Effect of piezoelectric transducer nonlinearity on phase shift interferometry

Chiayu Ai and James C. Wyant

If the nonlinearity of the motion of a piezoelectric transducer (PZT) can be described as a quadratic function, the integrated intensity of one frame in phase shift interferometry can be calculated using the Fresnel integral. For a PZT with smaller nonlinearity, the rms phase error is almost linearly proportional to the quadratic coefficient. The effects of PZT nonlinearity on the three- and the four-bucket algorithms are compared.

I. Introduction

Phase shift interferometry (PSI) techniques have been widely used in modern phase measurement instruments. The basic idea of the PSI technique is that, if the phase difference between two beams is made to vary in some known manner, the initial phase can be derived from three or more intensity measurements. The most common way to vary the phase difference between two beams is to apply a voltage to a piezoelectric transducer (PZT) on which the reference mirror is mounted. The voltage applied is a step voltage with equal period and equal increment or a ramp voltage. Then a series of interferograms is sampled and the initial phase is solved.

Most current data-reduction algorithms associated with this technique require that the phase shift between each of these sampled interferograms equals a constant, $2\pi/N$, where N is an integer.[1,2] However PZT displacement is not always a linear function of applied voltage. Hence the phase shift between two samplings is no longer the constant $2\pi/N$. Moreover the amount of phase shift will vary for each of the N samplings. Therefore the error in the phase measurement is introduced when using those algorithms which require the constant phase shift between two samplings. In practice, the PZT nonlinearity is so small that a quadratic function is sufficient to describe the PZT displacement. For the integrating-bucket method,[2] the presence of these quadratic terms causes the intensity integrals to become the Fresnel integral. For small quadratic terms, the asymptotic form of Fresnel integrals applies.

Because the four-bucket, $N = 4$, and three-bucket, $N = 3$, algorithms are widely used, the effect of the PZT quadratic nonlinearity on the phase measurement with the four- and three-bucket methods will be presented. The errors caused by these quadratic terms are periodic as are other errors,[3,4] in terms of the initial phase of the measurement, with frequency twice that of the interference fringes. In certain cases, averaging measurements made from different initial phases can be used to reduce these errors.[3] It should be noted that in optical testing the wavefront aberration is measured relative to a reference wavefront. Therefore the absolute optical path difference is not important for most phase measurements. Hence the dc bias is removed when comparison between these phase errors is made.

II. Sampling the Intensity

Let the intensity of the interference pattern be given by

$$I(x,y,t) = a(x,y) + b(x,y) \cos[\phi(x,y,t)] , \tag{1}$$

where a and b are the bias and modulation of the intensity at (x,y). $\phi(x,y,t)$ is the phase difference between two beams at (x,y) at a given time t. Because the reference mirror is mounted on a PZT, as shown in Fig. 1, the phase difference $\phi(t)$ can be expressed as

$$\phi(x,y,t) = \phi_0(x,y) + s_1 v(t) + s_2 v^2(t) , \tag{2}$$

where ϕ_0, s_1, and s_2 are the initial phase and the linear and quadratic coefficients of the PZT motion sensitivity, respectively. The voltage $v(t)$ is a step voltage or a ramp voltage. From now on, for simplicity, the arguments x and y will be omitted.

The authors are with University of Arizona, Optical Sciences Center, Tucson, Arizona 85721.

Received 17 July 1986.
0003-6935/87/061112-05$02.00/0.

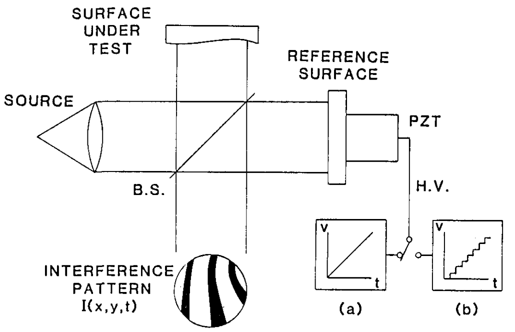

Fig. 1. Scheme of the Twyman-Green interferometer with the phase measurement technique: (a) ramp voltage in the integrating-bucket method; (b) step voltage in the phase-stepping method.

