11-1-1990

# An Analysis of Polarization Mixing Errors in Distance Measuring Interferometers

Walter Augustyn

Paul Davis
*Worecester Polytechnic Institute*, pwdavis@wpi.edu

# An analysis of polarization mixing errors in distance measuring interferometers

Walter Augustyn
*CMX Systems, Inc., 130 Research Parkway, Meriden, Connecticut 06450*

Paul Davis
*Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, Massachusetts 01609*

(Received 29 May 1990; accepted 1 August 1990)

Many distance measuring interferometers use orthogonally polarized beams in their reference and measurement arms. In practice, instrumental imperfections corrupt the polarized beam in each arm with a small amplitude signal from the other arm having the perpendicular polarization. The impact of this polarization mixing on the absolute accuracy of a single frequency laser interferometer is shown to be significantly smaller than on the accuracy of a two-frequency interferometer. Certain signals obtained from a single-frequency interferometer have *no* phase error due to polarization mixing. Other signals are affected in proportion to the ratio between the small mixing amplitudes and the primary polarization amplitudes. Roughly, those errors in the single frequency device depend on the *square* of the small mixing ratios while the error in the two frequency interferometer depends on a larger quantity, the *sum* of these ratios. For a mixing ratio of 0.1, the error in the single-frequency interferometer is about 20 times smaller than in the heterodyne interferometer.

## I. INTRODUCTION

Helium–neon distance measuring interferometers (DMIs) are widely used in several industrial markets for accurate linear position measurements as well as for other metrology purposes. Most commercial instruments available today use either heterodyne detection methods for sensing velocity or fringe counting methods for sensing position directly.

Fringe counting or single frequency DMIs use polarization amplitude splitting to generate orthogonally polarized measuring and reference beams. These beams are then recombined with the aid of additional polarizing devices in front of two or more detectors to produce the observable interference signals.

Heterodyne or two frequency DMIs generate two orthogonally polarized beams of slightly different frequencies. Typically, these frequency differences are in the range of 2–20 MHz or greater. These two beams are combined with the aid of additional polarizing devices in front of a single detector to produce the observable interference signal.

Both technologies convert the ac component of the interference signal to square waves. In single-frequency systems, the edges of these pulses correspond to position changes of the target. In two-frequency systems, the difference in frequency between the measurement interference signal and a reference frequency interference signal, which is generated internal to the laser source, is proportional to the target velocity.

In either technology, the edge positions of the wave trains are determined by the zero crossings of the ac component of the interference signal. If the ac signals include harmonic distortion, then edge position errors will occur. This results in a reduction of absolute accuracy. Additional signal processing with either technology cannot offset this error since it is present in the original optical interference signals.

For the past several years, the demand for commercial systems possessing absolute accuracies of less than 10 nm ($\lambda/64$) has been increasing.

Several error sources limit absolute accuracy. One is path length changes caused by atmospheric variations. Instrumental nonlinearites, rather than environmental changes, form another source. The latter are the subject of this paper.

These nonlinearities were first described for heterodyne systems by Bobroff.[1] They arise from polarization mixing, which causes a small fraction of the reference frequency to travel over the measurement path and vice versa. This mixing introduces distortion in the interference signal which can be calculated explicitly.

In two frequency systems the mixing can arise from the nonorthogonality of the electric vector of each frequency, misalignment of the laser source relative to the interferometer, and imperfections in the polarizing beam splitter. In single-frequency systems, they arise only from imperfections in the polarizing beam splitter.

In Sec. IV below, we carry out the calculations needed to exhibit in detail the signal distortion which arises in a single-frequency interferometer from polarization mixing. In Sec. V, our analysis demonstrates that some signals in a single frequency device remain free of zero crossing error while others show a small error. The expressions we obtain for this error reveal that it is always significantly smaller than the heterodyne error.

Section II provides background, including a description of the particular single-frequency device we shall study. Section III analyzes the single-frequency interferometer in the ideal case, without polarization mixing, when the zero crossings of the combined interference wave carry precisely the position changes of the target.

## II. BACKGROUND

Bobroff[1] showed that when polarization mixing occurs in a heterodyne system, the phase distortion between the cor-

 0734-211X/90/062032-05$01.00  

rupted interference signal and the reference signal generated internally from the laser source is proportional to a zero crossing error $\delta$,

$$\tan \delta = \alpha/B + \beta/A. \tag{1}$$

Here $A$ and $B$ are the amplitudes of the reference frequency and measuring frequency, respectively, and $\alpha$ and $\beta$ are the amplitudes of the reference frequency and measuring frequency which result from instrumental imperfections or misalignments.

