $k = j/2 + 1$ disappear, since $\cos(j/2 + 1)\alpha_{0r} = 0$.

c2. For even j and odd m, Eqs. (24) and (26) for $k = j/2 + 1$ disappear, since $\sin(j/2 + 1)\alpha_{0r} = 0$.

From these results we now count the number of independent equations among Eqs. (24)–(27). First, we consider these equations for a single value of parameter q.

For odd q the nontrivial equations are Eqs. (25) and (26) from result a1. From results b1–bn these $2(j - 1)$ equations reduce to $2[j/2]$ equations. Moreover, from results c1 and c2, one equation among the $2[j/2]$ equations disappears when order j is even. Therefore, for odd q, $j - 1$ equations among Eqs. (24)–(27) are nontrivial.

For even q the nontrivial equations are Eqs. (24) and (27) from result a2. With the use of results b1–bn, these $2(j - 1)$ equations reduce to $2[j/2]$ equations. From results c1 and c2, one equation among the $2[j/2]$ equations disappears for even j. Therefore, for even q, $j - 1$ equations among Eqs. (24)–(27) are nontrivial.

Finally, we sum up the number of nontrivial equations over parameter q ($q = 1, 2, \ldots, p$), which results in $p(j - 1)$ nontrivial equations. Therefore the number of independent equations among Eqs. (24)–(27) is equal to $p(j - 1)$ when the phase-shift interval is $2\pi/(j + 2)$ rad.

The total number of samples necessary for constructing group III algorithms for a nonuniform phase shift is then calculated to give

$$\begin{aligned} m &= j + [3p/2] + 2 + p(j - 1) \\ &= (p + 1)j + [p/2] + 2, \end{aligned} \tag{34}$$

where the phase-shift interval is $2\pi/(j + 2)$ rad.

Table 2 summarizes the number of independent equations for constructing group II and group III algorithms. For linear phase-shift miscalibrations ($p = 1$), the necessary numbers of samples are $j + 3$ for group II and $2j + 2$ for group III, regardless of whether the phase shift is uniform or not. These results are consistent with the results of Larkin and Oreb[8] and Surrel[9] for group II ($p = 1$) and with the result of Surrel[13] for group III.

For a quadratic and nonuniform phase shift ($p = 2$), the necessary numbers of samples are $j + 5$ for group II and $3j + 3$ for group III.

Note that there is also an exception for group III, similar to that mentioned in Subsection 2.E, in which the number of samples is not equal to the number of independent equations.

In our previous paper[12] on group III algorithms, we used a different definition for the offset value of the phase shift:

$$\alpha_{0r} = \begin{cases} \dfrac{2\pi}{n}\left(r - \dfrac{m}{2}\right) & \text{for even } m \\[2ex] \dfrac{2\pi}{n}\left(r - \dfrac{m + 1}{2}\right) & \text{for odd } m \end{cases}. \tag{35}$$

Since the phase-shift values for an even number of samples m were not symmetric ($\alpha_{01} \neq -\alpha_{0m}$), the resultant sampling amplitudes did not have the symmetries necessary for minimizing the number of samples (see Subsection 2.B). Therefore the number of samples for group III ($p = 1$) in our previous paper was not minimized and was equal to $2j + 3$. The number of samples for group III with $p = 1$ is $2j + 2$ from Table 2.

3. EXAMPLES OF PHASE-SHIFTING ALGORITHMS

A. Compensation of Quadratic and Spatially Nonuniform Phase Shift (Group II with $p = 2$ and $j = 1$)

Here we will derive a specific form of a new algorithm that compensates for a quadratic and spatially nonuniform phase shift ($p = 2$) when the signal is sinusoidal ($j = 1$).

In Table 2 six samples are necessary to construct the group II algorithms with $p = 2$ and $j = 1$. Since the signal is sinusoidal, the number of samples is 6 for any phase-shift interval. Here we assume that the phase-shift interval is $\pi/3$ rad.

From the result of Subsection 2.B, we can assume the following symmetric and asymmetric sampling amplitudes:

$$\{a_r\} = (a_1, a_2, a_3, a_3, a_2, a_1),$$
$$\{b_r\} = (b_1, b_2, b_3, -b_3, -b_2, -b_1), \tag{36}$$

where the phase shift is defined by $\alpha_{0r} = \pi(r - 7/2)/3$.