For the integrating-bucket method, the voltage applied is a ramp voltage, therefore the phase difference at a given time t is

$$\phi(t) = \phi_0 + 2\pi ct/T + 2\pi dt^2/T^2, \tag{3}$$

where c and d are the linear and quadratic coefficients, and T is the period within which the N integrating intensities are measured. The constant 2π is used for normalization. The normalization is referred to a situation where a linear PZT is exactly calibrated in order that the total phase shift is 2π in period T. In that situation c is defined as unity and d is zero. Thus for a linear PZT, which is not exactly calibrated, d is equal to 0 but c is no longer equal to 1, and a calibration error will be present in the phase measurement.[3] For a quadratic nonlinear PZT, e.g., $d = 0.01$, the PZT displacement will be one-hundredth of a wavelength longer than that when d is zero. In this paper we will investigate the effect of PZT quadratic nonlinearity on the phase measurement. Therefore emphasis will be put on the errors caused by various quadratic coefficient d when the linear coefficient c is 1. However the following derivation is not restricted to the case in which c is 1. Hence the result in this paper can be extended to the case where c is not equal to 1.

From Eqs. (1) and (3), the integrating intensity can be expressed as

$$I_n = \frac{1}{\Delta t}\int_{n\Delta t}^{(n+1)\Delta t} \left\{a + b\cos\left[\phi_0 + 2\pi c\left(\frac{t}{T}\right) + 2\pi d\left(\frac{t}{T}\right)^2\right]\right\}dt, \tag{4}$$

where $\Delta t = T/N$ for $n = 0,1,2,\ldots,N-1$. For the phase-stepping method the voltage applied is a step voltage with equal period and equal increment, so the intensity is given as

$$I_n = a + b\cos\left[\phi_0 + 2\pi c\frac{n}{N} + 2\pi d\left(\frac{n}{N}\right)^2\right], \tag{5}$$

where n is the order of step $n = 0,1,2,\ldots,N$.

III. Fresnel Integral

For the integrating-bucket method, the integrated intensity will become much more complicated than the form of Eq. (5) if a quadratic term is present in the PZT displacement. In this section we will show that the integration of each bucket can be expressed in terms of Fresnel integrals. From Eq. (4), the intensity of the first bucket is given as

$$I_0 = \frac{1}{\Delta t}\int_0^{\Delta t}\left\{a + b\cos\left[\phi_0 + 2\pi c\left(\frac{t}{T}\right) + 2\pi d\left(\frac{t}{T}\right)^2\right]\right\}dt$$

$$= a + \frac{b}{\Delta t}\int_0^{\Delta t}\cos\left[\phi_0 - \frac{\pi c^2}{2d} + 2\pi d\left(\frac{t}{T} + \frac{c}{2d}\right)^2\right]dt. \tag{6}$$

Changing the variable, the intensity can be expressed as follows: Let $x = t/T + c/2d$, then

$$I_0 = a + \frac{b}{\Delta t}\int_{\frac{c}{2d}}^{\frac{c}{2d}+\frac{1}{N}}\cos\left(\phi_0 - \frac{\pi c^2}{2d} + 2\pi dx^2\right)Tdx$$

$$= a + Nb\cos(s)\int_{\frac{c}{2d}}^{\frac{c}{2d}+\frac{1}{N}}\cos(2\pi dx^2)dx$$

$$- Nb\sin(s)\int_{\frac{c}{2d}}^{\frac{c}{2d}+\frac{1}{N}}\sin(2\pi dx^2)dx,$$

where $s = \phi_0 - (\pi c^2)/(2d)$. The integral can be simplified by letting $\tau = 2x\sqrt{d}$:

$$I_0 = a + \frac{Nb\cos(s)}{2\sqrt{d}}\int_{\omega_0}^{\omega_1}\cos\left(\frac{\pi}{2}\tau^2\right)d\tau - \frac{Nb\sin(s)}{2\sqrt{d}}\int_{\omega_0}^{\omega_1}\sin\left(\frac{\pi}{2}\tau^2\right)d\tau$$