This paper derives similar expressions for zero crossing errors for fringe counting interferometers in order to compare the accuracy of the two approaches. We consider the system described in Ref. 2.

A fringe counting interferometer employs a polarizing beam splitter to create linearly polarized, orthogonal reference, and target beams of the same frequency. These are illustrated in Fig. 1, where the $P$ polarized light vector is transmitted toward the target reflector.

The position of the target is established in the fringe-counting optical phase decoder by measuring the phase difference between the two polarized components after they pass through a quarter-wave plate. The quarter-wave plate is positioned so that the two beams become circularly polarized with opposite parity. This measurement is facilitated by introducing a specific, fixed phase relation between the two beams through linear polarizers mounted in front of each of the two detectors. The theory derived below is based on the operational principles of the CMX Systems, Inc. Model 502A Optical Phase Decoder (OPD).

## III. THE IDEAL CASE

Suppose two parallel plane waves have been linearly polarized so that their electric vectors are orthogonal. Suppose that the phase difference between these waves is $d$. After passing through a quarter-wave plate oriented with its axis at 45° to the two vectors, one will be right circularly polarized and the other will be left circularly polarized. By correctly positioning a coordinate system with its $z$ axis along the direction of propagation of the two waves, we may write the $x$ and $y$ components of the electric vectors of these two waves as

$$E_{R,x} = A\cos\omega t, \quad E_{R,y} = -A\sin\omega t, \tag{2}$$

and

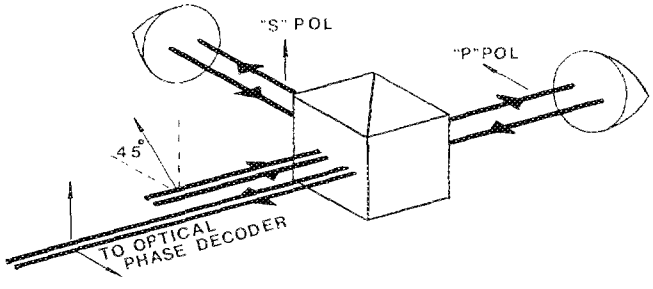


FIG. 1. A diagram of the single-frequency interferometer system. The $P$ polarized light vector is transmitted toward the target reflector on the right.

$$E_{L,x} = B\cos(\omega t + d), \quad E_{L,y} = B\sin(\omega t + d). \tag{3}$$

Here the subscripts $R$ and $L$ denote right- and left-circularly polarized, respectively, and the subscripts $x$ and $y$ denote the components of the electric vector in those coordinate directions.

Since the amplitudes of the two sets of waves add, the intensity of the interfering wave is the intensity of the wave whose electric vector is the vector sum of the two vectors $\mathbf{E}_R$ and $\mathbf{E}_L$. That is, the total intensity $I$ at the optical phase decoder in Fig. 1 is

$$I = |\mathbf{E}_R + \mathbf{E}_L|^2, \tag{4}$$

where $|\cdots|$ denotes the magnitude of a vector. (Alternatively, we could represent the total electric vector in complex notation and compute intensity as the product of the electric vector with its complex conjugate.) Carrying out the necessary calculations yields the time-varying intensity of the sum of the right- and left-circularly polarized beams,

$$I(t) = A^2 + B^2 + 2AB\cos(2\omega t + d). \tag{5}$$

To recover the phase information in this combined wave, we pass it through a polarizer set at an angle $\theta$ to the $x$ axis. The electric vector $\mathbf{E}_\theta$ after the polarizer will be the projection of the original electric vector onto the polarizer direction $\theta$. That is, the electric vector after the polarizer has magnitude

$$|\mathbf{E}_\theta| = (E_{R,x} + E_{L,x})\cos\theta + (E_{R,y} + E_{L,y})\sin\theta,$$

and its direction is that of the polarizer angle $\theta$. Just as in Eq. (4), we find that the intensity after the polarizer is

$$\begin{aligned} I_\theta(t) &= |\mathbf{E}_\theta|^2 \\ &= [(E_{R,x} + E_{L,x})\cos\theta + (E_{R,y} + E_{L,y})\sin\theta]^2 \\ &= E_x^2\cos^2\theta + 2E_xE_y\cos\theta\sin\theta + E_y^2\sin^2\theta, \end{aligned} \tag{6}$$

where

$$E_x = A\cos\omega t + B\cos(\omega t + d), \tag{7}$$

$$E_y = -A\sin\omega t + B\sin(\omega t + d). \tag{8}$$

Detectors used in these interferometers actually record time-averaged intensities, which we denote by $\langle I\rangle$. Analytically, we define (e.g., see Ref. 3)

$$\langle I\rangle = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} I(t)dt. \tag{9}$$