Substituting Eqs. (36) into Eqs. (8)–(11) and (19)–(21), we obtain six equations for the sampling amplitudes:

$$a_1 + a_2 + a_3 = 0, \qquad -a_1 + a_3 = 1/\sqrt{3},$$
$$b_1 + 2b_2 + b_3 = -1,$$
$$25a_1 - a_3 + 25b_1 + 18b_2 + b_3 = 0,$$
$$25a_1 - a_3 - 25b_1 - 18b_2 - b_3 = 0,$$
$$5a_1 + 6a_2 + a_3 + 5\sqrt{3}b_1 - \sqrt{3}b_3 = 0. \tag{37}$$

The solutions for Eqs. (37) are given by

$$\{a_r\} = (\sqrt{3}/72,\ -13\sqrt{3}/36,\ 25\sqrt{3}/72,\ 25\sqrt{3}/72,\ -13\sqrt{3}/36,\ \sqrt{3}/72),$$
$$\{b_r\} = (5/24,\ -1/4,\ -17/24,\ 17/24,\ 1/4,\ -5/24). \tag{38}$$

The resultant algorithm is

$$\phi = \arctan\left[\frac{\sqrt{3}(5I_1 - 6I_2 - 17I_3 + 17I_4 + 6I_5 - 5I_6)}{I_1 - 26I_2 + 25I_3 + 25I_4 - 26I_5 + I_6}\right]. \tag{39}$$

When the phase shift is quadratic and given by

$$\alpha_r = \frac{\pi}{3}\left(r - \frac{7}{2}\right)\left[1 + \epsilon_1 + \frac{\epsilon_2}{3}\left(r - \frac{7}{2}\right)\right], \tag{40}$$

the residual phase errors given by this algorithm can be calculated to give

$$\begin{aligned}\Delta\phi &= \phi - \phi_1 \\ &= -\frac{1975}{124{,}416}(\pi\epsilon_2)^2 \sin(2\phi_1) \\ &\quad - \pi^2\epsilon_1\epsilon_2\left[\frac{56}{864\sqrt{3}}\cos(2\phi_1) + \frac{265}{864\sqrt{3}}\right] + o(\epsilon^3).\end{aligned} \tag{41}$$

From Eq. (41) it can be observed that the algorithm eliminates the dc error components linear with ϵ_1 or ϵ_2.

We now numerically estimate the residual error for the new algorithm. Table 3 shows the peak-to-valley phase errors of the new algorithm, including the dc error components, for several values of ϵ_1 and ϵ_2. The variable components of the errors are shown in the parentheses, which is useful when the phase shift is spatially uniform. For comparison, Table 3 also shows the phase errors for the five-sample algorithm[10] (by Schmit and Creath)

$$\phi = \arctan\left(\frac{I_1 - 4I_2 + 4I_4 - I_5}{-I_1 - 2I_2 + 6I_3 - 2I_4 - I_5}\right) \tag{42}$$

and for the seven-sample algorithm (by de Groot[11]) as defined by Eq. (1).

In Table 3 the phase errors, including dc components, are well minimized by the new algorithm for all ϵ_2 ($\neq 0$) values. Also, the new algorithm gives better results for phase errors that are due to a spatially uniform phase shift (see the values in the parentheses in Table 3), except for $\epsilon_2 = 0$. This is because the coefficient of $\epsilon_2{}^2$ in the residual errors of the new algorithm [see Eq. (41)] is smaller than those of the two other algorithms.

B. Group II Algorithm for $p = 2$ and $j = 2$

From Table 2 a group II algorithm that compensates for a quadratic nonuniform phase shift and a second-harmonic signal ($p = 2$ and $j = 2$) needs at least seven samples when the phase-shift interval is $\pi/2$ rad. Substituting the parameters $j = 2$ and $p = 2$, the phase-shift values $\alpha_{0r} = \pi(r - 4)/2$, and the symmetric sampling amplitudes specified by Eqs. (29) with seven samples into Eqs. (8)–(11) and (19)–(21), we obtain seven linear equations:

$$2a_1 + 2a_2 + 2a_3 + a_4 = 0, \qquad -2a_2 + a_4 = 1,$$
$$-2a_1 + 2a_2 - 2a_3 + a_4 = 0, \qquad b_1 - b_3 = 1/2,$$
$$-3a_1 + a_3 + 2b_2 = 0, \qquad 8a_2 - 9b_1 + b_3 = 0,$$
$$8a_2 + 9b_1 - b_3 = 0. \tag{43}$$

Table 3. Peak-to-Valley Phase Errors That Are Due to a Quadratic and Spatially Nonuniform Phase Shift for the New Six-Sample and for the Previously Reported Seven-Sample and Five-Sample Algorithms

ϵ_1	ϵ_2	Six-Sample	Seven-Sample (de Groot)[11]	Five-Sample (Schmit and Creath)[10]
		(π rad)		
0.1	0	0.00011	0.00002	0.00031
		(0.00011)[a]	(0.00002)	(0.00031)
0	0.2	0.0030	0.10	0.055
		(0.0030)	(0.013)	(0.012)
0.1	0.2	0.012	0.099	0.062
		(0.0046)	(0.013)	(0.016)
0	0.4	0.012	0.20	0.12
		(0.012)	(0.060)	(0.049)
0.1	0.4	0.026	0.19	0.13
		(0.010)	(0.068)	(0.047)

[a] The error values in the lower parentheses are for the cases in which the phase shift is spatially uniform and dc components of the errors are neglected.