$$= a + \frac{Nb\cos(s)}{2\sqrt{d}}\left[\int_0^{\omega_1}\cos\left(\frac{\pi}{2}\tau^2\right)d\tau - \int_0^{\omega_0}\cos\left(\frac{\pi}{2}\tau^2\right)d\tau\right]$$

$$- \frac{Nb\sin(s)}{2\sqrt{d}}\left[\int_0^{\omega_1}\sin\left(\frac{\pi}{2}\tau^2\right)d\tau - \int_0^{\omega_0}\sin\left(\frac{\pi}{2}\tau^2\right)d\tau\right], \tag{7}$$

where

$$s = \phi_0 - \frac{\pi c^2}{2d},$$

$$\omega_0 = \frac{c}{\sqrt{d}},$$

$$\omega_1 = \frac{c}{\sqrt{d}} + \frac{2\sqrt{d}}{N},$$

$$\tau = 2x\sqrt{d} = \frac{2t\sqrt{d}}{T} + \frac{c}{\sqrt{d}}.$$

Both Eqs. (6) and (7) could be expressed for the nth bucket, instead of restricting both to the first bucket. The integrals in the brackets of Eq. (7) are Fresnel integrals. Hence the integrating intensity can be expressed in terms of Fresnel integrals as

$$I_0 = a + \frac{Nb\cos(s)}{2\sqrt{d}}[\mathcal{C}(\omega_1) - \mathcal{C}(\omega_0)] - \frac{Nb\sin(s)}{2\sqrt{d}}[\mathcal{S}(\omega_1) - \mathcal{S}(\omega_0)], \tag{8}$$

where

$$\mathcal{C}(\omega) \equiv \int_0^{\omega}\cos\left(\frac{\pi}{2}\tau^2\right)d\tau,$$

$$\mathcal{S}(\omega) \equiv \int_0^{\omega}\sin\left(\frac{\pi}{2}\tau^2\right)d\tau$$

are known as the Fresnel integrals.[5] Similarly, the integrating intensity of the nth bucket can be expressed in terms of Fresnel integrals, as given by

$$I_n = a + \frac{Nb\cos(s)}{2\sqrt{d}}\left[\mathcal{C}(\omega_{n+1}) - \mathcal{C}(\omega_n)\right] - \frac{Nb\sin(s)}{2\sqrt{d}}\left[\mathcal{S}(\omega_{n+1}) - \mathcal{S}(\omega_n)\right], \quad (9)$$

where

$$\omega_n = \frac{c}{\sqrt{d}} + n\frac{2\sqrt{d}}{N}$$

for $n = 0,1,2,3,\ldots,M$. The value of M can be larger or $\leq N$.

For the four-bucket method, letting $N = 4$ in Eq. (9), the intensities of the consequent buckets can be expressed as

$$I_n = a + \frac{2b\cos(s)}{\sqrt{d}}\left[\mathcal{C}(\omega_{n+1}) - \mathcal{C}(\omega_n)\right] - \frac{2b\sin(s)}{\sqrt{d}}\left[\mathcal{S}(\omega_{n+1}) - \mathcal{S}(\omega_n)\right], \quad (10)$$

where

$$\omega_n = \frac{c}{\sqrt{d}} + n\frac{\sqrt{d}}{2}$$

for $n = 0,1,2,3,4$ and $s = \phi_0 - (\pi c^2)/(2d)$.

The Fresnel integrals $\mathcal{C}(\omega)$ and $\mathcal{S}(\omega)$ can be evaluated numerically using a series expression. In practice the normalized linear coefficient c is almost equal to 1 in order that the total phase shift is almost 2π in period T, as explained in Sec. II. Therefore the quadratic coefficient d is $\ll 1$ and ω_n is $\gg 1$. Hence the asymptotic expression for the Fresnel integral is necessary to speed up the calculation.

