Investigation and compensation of the nonlinearity of heterodyne interferometers

Wenmei Hou* and Günter Wilkening‡
*Research fellow of the Alexander von Humboldt-Stiftung, Bonn, and
‡Physikalisch-Technische Bundesanstalt, Braunschweig, Germany

The nonlinearity of heterodyne interferometers is mainly influenced by the nonorthogonality and ellipticity of the linearly polarized partial beams of the laser and by the limited extinction capability of the polarizing beam splitters. The phase errors at the interferometer output can be detected by measuring the phase difference between the two orthogonal directions in the output beam. This arrangement can be used to check an interferometer setup for causes of nonlinearity and to eliminate them by adjustment, replacement, etc. The arithmetical mean value of the phases of the two orthogonal output signals is linear, except for a residual error. These detection and compensation capabilities of the proposed phase measurement arrangement are demonstrated by experimental investigations.

Keywords: *heterodyne interferometry; nonorthogonality; displacement interferometry; high-accuracy interferometry*

Heterodyne interferometer

The laser interferometer has become firmly established in length measuring technique. As the wavelength of the light is used as a measure, a relatively simple, direct linkup with the definition of the length is possible. In the acceptance and calibration of machine tools and coordinate measuring machines, laser interferometers are used as a universal means of measuring length, angle, and planeness. Laser interferometers are used as measuring systems integrated into highly precise machines for microfinishing and for the manufacture of integrated circuits. The manufacture of ever smaller structures implies that the uncertainty requirements to be met by length measurements in the submicrometer range become increasingly higher. The light wavelength by no means suffices as an incremental unit, and subdivisions down to the nanometer region are therefore necessary. The two-frequency interferometer has special advantages for this purpose, as it allows a very stable subdivision of the light wavelength with high resolution on the basis of a phase measurement.

The basic setup of a two-frequency interferometer is shown in *Figure 1*. In a beam of Gaussian profile, the laser (so far, He-Ne lasers are used exclusively) emits two partial beams of different frequencies (f_1, f_2) polarized orthogonally to each other (*Figure 1b*). These two partial beams are described as plane waves:

$$\begin{aligned} E_1 &= E_{01}\sin(2\pi f_1 t + \varphi_{r1}) \\ E_2 &= E_{02}\sin(2\pi f_2 t + \varphi_{r2}) \end{aligned} \tag{1}$$

φ_{r1}, φ_{r2} are initial, constant phases.

One part of the laser beam is split off in front of the interferometer, passes through a polarizing filter (at 45° to the polarizing directions of partial beams x_{1L}, x_{2L}), and falls on diode D_r. The partial beams are superimposed and generate a photoelectric current I_r in the diode:

$$\begin{aligned} I_r' &\sim E_r^2 \\ &\sim \tfrac{1}{2}[E_{01}\sin(2\pi f_1 t + \varphi_{r1}) + E_{02}\sin(2\pi f_2 t + \varphi_{r2})]^2 \\ &\sim \tfrac{1}{2}\{E_{01}^2\sin^2(2\pi f_1 t + \varphi_{r1}) \\ &\quad + E_{02}^2\sin^2(2\pi f_2 t + \varphi_{r2}) \\ &\quad + E_{01}E_{02}\cos[2\pi(f_1 - f_2)t + (\varphi_{r1} - \varphi_{r2})] \\ &\quad - E_{01}E_{02}\cos[2\pi(f_1 + f_2)t + (\varphi_{r1} + \varphi_{r2})]\} \\ &= I(f_1) + I(f_2) + I(f_1 - f_2) + I(f_1 + f_2) \end{aligned} \tag{2}$$

The photodiodes can transform only the difference term $I(f_1 - f_2)$ into an AC signal. The other terms lead to DC contributions without information

† *Permanent address: Beijing Institute of Machine Building Industry, Beijing, People's Republic of China.*
Address reprint requests to Günter Wilkening, Physikalisch-Technische Bundesanstalt, Braunschweig, Germany.