In practice, of course, the infinite limit is not taken. It suffices merely to record intensities over a time interval of length $2T$ which is much greater than the period $2\pi/\omega$ of the incident signal.

Applying the averaging operation of Eq. (9) to the intensity $I(t)$ before the polarizer, which is given by Eq. (5), we find

$$\langle I\rangle = A^2 + B^2. \tag{10}$$

This part of the signal is used to subtract the dc offset of the interferometer's signals in the optical phase decoder.

Since the polarizer angle terms in Eq. (6) are independent of time, the time-averaged intensity after a polarizer at angle $\theta$ is

$$\langle I_\theta \rangle = \langle E_x^2 \rangle \cos^2\theta + 2\langle E_x E_y \rangle \cos\theta \sin\theta + \langle E_y^2 \rangle \sin^2\theta, \tag{11}$$

where the time-averaged terms follow from using Eqs. (7) and (8) in (9):

$$\langle E_x^2 \rangle = \tfrac{1}{2}(A^2 + B^2) + AB\cos d,$$
$$\langle E_y^2 \rangle = \tfrac{1}{2}(A^2 + B^2) - AB\cos d,$$
$$\langle E_x E_y \rangle = AB \sin d.$$

Equation (11) immediately reveals that the intensity recorded after a polarizer set parallel to the $x$ axis (i.e., at $\theta = 0$) is

$$\langle I_0 \rangle = \langle E_x^2 \rangle = \tfrac{1}{2}(A^2 + B^2) + AB\cos d. \tag{12}$$

The intensity recorded after a polarizer set at a 45° angle is

$$\langle I_{45} \rangle = \tfrac{1}{2}(\langle E_x^2 \rangle + 2\langle E_x E_y \rangle + \langle E_y^2 \rangle)$$
$$= \tfrac{1}{2}(A^2 + B^2) + AB \sin d. \tag{13}$$

Combinations of the three time-averaged intensities (10), (12), and (13) then produce four signals whose zero crossings identify phase values of $d = 0°, 90°, 45°$, and 135°, respectively. These signals are

$$S_1 = \langle I_{45} \rangle - \tfrac{1}{2}\langle I \rangle = AB \sin d, \tag{14}$$
$$S_2 = \langle I_0 \rangle - \tfrac{1}{2}\langle I \rangle = AB \cos d, \tag{15}$$
$$S_1 - S_2 = AB \sin d - AB \cos d, \tag{16}$$
$$S_1 + S_2 = AB \sin d + AB \cos d. \tag{17}$$

The incoming wave is divided in the optical phase decoder so that the actual intensities incident upon the reference detector, which records $\langle I \rangle$, and upon the detectors after the 0° and 45° polarizers are reduced by one third. Hence, the amplitudes of the four signals just defined are also one-third of the values shown. Of course, this constant multiplier does not affect our analysis of the zeros of these signals, and we omit it for simplicity.

In the next section, we describe how these signals are corrupted by polarization mixing. We then analyze in detail the error in phase angle introduced by that mixing.

## IV. MIXED POLARIZATION

Now suppose that the electric vector of each linearly polarized plane wave includes a small component of the electric vector of the *other* wave, as can occur in interferometers when a small fraction of the $P$ polarization mixes with the $S$ and vice versa.

Polarizing beam splitters are characterized by a figure of merit known as the *extinction ratio.* It is the ratio of the intensity of the $S$ polarized light reflected to that of the $P$ polarized light along the $S$ channel; this ratio is applied to the transmission channel as well. Polarization mixing can therefore occur in beam splitters with extinction ratios as high as 1000:1.