IV. Result

For a four-bucket method, the simple arctangent formula is widely used to solve the phase,[2]

$$\phi = \tan^{-1}\frac{I_3 - I_1}{I_0 - I_2}. \quad (11)$$

Substituting the intensities in Eq. (10) into this formula, the resulting phase ϕ differs from the initial phase ϕ_0. Because the intensity is a function of the initial phase ϕ_0, the phase error $\phi - \phi_0$ is also a function of ϕ_0. Varying the initial phase ϕ_0 from 0° to 360° and evaluating the phase error $\phi - \phi_0$, a nearly sinusoidal phase error is obtained, as illustrated in Fig. 2, curves $T1$ and $T2$. Curve $T1$ is obtained using the first four-bucket intensities, I_0, I_1, I_2, and I_3 of Eq. (10); curve $T2$ is from the last four, I_1, I_2, I_3, and I_4. The initial phase for the first four intensities differs from the initial phase for the last four intensities by ~90°. The amplitudes of two phase error curves $T1$ and $T2$ are not equal, but they are ~180° out of phase. Therefore averaging two measurements made from the first four-bucket and the last four-bucket intensities will tremendously reduce the phase error. It should be noted that there is a dc bias for each phase error curve. Because the absolute phase value is not important, the dc bias is removed when comparing the phase error curves. A similar result is obtained if the intensities of the phase-stepping method in Eq. (5) are used.[4]

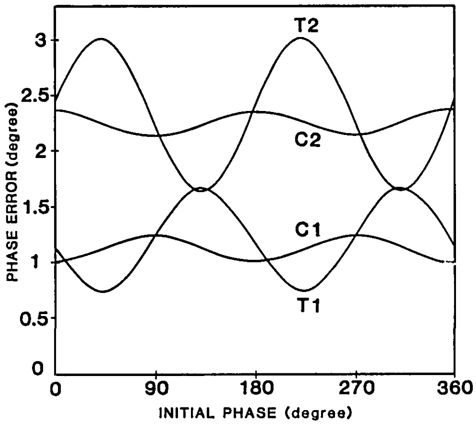

Fig. 2. Phase error caused by quadratic nonlinearity. Curve $T1$ is the phase error for various initial phases, obtained using Eq. (11) and the first four-bucket intensities; curve $T2$ obtained from the last four-bucket intensities instead. Curves $C1$ and $C2$ are the corresponding phase errors from Eq. (12). The linear and quadratic coefficients c and d equal 1 and 0.01, respectively.

Carré's formula, shown below, has been used to correct the linear calibration error[6,7,8]

$$\phi = \tan^{-1}\left[\frac{\sqrt{[(I_1 - I_2) + (I_0 - I_3)][3(I_1 - I_2) - (I_0 - I_3)]}}{(I_1 + I_2) - (I_0 + I_3)}\right]. \quad (12)$$

From this formula and the intensities in Eq. (10), similar results are obtained, as illustrated in Fig. 2, curves $C1$ and $C2$. Therefore averaging two measurements can be used to reduce the phase error as for curves $T1$ and $T2$. It should be noted that the amplitudes of curves $C1$ and $C2$ are much smaller than those of curves $T1$ and $T2$.

The rms of phase errors for the initial phase ϕ_0 ranging from 0° to 360° can be used to represent the PSI rms error caused by the PZT quadratic nonlinearity. The rms phase error, obtained using Eq. (11) or Eq. (12) and removing the dc bias, for various quadratic coefficients is shown in Fig. 3. Curves A and A' are the rms phase error from Eq. (11). Curve A' is the result of averaging two measurements made from the first four and the last four buckets. When the quadratic coefficient d is <0.1, the rms phase error is almost linearly proportional to the value of d for both curves A and A'. Similar results can be perceived for curves B and B' obtained using Eq. (12) instead.