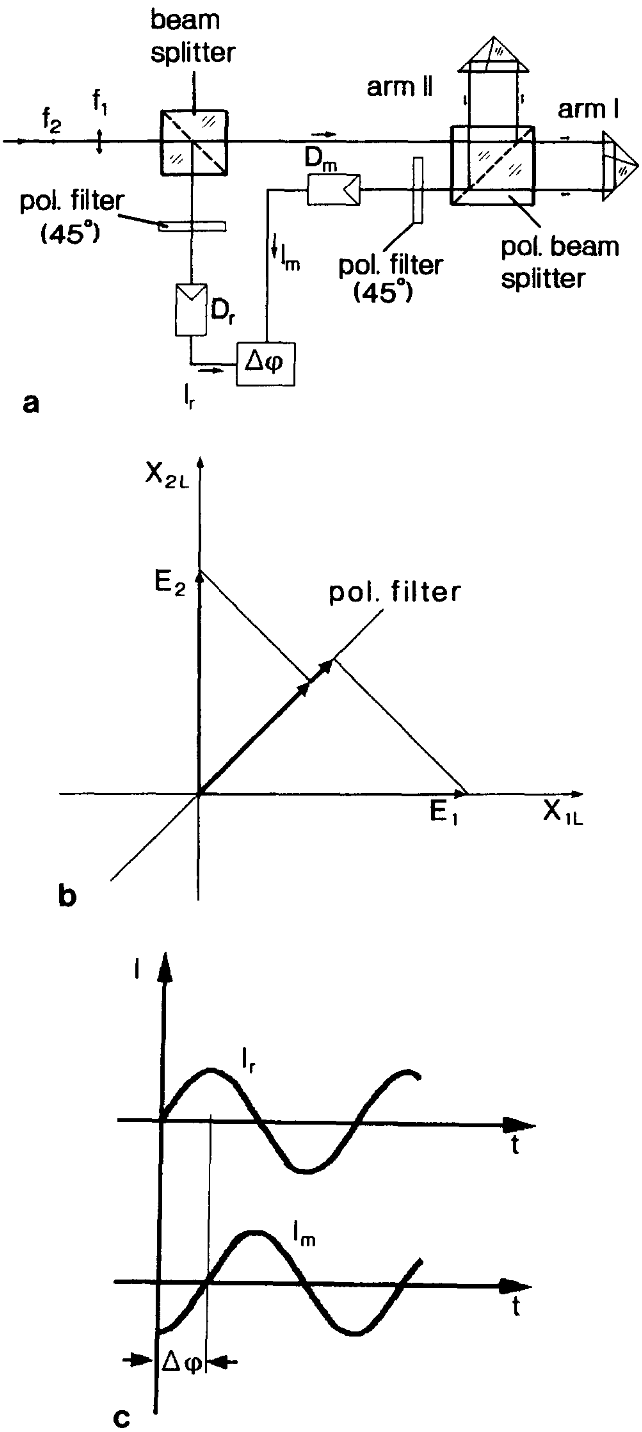


Figure 1 (*a*) Heterodyne interferometer, principle. (*b*) Components of the orthogonally linear polarized beam at the photodiode. (*c*) Phase difference $\Delta\varphi$ between measurement and reference signal

and are suppressed along the signal path.

$$I_r = I(f_1 - f_2) \sim I_0 \cos[2\pi(f_1 - f_2)t + \varphi_{0r}] \qquad (3)$$

$$I_0 = \tfrac{1}{2}E_{01}E_{02}$$

Here φ_{0r} is the initial, constant phase shift.

This signal serves as a reference. The polarizing beam splitter of the interferometer splits the incident beam up into directions I and II ("arms" of the interferometer), in compliance with the polarization. It fixes the polarization directions (x_1, x_2) in the interferometer. It will be assumed in the following that these directions are orthogonal. They are used as a reference. If the directions of the laser system x_{1L}, x_{2L} agree with those of the interferometer (given by the polarizing splitter), the frequencies are completely separated. The partial beams cover paths of different length l_1 and l_2, are recombined in the beam splitter, and detected by diode D_m. The photocurrent is obtained:

$$\begin{aligned} I'_m &\sim E_m^2 \\ &\sim \tfrac{1}{2}[E_{01}\sin(2\pi f_1 t + \varphi_{01} + \varphi_1) \\ &\quad + E_{02}\sin(2\pi f_2 t + \varphi_{02} + \varphi_2)]^2 \\ &= I(f_1) + I(f_2) + I(f_1 - f_2) + I(f_1 + f_2) \qquad (4) \end{aligned}$$

For the interferometer arrangement shown in *Figure 1a*,

$$\varphi_1 = \frac{4\pi n l_1}{\lambda_1}; \qquad \varphi_2 = \frac{4\pi n l_2}{\lambda_2}$$

where n is the refractive index; λ_1 and λ_2 are the wavelengths belonging to f_1 and f_2; φ_{01} and φ_{02} are initial, constant phases.