Because of the polarization mixing, the right-rotating electric vector $\mathbf{E}_R$ that emerges from the quarter-wave plate includes a small component which is phase shifted by $d$, while the left-rotating vector $\mathbf{E}_L$ includes a small component which is not shifted. That is, Eqs. (2) and (3) are replaced by

$$E_{R,x} = A\cos\omega t + \beta\cos(\omega t + d),$$
$$E_{R,y} = -A\sin\omega t - \beta\sin(\omega t + d), \tag{18}$$

and

$$E_{L,x} = B\cos(\omega t + d) + \alpha\cos\omega t,$$
$$E_{L,y} = B\sin(\omega t + d) + \alpha\sin\omega t. \tag{19}$$

Each circularly polarized wave has been corrupted by a small out of phase signal, whose amplitude is denoted by the corresponding lower case Greek letter. (The reference arm amplitudes are $A$, $\alpha$. The measurement arm amplitudes are $B$, $\beta$. The signals with amplitudes $A$, $\beta$, are $S$ polarized before the quarter-wave plate. The signals with amplitudes $B$, $\alpha$, are $P$ polarized before the quarter-wave plate.)

We can now repeat the analysis of the preceding section with the added algebraic complexity of the polarization mixing amplitudes appearing in Eqs. (18) and (19).

The time-varying intensity $I(t)$ that results from combining these two corrupted waves is computed just as in Eq. (5) to obtain

$$I(t) = [(A+\alpha)\cos\omega t + (\beta + B)\cos(\omega t + d)]^2$$
$$+ [(-A+\alpha)\sin\omega t + (-\beta + B)\sin(\omega t + d)]^2$$
$$= A^2 + B^2 + \alpha^2 + \beta^2$$
$$+ 2A\alpha\cos 2\omega t + 2\beta B\cos(2\omega t + 2d)$$
$$+ 2(A\beta + B\alpha)\cos d + 2(AB + \alpha\beta)\cos(2\omega t + d). \tag{20}$$

The time-averaged intensity before the polarizer is now

$$\langle I \rangle = A^2 + B^2 + \alpha^2 + \beta^2 + 2(\alpha B + \beta A)\cos d \tag{21}$$

in place of Eq. (10). Note that polarization mixing has added a small modulation to what was a dc signal in the ideal case.

After passing this combined wave through a polarizer set at angle $\theta$, we obtain the time-varying intensity expression $I_\theta(t)$ given by Eq. (6) but with the following expressions for $E_x$, $E_y$ in place of Eqs. (7) and (8):

$$E_x = (A+\alpha)\cos\omega t + (B+\beta)\cos(\omega t + d), \tag{22}$$
$$E_y = (-A+\alpha)\sin\omega t + (B-\beta)\sin(\omega t + d). \tag{23}$$

The corresponding time averages required for Eq. (11) are

$$\langle E_x^2 \rangle = \tfrac{1}{2}[(A+\alpha)^2 + (B+\beta)^2]$$
$$+ (A+\alpha)(B+\beta)\cos d,$$
$$\langle E_y^2 \rangle = \tfrac{1}{2}[(A-\alpha)^2 + (-B+\beta)^2]$$
$$+ (A-\alpha)(-B+\beta)\cos d,$$
$$\langle E_x E_y \rangle = (AB - \alpha\beta)\sin d.$$

From Eq. (11), we then find that the time-averaged intensity after a polarizer set at $\theta = 0$ is

$$\langle I_0 \rangle = \tfrac{1}{2}[(A+\alpha)^2 + (B+\beta)^2] + (A+\alpha)(B+\beta)\cos d; \tag{24}$$

cf. Eq. (12). In place of Eq. (13), we find that the intensity recorded after a polarizer set at a 45° angle is

$$\langle I_{45}\rangle = \tfrac{1}{2}(A^2 + B^2 + \alpha^2 + \beta^2) + (\beta A + \alpha B)\cos d + (AB - \alpha\beta)\sin d. \quad (25)$$

Combinations of these three time-averaged intensities (21), (24), and (25) then produce the four signals whose zero crossings identify phase values of $d = 0°$, 90°, 45°, and 135°, respectively. However, we do not directly implement the forms of $S_1$ and $S_2$ given by Eqs. (14) and (15) in the ideal case, for such signals would now contain dc terms arising from the additional modulation introduced by polarization mixing. Instead, we construct the signals

$$S_1 = \langle I_{45}\rangle - f_{45}\langle I\rangle,$$
$$S_2 = \langle I_0\rangle - f_0\langle I\rangle,$$

where the attenuation factors $f_0, f_{45}$ are chosen so that $S_1, S_2$ each have zero dc level.