The three-bucket algorithm, which has a phase shift of ~120° for each bucket, is given as

$$\phi = \tan^{-1}\frac{\sqrt{3}(I_2 - I_0)}{I_0 - 2I_1 + I_2}, \quad (13)$$

where the integrated intensity I_n can be obtained by letting $N = 3$ in Eq. (9). Curve P in Fig. 4 is the phase error obtained using the first three-bucket intensities

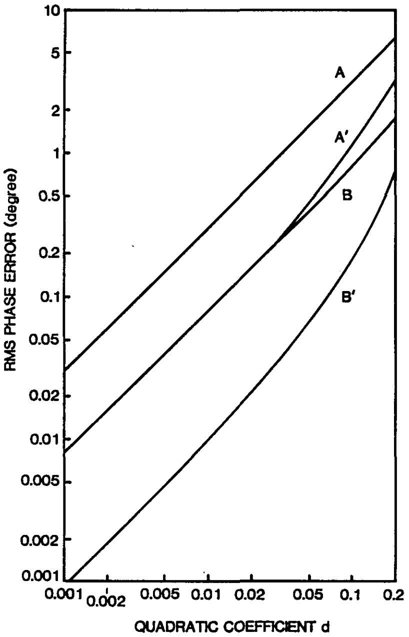

Fig. 3. Resulting rms of phase error for various quadratic coefficients. Curve A' is obtained using Eq. (11) and averaging two measurements in which the initial phase difference between the two measurements is ~90°. Curve A is obtained from Eq. (11) but without averaging. Curves B and B' are the corresponding results obtained using Eq. (12) instead.

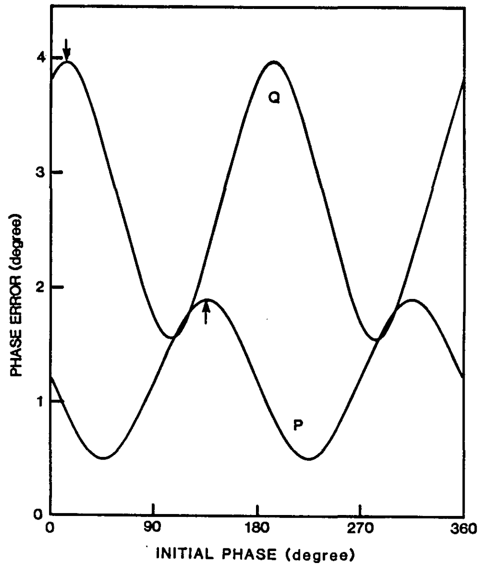

Fig. 4. Phase error caused by quadratic nonlinearity. Curves P and Q are the phase error, from Eq. (13), of two cases where the initial phases differ by ~120°. The arrow on curve Q corresponds to that on curve P; they are ~240° degree out of phase in terms of the sinusoidal curve. The linear and quadratic coefficients c and d equal 1 and 0.01, respectively.

I_0, I_1, and I_2; curve Q is from the last three buckets I_1, I_2, and I_3. The two sets have an ~120° difference in the initial phases, not 90° as in the cases of Fig. 2. The amplitudes of these two nearly sinusoidal curves are not equal and are ~120° (or 240°) out of phase with each other. Because of this 120° out of phase, instead of the 180° out of phase in Fig. 2, averaging two measurements made from the first three buckets and the last three buckets can only slightly reduce the error.

V. Discussion

Comparing the phase error curves $T1$ with $C1$ in Fig. 2, it is seen that the error caused by the quadratic nonlinearity when using Eq. (12) is smaller than the error when using Eq. (11). This can be explained as follows: In Fig. 2 the two coefficients c and d are 1 and 0.01, the quadratic PZT displacement better fits the straight line with $c > 1$ and $d = 0$. Therefore, Eq. (11), which assumes the phase shift for each bucket to 90°, will suffer from a larger error than Eq. (12) does, which can adjust for the linear calibration error.

From Figs. 2 and 4, the phase error curves are nearly sinusoidal with frequency twice that of the interference fringes. This could be due to the fact that the tangent function, used to calculate the phase, has a period of 180° which is $\lambda/2$ in terms of optical path difference. Therefore there are two ripples for each interference fringe. This explains why most of the [illegible]ous phase errors in the PSI have twice the spatial frequency.[3,4,8,9]

Because the phase error has frequency twice that of the interference fringes, if the phase shift is 90° for each bucket (i.e., $N = 4$), the phase error curves will be 180° out of phase for two measurements made from the first four and the last four buckets. Therefore averaging these two measurements can reduce the phase error. The effect of averaging two measurements on the reduction of the phase error is shown in Fig. 3.