Here, too, the difference term is retained:

$$\begin{aligned} I_m &= I(f_1 - f_2) \\ &\sim I_0 \cos[2\pi(f_1 - f_2)t + \varphi_{01} - \varphi_{02} + \Delta\varphi] \qquad (5) \end{aligned}$$

$$\Delta\varphi = \varphi_1 - \varphi_2 \approx 4\pi(nl_1 - nl_2)/\lambda_m$$

λ_m is the mean wavelength.*

Compared with the (constant) signal I_r, I_m is affected by a phase shift $\Delta\varphi$, which varies with the path difference in the interferometer and which can be obtained by measuring the phase difference between I_r and I_m (*Figure 1c*). (If, for example, the resolution of the phase measurement is 1°, in the interferometer shown in *Figure 1* a length resolution of 0.9 nm can be attained for a displacement of the reflector.)

Several solutions for the phase measurement have been suggested and put into practice[1–4] some of which allow resolutions down to the range of 0.1 nm.[3,4]

Nonlinearities in the heterodyne interferometer

Equation (5) is exactly valid when only one frequency (and one polarization state) occurs in each interferometer arm. Due to various influences, mixed states are often found in both interferometer arms. This

* *For the sake of simplicity, a mean wavelength is introduced here. However, it should be noted that for interferometer systems with large frequency differences instead of the mean value, the wavelength of the interferometer arm in which the displacement occurs must be taken.*

results in a nonlinear relation between the measured phase difference and the respective displacement.[2,4–8] Four of the most frequent causes are investigated in the following:

- mixing due to nonorthogonality of the polarization directions of the incident radiation;
- mixing due to elliptic polarization of the incident radiation;
- mixing due to incomplete separation of the frequency components in a nonideal polarizing splitter; and
- amplitude variations in coherent transmission between interferometer arms.

Imperfect or incorrectly adjusted polarizing optics behind the beam splitter (e.g., retardation plates) or optics that show polarizing behavior (e.g., reflectors) do not cause first-order nonlinearity. For the following discussion it is assumed that the polarization axes in the interferometer are orthogonal.

Mixing due to nonorthogonality of the polarization directions of the incident radiation

It is assumed that the two polarization directions deviate from orthogonality by α (*Figure 2*). As a result, in addition to frequency f_1, a part of frequency f_2 is also found in arm *I* of the interferometer. The following relationships hold for the partial beams in the two interferometer arms (i.e., for the two polarization directions x_1, x_2):

$$E_{x1} = E_{01}\sin(2\pi f_1 t + \varphi_{01} + \varphi_1) + E_{02}\sin\alpha\sin(2\pi f_2 t + \varphi_{02} + \varphi_1) \qquad (6)$$

$$E_{x2} = E_{02}\cos\alpha\sin(2\pi f_2 t + \varphi_{02} + \varphi_2)$$

The following relation then results for the AC signal of diode D_m:

$$I_m \sim I_0 A^* \cos[2(f_1 - f_2)t + \varphi_{0m} + \Delta\varphi - \gamma] \qquad (7)$$

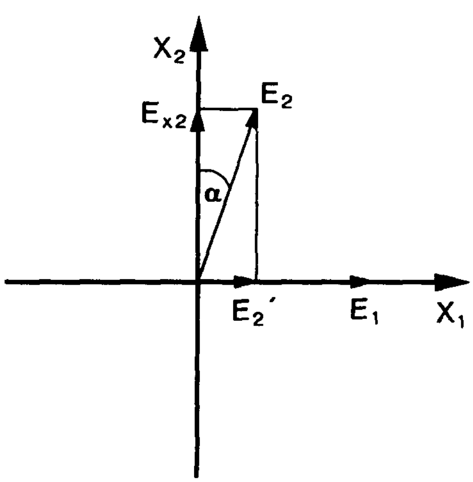


Figure 2 Beam components in the orthogonal reference system of the interferometer (x_1, x_2); deviation from orthogonality