We can easily see that Eqs. (43) do not have self-consistent solutions. Therefore Eqs. (8)–(11) and (19)–(21) cannot be satisfied with seven samples; the number of samples necessary for constructing the algorithm is not equal to the number of independent equations.

These equations can be satisfied if we take eight samples instead of seven. Substituting the new phase-shift values $\alpha_{0r} = \pi(r - 9/2)/2$ and the symmetric sampling amplitudes specified by Eqs. (29) with eight samples into Eqs. (8)–(11) and (19)–(21), we obtain seven linear equations for eight components of the sampling amplitudes. Since the solutions are not unique, we choose the solutions so that the square sum $\Sigma_{r=1}^{8}(a_r^2 + b_r^2)$ is minimum. Since the square sum is proportional to the algorithm's susceptibility to random noise,[17] this choice is an optimization of the algorithm.

The solutions can be calculated to give

$$\{a_r\} = \left(-\frac{3}{32\sqrt{2}}, \frac{1}{32\sqrt{2}}, -\frac{17}{32\sqrt{2}}, \frac{19}{32\sqrt{2}}, \frac{19}{32\sqrt{2}}, -\frac{17}{32\sqrt{2}}, \frac{1}{32\sqrt{2}}, -\frac{3}{32\sqrt{2}}\right),$$

$$\{b_r\} = \left(-\frac{4}{32\sqrt{2}}, \frac{2}{32\sqrt{2}}, -\frac{14}{32\sqrt{2}}, -\frac{20}{32\sqrt{2}}, \frac{20}{32\sqrt{2}}, \frac{14}{32\sqrt{2}}, -\frac{2}{32\sqrt{2}}, \frac{4}{32\sqrt{2}}\right), \quad (44)$$

and the resultant algorithm is

$$\phi = \arctan\left[\frac{-4(I_1 - I_8) + 2(I_2 - I_7) - 14(I_3 - I_6) - 20(I_4 - I_5)}{-3(I_1 + I_8) + I_2 + I_7 - 17(I_3 + I_6) + 19(I_4 + I_5)}\right]. \quad (45)$$

C. Group III Algorithm for $p = 2$ and $j = 2$

From Table 2 a group III algorithm that compensates for a quadratic nonuniform phase shift and a second-harmonic signal ($p = 2$ and $j = 2$) needs at least nine samples when the phase shift is $\pi/2$ rad. Substituting the parameter values $j = 2$ and $p = 2$, the phase-shift values $\alpha_{0r} = \pi(r - 5)/2$, and the symmetric sampling amplitudes specified by Eqs. (29) with nine sampling amplitudes into Eqs. (8)–(11), (19)–(21), and (24)–(27), we obtain nine linear equations.

The solutions are calculated to give

$$\{a_r\} = \left(-\frac{1}{16}, -\frac{1}{4}, -\frac{1}{4}, \frac{1}{4}, \frac{5}{8}, \frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{16}\right),$$

$$\{b_r\} = \left(\frac{1}{32}, -\frac{1}{16}, -\frac{7}{16}, -\frac{9}{16}, 0, \frac{9}{16}, \frac{7}{16}, \frac{1}{16}, -\frac{1}{32}\right), \quad (46)$$

and the resultant algorithm is

$$\phi = \arctan\left[\frac{(1/2)(I_1 - I_9) - (I_2 - I_8) - 7(I_3 - I_7) - 9(I_4 - I_6)}{-(I_1 + I_9) - 4(I_2 + I_8) - 4(I_3 + I_7) + 4(I_4 + I_6) + 10I_5}\right]. \quad (47)$$

D. Reduction of the Number of Samples by Zero Solutions

In Subsection 3.E we mentioned that the number of samples necessary for constructing an algorithm is not always equal to the number of independent equations indicated in Table 2 but is reduced by unexpected zero solutions for the equations. Here we show an example for this reduction.

Consider a group II algorithm with phase-shift interval of $\pi/3$ rad that compensates for a quadratic uniform phase shift and a second-harmonic signal ($p = 2$ and $j = 2$). As shown in Table 2, the algorithm needs six samples when the phase-shift interval is $\pi/2$ rad. For smaller phase-shift intervals, the number of independent equations among Eqs. (8)–(11) is 5, corresponding to the fact that the discrete Fourier algorithm for a second-harmonic signal needs at least five samples with a phase-shift interval smaller than $\pi/2$ rad.[13] Since the number of nontrivial equations among Eqs. (20) and (21) for $p = 2$ is 2, the total number of independent equations among Eqs. (8)–(11), (20), and (21) is now 7.