For the larger d, e.g., $d > 0.1$, the rms error increases at a rate larger than that for $d < 0.1$; curve A' is not straight but curved upward. The reason is that for large d, e.g., 0.1, the initial phase difference between two measurements made from the first four buckets and the last four buckets is quite different from 90°. Therefore the nearly sinusoidal phase error curves are not ~180° out of phase with each other. Hence, for larger d averaging two measurements cannot reduce the error as much as for small d. This also explains curves B and B' obtained by using Eq. (12).

In Fig. 2, the initial phase difference between the first four buckets and the last four buckets is ~90°. Therefore, in principle, curve $T2$ can be obtained by shifting curve $T1$ to the left by 90° in terms of initial phase, i.e., x axis. The amplitude difference between curves $T1$ and $T2$ is due to the fact that the first four buckets have a smaller error in the phase shift for each bucket than the last four buckets. The phase shift sum should be 360° for either four buckets, but actually the sum is 363.6° for the first four buckets and 365.4° for the last four buckets. Therefore the amplitude of curve $T1$ is smaller than curve $T2$. This is also true for curves $C1$ and $C2$.

Similarly, in Fig. 4, the initial phase difference between the two sets of measurements, which result in curves P and Q, respectively, is ~120°. Therefore, curve Q can be obtained by shifting curve P to the left by 120°. Due to twice the spatial frequency in the phase error, phase error curves P and Q should be 240° out of phase with each other, as illustrated by the arrows in Fig. 4. Thus averaging two sets of measurements, obtained using Eq. (13), does not have as much effect on error reduction as using Eq. (11). Therefore the three-bucket method is not a good choice from the point of error reduction.

VI. Conclusion

If the PZT displacement can be described as a quadratic function, the integrating intensity can be calculated using the Fresnel integral. Hence the phase error caused by the PZT nonlinearity can be analyzed. The error is periodic with frequency twice that of the interference fringes. Because the absolute phase is not important in optical testing, the dc bias is removed. For the PZT with smaller nonlinearity, the rms phase error is almost linearly proportional to the quadratic coefficient d. Because of the 120° phase shift for the three-bucket method, the averaging cannot reduce the phase error as much as with the four-bucket method. With the four-bucket method, using the simple arctangent formula and averaging two measurements can have as small an amount of error as using the Carré formula without averaging. However the argument of arctangent in the Carré formula is much more complicated than that in the simple arctangent formula. Therefore a convenient way to reduce this error is to take two measurements with the four-bucket method in the simple arctangent formula, and average them.

References

1. J. H. Bruning, "Fringe Scanning Interferometers," in *Optical Shop Testing*, D. Malacara, Ed. (Wiley, New York, 1978).
2. J. C. Wyant, "Use of an ac Heterodyne Lateral Shear Interferometer with Real-Time Wavefront Correction Systems," Appl. Opt. **14**, 2622 (1975).
3. J. C. Wyant, "Review of phase shifting interferometry," in *Technical Digest, OSA Spring Meeting* (Optical Society of America, Washington, DC, 1984).
4. C. L. Koliopoulos, "Interferometric Optical Phase Measurement Techniques," Ph.D. Dissertation, U. Arizona, Tucson (1981).
5. M. Born and E. Wolf, *Principles of Optics* (Pergamon, New York, 1975), p. 430.
6. P. Carré, "Installation et utilisation du comparateur photoélectrique et interférentiel du Bureau International des Poids et Mesures," Metrologia **2**, 13 (1966).
7. P. Hariharan, B. F. Oreb, and N. Brown, "Real-Time Holographic Interferometry: A Microcomputer System for The Measurement of Vector Displacements," Appl. Opt. **22**, 876 (1983).
8. Y.-Y. Cheng and J. C. Wyant, "Phase Shifter Calibration in Phase-Shifting Interferometry," Appl. Opt. **24**, 3049 (1985).
9. J. Schwider, R. Burow, K.-E. Elssner, J. Grzanna, R. Spolaczyk, and K. Merkel, "Digital Wave-Front Measuring Interferometry: Some Systematic Error Sources," Appl. Opt. **22**, 3421 (1983).