Practically, the attenuation factors $f_0$ and $f_{45}$ are set by adjusting potentiometers in the optical phase decoder circuitry so that $S_1$ and $S_2$ show zero dc level on an oscilloscope. Analytically, the same result is achieved by choosing these factors so that the average of the maximum and minimum values of each signal is zero. From the latter procedure, we obtain

$$f_0 = \frac{1}{2} + \frac{\alpha A + \beta B}{A^2 + B^2 + \alpha^2 + \beta^2}, \quad (26)$$

$$f_{45} = \tfrac{1}{2}. \quad (27)$$

The ideal values which appear in Eqs. (14) and (15) are $f_0 = f_{45} = 1/2$.

Note that the calibration of these attentuation factors is not affected by subsequent variations in the intensity of the laser source; $f_{45}$ is independent of amplitude, and $f_0$ is invariant because the four amplitudes $A$, $B$, $\alpha$, and $\beta$ all change by the same proportion.

Using the values of the attenuation factors given in Eqs. (26) and (27), we obtain

$$S_1 = (AB - \alpha\beta)\sin d. \quad (28)$$

$$S_2 = (AB - \alpha\beta)\frac{1 - r}{1 + r}\cos d, \quad (29)$$

$$S_1 - S_2 = (AB - \alpha\beta)\left(\sin d - \frac{1 - r}{1 + r}\cos d\right), \quad (30)$$

$$S_1 + S_2 = (AB - \alpha\beta)\left(\sin d + \frac{1 - r}{1 + r}\cos d\right), \quad (31)$$

where $r = (\alpha^2 + \beta^2)/(A^2 + B^2)$. In Sec. V, we analyze the errors in the zero crossings of these signals which are induced by the additional cross-polarization terms containing $\alpha$ or $\beta$.

## V. ERRORS FROM POLARIZATION MIXING

The zero crossings of the four signals $S_1, S_2, S_1 \pm S_2$ defined by Eqs. (28)–(31) are used to determine when the phase difference $d$ crosses 0°, 90°, 45°, and 135°, respectively. The actual corresponding target position changes depend on the interferometer configuration. For a four-pass interferometer, the position change would be $\lambda/64$. Errors in the zero crossing values produce proportionate errors in these relative position values. In this section, we obtain expressions for the zero crossing errors in these signals.

The zero crossings of $S_1$ and $S_2$ are determined by setting the expressions in Eqs. (28) and (29) to zero and solving for $d$. We see immediately that even in the face of polarization mixing, the zero crossings of $S_1$ and $S_2$ are *error free*:

$S_1 = 0$ *exactly* at $d = 0°$,

$S_2 = 0$ *exactly* at $d = 90°$.

That is, the accuracy of the 0° and 90° position measurements is not affected by polarization mixing, in contrast to the heterodyne interferometer, whose error is given by Eq. (1).

By setting Eqs. (30) and (31) to zero, we find that the zero crossings of $S_1 \pm S_2$ occur when

$$\tan d = \mp(1 - r)/(1 + r), \quad (32)$$

where, as defined in the preceding section,

$$r = (\alpha^2 + \beta^2)/(A^2 + B^2).$$

When $\alpha$ and $\beta$ are small relative to $A$ and $B$, then $r \ll 1$ and $\tan d \approx \mp 1$. In the ideal case, we see from Eqs. (16) and (17) that the zero crossings occur precisely when $\tan d = \mp 1$.

To determine the zero crossing error $\delta$ for $S_1 - S_2$, write $d = 45° - \delta$. Then Eq. (32) and the identity for the tangent of the sum of two angles together yield an expression for the zero crossing error $\delta$ in $S_1 - S_2$,

$$\tan\delta = r = (\alpha^2 + \beta^2)/(A^2 + B^2). \quad (33)$$

Similarly, for the signal $S_1 + S_2$ we define its zero crossing error $\delta$ by $d = 135° + \delta$ and find that $\delta$ also satisfies Eq. (33).

Equation (33) defines the *zero crossing error* due to polarization mixing for the 45° and 135° signals in the single frequency interferometer. (Recall that the zero crossings of the 0° and 90° signals are completely unaffected by polarization mixing.) From the definitions of $\delta$ for the 45° and 135° zero crossings, we see that the effect of polarization mixing is to lead the 45° crossing and to lag the 135° crossing. Note that the error expression (33) is exact; no approximations were made in its derivation.