Substituting $p = 2$, $j = 2$, $\alpha_{0r} = \pi(r - 4)/3$, and the symmetric sampling amplitudes with seven samples into Eqs. (8)–(11), (20), and (21) gives unique solutions

$$\{a_r\} = \left(0, -\frac{1}{2}, \frac{1}{2}, 0, \frac{1}{2}, -\frac{1}{2}, 0\right),$$

$$\{b_r\} = \left(\frac{1}{3\sqrt{3}}, -\frac{1}{2\sqrt{3}}, -\frac{1}{2\sqrt{3}}, 0, \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, -\frac{1}{3\sqrt{3}}\right), \quad (48)$$

and the resultant algorithm is given by

$$\phi = \arctan\left[\frac{-I_2 - I_3 + I_5 + I_6 + (2/3)(I_1 - I_7)}{\sqrt{3}(-I_2 + I_3 + I_5 - I_6)}\right]. \quad (49)$$

Note that the fourth amplitude a_4 becomes zero and that the algorithm is constructed with six samples, while the number of independent equations is 7.

E. Generic Solutions of Group II for Quadratic Phase Shift ($p = 2$ and Arbitrary j)

Generic solutions, which compensate for a certain order of phase-shift error and the harmonic signals up to the arbitrary order j, offer many insights into the performance of the algorithms, such as the systematic and random errors by the algorithms as a function of the number of samples.[17] Generic solutions for linear phase-shift miscalibrations ($p = 1$) were derived by several authors.[8,9,13] However, to date, generic solutions for a nonlinear phase shift ($p \geqslant 2$) have not been reported.

Generic solutions for a spatially nonuniform phase shift are somewhat difficult to derive, since the linear equations specified in Table 2 do not always have self-consistent solutions for all values of order j. We have shown an example of this in Subsection 3.B.

In contrast, generic solutions for a spatially uniform phase shift are relatively easy to derive. Here, as an example, we will derive a generic algorithm of group II that compensates for a quadratic ($p = 2$) uniform phase shift and the harmonic signals up to the jth order.

From Table 2 the algorithm needs at least $j + 4$ samples when the phase-shift interval is $2\pi/(j + 2)$ rad. We start with the sampling amplitudes with $j + 4$ components:

$$\begin{aligned}\{a_r\} &= (0, a_2, a_3, \ldots, a_{j+2}, a_{j+3}, 0)\\ &\quad + (e_1, -e_1, 0, 0, \ldots, 0, -e_1, e_1),\\ \{b_r\} &= (0, b_2, b_3, \ldots, b_{j+2}, b_{j+3}, 0)\\ &\quad + (e_2, e_2, 0, 0, \ldots, 0, -e_2, -e_2),\end{aligned} \tag{50}$$

where e_i (for $i = 1, 2$) are the unknowns to be determined and a_r and b_r are the discrete Fourier amplitudes defined by

$$a_r = \frac{2}{j+2}\cos \alpha_{0r}, \qquad b_r = \frac{2}{j+2}\sin \alpha_{0r} \quad \text{for } r = 2, 3, \ldots, j+3, \tag{51}$$

where the phase shift is defined by

$$\alpha_{0r} = \frac{2\pi}{j+2}\left(r - \frac{j+5}{2}\right). \tag{52}$$

The sampling amplitudes are shown to satisfy Eqs. (8)–(11) from the orthogonality of the Fourier amplitudes:

$$\frac{2}{j+2}\sum_{r=2}^{j+3} \sin(k\alpha_{0r})\sin \alpha_{0r} = \frac{2}{j+2}\sum_{r=2}^{j+3} \cos(k\alpha_{0r})\cos \alpha_{0r} = \delta(k, 1), \tag{53a}$$

$$\frac{2}{j+2}\sum_{r=2}^{j+3} \sin(k\alpha_{0r})\cos \alpha_{0r} = \frac{2}{j+2}\sum_{r=2}^{j+3} \cos(k\alpha_{0r})\sin \alpha_{0r} = 0. \tag{53b}$$

Substituting the sampling amplitudes into Eqs. (20) and (21) for $p = 2$ and using the relations $\alpha_{0,j+3} = \alpha_{01} + 2\pi$ and $\alpha_{0,j+4} = \alpha_{02} + 2\pi$, we have two linear equations for e_i:

$$\begin{aligned} e_1(\sin \alpha_{01} - \sin \alpha_{02}) + e_2(\cos \alpha_{01} + \cos \alpha_{02}) &= \frac{1}{\pi(j+2)}\sum_{r=2}^{j+3} \alpha_{0r}\sin(2\alpha_{0r}),\\ e_1(\cos \alpha_{01} - \cos \alpha_{02}) - e_2(\sin \alpha_{01} + \sin \alpha_{02}) &= \frac{1}{\pi(j+2)}\sum_{r=2}^{j+3} \alpha_{0r}\cos(2\alpha_{0r}),\end{aligned} \tag{54}$$

which have solutions[20]

$$e_1 = \frac{1}{2(j+2)}\frac{\cos\left(\dfrac{3\pi}{j+2}\right)}{\sin^2\left(\dfrac{2\pi}{j+2}\right)}, \qquad e_2 = \frac{1}{2(j+2)}\frac{\sin\left(\dfrac{3\pi}{j+2}\right)}{\sin^2\left(\dfrac{2\pi}{j+2}\right)}, \tag{55}$$

and the resultant algorithm is given by

$$\phi = \arctan\left\{\frac{\dfrac{1}{4}(I_1 + I_2 - I_{j+3} - I_{j+4})\dfrac{\sin\left(\dfrac{3\pi}{j+2}\right)}{\sin^2\left(\dfrac{2\pi}{j+2}\right)} + \displaystyle\sum_{r=2}^{j+3} I_r\left[\sin\left(\frac{2\pi}{j+2}\right)\left(r - \frac{j+5}{2}\right)\right]}{\dfrac{1}{4}(I_1 - I_2 - I_{j+3} + I_{j+4})\dfrac{\cos\left(\dfrac{3\pi}{j+2}\right)}{\sin^2\left(\dfrac{2\pi}{j+2}\right)} + \displaystyle\sum_{r=2}^{j+3} I_r\left[\cos\left(\frac{2\pi}{j+2}\right)\left(r - \frac{j+5}{2}\right)\right]}\right\}. \tag{56}$$

For $j = 2$ Eq. (56) reduces to the six-sample $\pi/2$-interval algorithm derived by Schmit and Creath[10] as

$$\phi = \arctan\left(\frac{I_1 - 3I_2 - 4I_3 + 4I_4 + 3I_5 - I_6}{-I_1 - 3I_2 + 4I_3 + 4I_4 - 3I_5 - I_6}\right). \tag{57}$$

The generic solution for a quadratic phase shift has a similar structure to that of the generic solutions[8,9] for a linear phase-shift miscalibration ($p = 1$) in that the sampling amplitudes consist of discrete Fourier amplitudes and several additional correction terms in both the numerator and the denominator. However, these two generic algorithms for $p = 1$ and $p = 2$ have quite different statistical properties with respect to random noise.[17]

4. FOURIER REPRESENTATION OF THE ALGORITHM

The error-compensating capability of a phase-shifting algorithm can be visualized and well understood if we take a Fourier representation of the sampling functions of the algorithm.[8,10–13,16] Freischlad and Koliopoulos showed[16] that for any phase-shifting algorithm the Fourier transforms of its two sampling functions have matched values at the fundamental frequency $\nu_0 = 2\pi/T$, where T is the period of the phase-shift parameter α, and their phases are in quadrature at this frequency. We have shown[8,12] that when a phase-shifting algorithm is insensitive to a linear phase-shift miscalibration, the first derivatives of the Fourier transforms also have matched values at frequency ν_0, and their phases are in quadrature. Surrel showed[13] that when an algorithm gives a residual error $o(\epsilon_1{}^p)$, a linear combination of the Fourier transforms has a $(p - 1)$ st-order continuity of its derivative at frequency ν_0.

In Subsection 2.C we have derived a new condition of Eq. (19) for an algorithm to become insensitive to a spatially nonuniform phase shift. Here we derive a Fourier representation of this new condition.

The sampling functions of the numerator (f_1) and the denominator (f_2) of an algorithm given by Eq. (5) are defined by

$$f_1(\alpha) = \sum_{r=1}^{m} b_r \delta(\alpha - \alpha_r), \tag{58}$$

$$f_2(\alpha) = \sum_{r=1}^{m} a_r \delta(\alpha - \alpha_r), \tag{59}$$

where $\delta(\alpha)$ is the Dirac delta function and the other parameters are as defined in Subsection 2.A. The Fourier transforms of these two functions are

$$F_1(\nu) = \sum_{r=1}^{m} b_r \exp(-i\alpha_r \nu), \tag{60}$$

$$F_2(\nu) = \sum_{r=1}^{m} a_r \exp(-i\alpha_r \nu), \tag{61}$$

where i is the imaginary unit and ν is the frequency variable. It was shown[16] that, in order to give a correct object phase, functions F_1 and F_2 need to satisfy the two following equations:

$$F_1(0) = F_2(0) = 0, \tag{62}$$

$$F_1(\nu_0) + iF_2(\nu_0) = 0. \tag{63}$$

When there is a phase-shift error, the phase-shift value α_r is shifted from the correct value α_{0r}. When an algorithm is insensitive to a nonlinear and spatially uniform phase shift of order p, the derivatives of the unperturbed function $F_{01} + iF_{02}$ take zero value at frequency ν_0.[12,13] That is,

$$\frac{\mathrm{d}^q F_{01}}{\mathrm{d}\nu^q} + i \frac{\mathrm{d}^q F_{02}}{\mathrm{d}\nu^q} = 0 \quad \text{at} \quad \nu = \nu_0 \quad \text{for} \quad q = 1, \ldots, p, \tag{64}$$

where $F_{01}(\nu)$ and $F_{02}(\nu)$ are the Fourier functions with unperturbed phase shift $\alpha_r = \alpha_{0r}$.