The error expression (33) for a single-frequency system is analogous to (1) for a heterodyne device. The single-frequency error term (33) is smaller because it depends upon squares of the relative polarization mixing amplitudes rather than upon a sum, as in the heterodyne expression (1).

To provide a numerical comparison between the error in the two devices, consider the cross polarization ratio values $\alpha/B = 0.11$ and $\beta/A = 0.09$ used in Ref. 1. The heterodyne error from (1) is 11.31°, but the single frequency error from (33) (assuming $A = B$) is just 0.58°, nearly *20 times smaller.*

We can also provide upper bounds on the error in both devices in terms of the larger of the two cross-polarization ratios. Let $\epsilon_{\max}$ denote this larger ratio; i.e.,

$$\epsilon_{\max} = \max(\alpha/B, \beta/A).$$

Then from Eq. (1), the phase angle error in the heterodyne device satisfies

$$\tan\delta_{\text{het}} \leqslant 2\epsilon_{\max}. \quad (34)$$

Since $\alpha^2 \leqslant A^2\epsilon_{\max}^2$ and $\beta^2 \leqslant B^2\epsilon_{\max}^2$, we find from Eq. (33)

that the phase angle error in the single frequency interferometer is bounded by

$$\tan \delta_{sf} \leqslant \epsilon_{\max}^2. \tag{35}$$

Since the single-frequency error bound is smaller than the heterodyne error bound by a factor of $\epsilon_{\max}/2$, the single-frequency zero crossing is more accurate by a factor of $2/\epsilon_{\max}$, or about a factor of 20 when the cross-polarization ratios are in the range of 1/10. Since the larger cross-polarization ratio is effectively the square root of the extinction ratio of the beam splitter, we can say that the single-frequency phase angle is more accurate than the heterodyne phase angle by a factor of twice the square root of the beam splitter's extinction ratio.

For example, with a beam splitter with a high extinction ratio like 1000:1, we expect the single-frequency system to be about $2\sqrt{1000} \simeq 63$ times more accurate than the heterodyne interferometer. Explicitly, if the extinction ratio is 1000:1, then $\epsilon_{\max} = \sqrt{0.001} \simeq 0.03$. From (34), the heterodyne error $\delta_{\text{het}}$ is bounded by 3.62° while Eq. (35) shows that the single frequency error $\delta_{sf}$ is bounded by 0.057°.

In this case, the single-frequency system is at least 60 times more accurate. The 3.62° error bound of the two-frequency system corresponds to a target position error of about $\lambda/200$ [In a single-pass interferometer, a phase difference of 360° corresponds to a target position change of $\lambda/2$. Hence, an error of 3.62° corresponds to a position error of $(3.62/360)(\lambda/2) \simeq \lambda/200$.] while the 0.057° error bound of the single frequency system is about $\lambda/12\,000$.

The error bounds (34) and (35) show that the accuracy advantage of the single-frequency system increases with the extinction ratio of the beam splitter (more precisely, like the square root of the extinction ratio).

## VI. CONCLUSIONS

We have derived expressions (28)–(31) for the four signals which determine the 0°, 90°, 45°, and 135° zero crossings in the single-frequency interferometer system described in Ref. 2. These show that the 0° and 90° zero crossings are *completely unaffected* by polarization mixing.

Furthermore, we find that the errors for the 45° and 135° zero crossings in the single-frequency system are significantly smaller than the phase angle error found in two frequency systems.[1] Explicit expressions for the error in the single frequency system show that it is more accurate than a heterodyne system by a factor of one-half the cross-polarization ratio. Hence, the relative advantage of the single-frequency interferometer improves with the extinction ratio of the beam splitter.

In summary, the significant result is that single-frequency fringe counting systems offer inherently higher absolute accuracy for critical measurements. Accuracies for polarization mixings on the order of 1% of peak intensities can be in the range of $\lambda/1200$, about 20 times more accurate than a two-frequency system.

[1] N. Bobroff, Appl. Opt. **26**, 2676 (1987).
[2] Model 502A Optical Phase Decoder Theory of Operation, CMX Tech. Bull. No. 6, CMX Systems Inc., Meriden, CT, May 12, 1989.
[3] M. Born and E. Wolf, *Principles of Optics* (Pergammon, New York, 1980).