Now we will derive the Fourier representation of Eq. (19). Taking the qth derivative of Eqs. (60) and (61) and substituting $\alpha_r = \alpha_{0r}$ into both equations give

$$\begin{aligned}\frac{\mathrm{d}^q F_{01}}{\mathrm{d}\nu^q} - i \frac{\mathrm{d}^q F_{02}}{\mathrm{d}\nu^q} = (-i)^q \sum_{r=1}^{m} \{&[-a_r \alpha_{0r}{}^q \sin(\alpha_{0r}\nu) \\ &+ b_r \alpha_{0r}{}^q \cos(\alpha_{0r}\nu)] \\ &- i[a_r \alpha_{0r}{}^q \cos(\alpha_{0r}\nu) \\ &+ b_r \alpha_{0r}{}^q \sin(\alpha_{0r}\nu)]\}.\end{aligned} \tag{65}$$

Since the period T of the phase-shift parameter is 2π, the fundamental frequency $\nu_0 = 2\pi/T$ is now equal to unity. With $\nu_0 = 1$ and from Eq. (65), Eq. (19) can be rewritten as the following equation:

$$\mathrm{Im}\left[i^q\left(\frac{\mathrm{d}^q F_{01}}{\mathrm{d}\nu^q} - i \frac{\mathrm{d}^q F_{02}}{\mathrm{d}\nu^q}\right)\right] = 0 \quad \text{at} \quad \nu = \nu_0, \tag{66}$$

where $\mathrm{Im}(x)$ denotes the imaginary part of complex x and order q satisfies $1 \leqslant q \leqslant p$.

Note that the signs of the imaginary unit before $\mathrm{d}^q F_{02}/\mathrm{d}\nu^q$ in Eqs. (64) and (66) are different. From Eqs. (64) and (66), we can derive that, when an algorithm is insensitive to a pth-order nonlinear and spatially nonuniform phase shift, the Fourier functions F_{01} and iF_{02} have the following properties:

1. The odd-order derivatives $\mathrm{d}^q F_{01}/\mathrm{d}\nu^q$ and $i(\mathrm{d}^q F_{02}/\mathrm{d}\nu^q)$ have a matched zero real part at the fundamental frequency $\nu = \nu_0$, where order q satisfies $1 \leqslant q \leqslant p$.
2. The even-order derivatives $\mathrm{d}^q F_{01}/\mathrm{d}\nu^q$ and $i(\mathrm{d}^q F_{02}/\mathrm{d}\nu^q)$ have a matched zero imaginary part at $\nu = \nu_0$, where order q satisfies $1 \leqslant q \leqslant p$.

We have noted in Subsection 2.C that if an algorithm has the symmetries $a_r = a_{m+1-r}$ and $b_r = -b_{m+1-r}$ in the sampling amplitudes, Eq. (19) for odd q is always satisfied. Therefore an algorithm with the symmetries always satisfies property 1. However, property 2 is not common in the error-compensating algorithms so far reported.

Figure 1 shows the real and imaginary parts of the functions $(\mathrm{d}^q F_{01}/\mathrm{d}\nu^q)$ and $i(\mathrm{d}^q F_{02}/\mathrm{d}\nu^q)$ for the new six-sample algorithm $(p = 2)$ derived in Subsection 3.A [see Eq. (39)]. Both first-order derivatives have zero real parts, since functions F_{01} and iF_{02} are pure imaginary because of the symmetries of the sampling amplitudes. Therefore property 1 is satisfied for this algorithm. The imaginary parts of the second-order derivatives have matched zero values at fundamental frequency ν_0, which is consistent with property 2.

Figures 2 and 3 show the functions F_{01} and iF_{02}, the imaginary part of the second-order derivatives for the new eight-sample [see Eq. (45)] and new nine-sample [see Eq. (47)] algorithms, both of which are insensitive to a

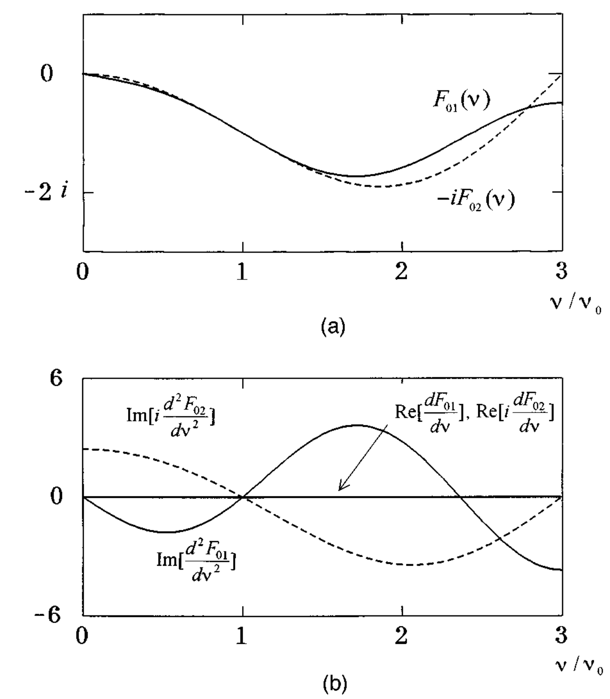

Fig. 1. (a) Functions $F_{01}(\nu)$ and $iF_{02}(\nu)$ and (b). Real and imaginary parts of their derivatives for the new six-sample algorithm, where $F_{01}(\nu)$ and $iF_{02}(\nu)$ are defined by $F_{01}(\nu) = (i/12) \times [5\sin(5\pi\nu/6) - 6\sin(\pi\nu/2) - 17\sin(\pi\nu/6)]$ and $iF_{02}(\nu) = (\sqrt{3}/36)i[\cos(5\pi\nu/6) - 26\cos(\pi\nu/2) + 25\cos(\pi\nu/6)]$.

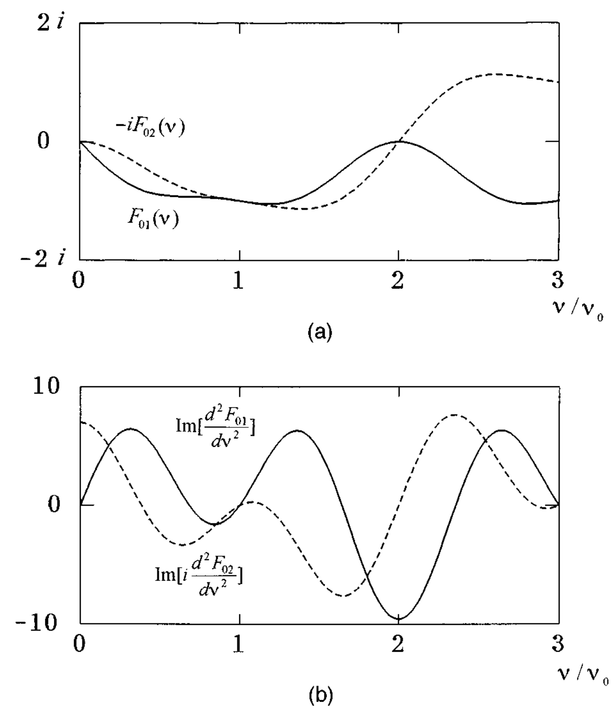

Fig. 2. (a) Functions $F_{01}(\nu)$ and $iF_{02}(\nu)$ and (b) imaginary parts of their second-order derivatives for the new eight-sample algorithm, where $F_{01}(\nu)$ and $iF_{02}(\nu)$ are defined by $F_{01}(\nu) = (\sqrt{2}/32)i[-4\sin(7\pi\nu/4) + 2\sin(5\pi\nu/4) - 14\sin(3\pi\nu/4) - 20 \times \sin(\pi\nu/4)]$ and $iF_{02}(\nu) = (\sqrt{2}/32)i[-3\cos(7\pi\nu/4) + \cos(5\pi\nu/4) - 17\cos(3\pi\nu/4) + 19\cos(\pi\nu/4)]$.

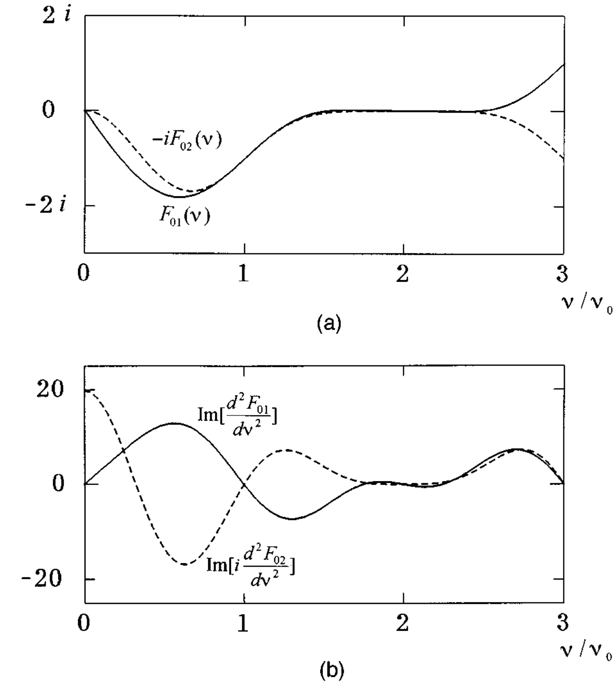

Fig. 3. (a) Functions $F_{01}(\nu)$ and $iF_{02}(\nu)$ and (b) imaginary parts of their second-order derivatives for the new nine-sample algorithm, where $F_{01}(\nu)$ and $iF_{02}(\nu)$ are defined by $F_{01}(\nu) = (i/16) \times [\sin(2\pi\nu) - 2\sin(3\pi\nu/2) - 14\sin(\pi\nu) - 18\sin(\pi\nu/2)]$ and $iF_{02}(\nu) = (i/8)[-\cos(2\pi\nu) - 4\cos(3\pi\nu/2) - 4\cos(\pi\nu) + 4\cos \times (\pi\nu/2) + 5]$.

quadratic and spatially nonuniform phase shift ($p = 2$). We can observe that both Fourier functions satisfy property 2.

5. CONCLUSION

The phase measurement errors in phase-shifting interferometry that are due to spatially nonuniform phase shifts were discussed. For linear phase-shift miscalibrations, error-compensating algorithms become insensitive to the spatial nonuniformity of the phase shift if the sampling amplitudes have symmetries in their values. However, for a nonlinear phase shift, the spatial nonuniformity gives significant errors for all error-compensating algorithms so far reported.

An analytical procedure for constructing phase-shifting algorithms that compensate for a nonlinear and spatially nonuniform phase shift and harmonic components of the signal was derived. As examples, new six-sample, eight-sample, and nine-sample algorithms were derived that are insensitive to a quadratic and nonuniform phase shift. Some characteristics of these algorithms were also discussed with the aid of a Fourier description.

REFERENCES

1. 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).

3. K. Creath, "Phase-measurement interferometry techniques," in *Progress in Optics*, E. Wolf, ed. (North-Holland Elsevier, Amsterdam, 1988), Vol. 26, pp. 349–393.
4. K. G. Larkin, "Efficient nonlinear algorithm for envelope detection in white light interferometry," J. Opt. Soc. Am. A **13**, 832–843 (1996).
5. P. Carre, "Installation et utilisation du comparateur photoelectrique et interferentiel du Bureau International des Poids et Mesures," Metrologia **2**, 13–23 (1966).
6. J. Schwider, R. Burow, K. E. Elssner, J. Grzanna, R. Spolaczyk, and K. Merkel, "Digital wavefront measuring interferometry: some systematic error sources," Appl. Opt. **22**, 3421–3432 (1983).
7. 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).
8. J. Schwider, O. Falkenstorfer, H. Schreiber, A. Zoller, and N. Streibl, "New compensating four-phase algorithm for phase-shift interferometry," Opt. Eng. **32**, 1883–1885 (1993).
9. K. G. Larkin and B. F. Oreb, "Design and assessment of symmetrical phase-shifting algorithms," J. Opt. Soc. Am. A **9**, 1740–1748 (1992).
10. Y. Surrel, "Phase stepping: a new self-calibrating algorithm," Appl. Opt. **32**, 3598–3600 (1993).
11. J. Schmit and K. Creath, "Extended averaging technique for derivation of error compensating algorithms in phase-shifting interferometry," Appl. Opt. **34**, 3610–3619 (1995).
12. P. de Groot, "Derivation of algorithms for phase-shifting interferometry using the concept of a data-sampling window," Appl. Opt. **34**, 4723–4730 (1995).
13. K. Hibino, B. F. Oreb, D. I. Farrant, and K. G. Larkin, "Phase-shifting for nonsinusoidal waveforms with phase-shift errors," J. Opt. Soc. Am. A **12**, 761–768 (1995).
14. Y. Surrel, "Design of algorithms for phase measurements by the use of phase stepping," Appl. Opt. **35**, 51–60 (1996).
15. P. J. de Groot, "Vibration in phase-shifting interferometry," J. Opt. Soc. Am. A **12**, 354–365 (1995).
16. J. Schwider, "Advanced evaluation techniques in interferometry," in *Progress in Optics*, E. Wolf, ed. (North-Holland Elsevier, Amsterdam, 1990), Vol. 28, pp. 271–359.
17. K. Freischlad and C. L. Koliopoulos, "Fourier description of digital phase-measuring interferometry," J. Opt. Soc. Am. A **7**, 542–551 (1990).
18. K. Hibino, "Susceptibility of error-compensating algorithms to random noise in phase-shifting interferometry," Appl. Opt. **36** (1997).
19. C. P. Brophy, "Effect of intensity error correlation on the computed phase of phase-shifting interferometry," J. Opt. Soc. Am. A **7**, 537–541 (1990).
20. K. A. Stetson and W. R. Brohinsky, "Electro-optic holography and its application to hologram interferometry," Appl. Opt. **24**, 3631–3637 (1985).
21. I. S. Gradshteyn and I. M. Ryzhik, "Trigonometric and hyperbolic functions," in *Table of Integrals, Series, and Products*, A. Jeffrey, ed. (Academic, London, 1980), p. 31.