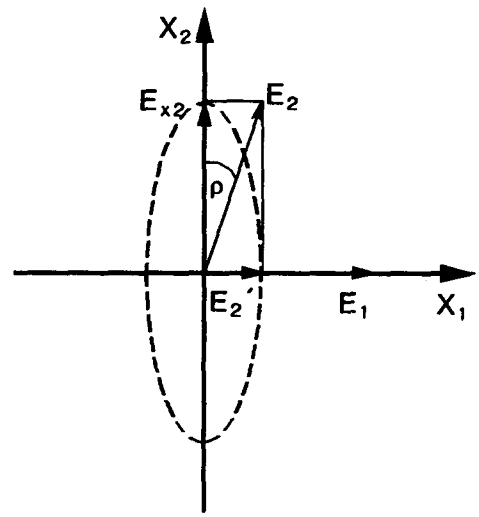


Figure 3 Beam components in the orthogonal reference system of the interferometer (x_1, x_2); ellipticity in one direction

$$\gamma = tg^{-1}\frac{\sin\alpha\sin\Delta\varphi}{\cos\alpha + \sin\alpha\cos\Delta\varphi} \qquad (8)$$

$$A^* = \sqrt{1 + \sin 2\alpha\cos\Delta\varphi} \qquad (9)$$

To simplify matters, the following is assumed:

$$\alpha \ll 1 \; (\alpha < 5°)$$

The following is then valid in first approximation:

$$\gamma = \alpha\sin\Delta\varphi \qquad (10)$$

Comparison between Equations (5) and (7) shows that the phase measurement is affected by a periodic systematic error, γ. The maximum error γ_{max} is about 2α. (If α is 5°, for example, the maximum error of phase measurement will be 10°. This corresponds to an error of 9 nm in the length measurement with a corner cube interferometer as shown in *Figure 1a*.) It should also be noted that the amplitude of the measured signal I_m is no longer constant (Equation [7]) but modulated by α and $\Delta\varphi$. When there is frequency mixing in both arms, the phase errors behave additively. Therefore, the simple rotation of the polarization directions of the laser beam relative to those of the interferometer theoretically do not cause a first-order phase error.[8]

Mixing due to elliptic polarization of the incident radiation

Frequency mixing effects also occur when the incident radiation is not ideally linear but elliptically polarized. On the assumption that component E_2 is elliptically polarized and the ratio of the ellipse axes described by the angle ρ (*Figure 3*), the following relations are valid for the partial beams in both interferometer arms:

$$E_{x1} = E_{01}\sin(2\pi f_1 t + \varphi_{01} + \varphi_1) + E_{02}\sin\rho\cos(2\pi f_2 t + \varphi_{02} + \varphi_1) \qquad (11)$$

$$E_{x2} = E_{02}\cos\rho\sin(2\pi f_2 t + \varphi_{02} + \varphi_2)$$

The AC component of the signal formed in photodiode D_m then is as follows:

$$I_m \sim I_0 U^* \cos[2\pi(f_1 - f_2)t + (\varphi_{01} - \varphi_{02}) + (\varphi_1 - \varphi_2) - \eta] \quad (12)$$

$$\eta = tg^{-1} \frac{\sin\rho \cos\Delta\varphi}{\cos\rho - \sin\rho \sin\Delta\varphi} \quad (13)$$

$$U^* = \sqrt{1 - \sin 2\rho \sin\Delta\varphi} \quad (14)$$

η and U^* correspond to γ and A^* for $\rho = \alpha$ in Equation (7). In case both partial beams are elliptically polarized but orthogonal (i.e., opposite ellipticity and equal amplitudes), no first-order phase error occurs.[8]

Mixing due to incomplete separation of the frequency components in a nonideal polarizing splitter

On the assumption that the fraction a of component E_2 reaches arm I, the following is valid for the splitting of the components:

$$\text{Arm I:} \quad E_{x1} + E_2' = E_{01}\sin(2\pi f_1 t + \varphi_{01} + \varphi_1) + E_{02} a \sin(2\pi f_2 t + \varphi_{02} + \varphi_1)$$

$$\text{Arm II:} \quad E_{x2} = E_{02}(1 - a)\sin(2\pi f_2 t + \varphi_{02} + \varphi_2) \quad (15)$$

In this case, both states of polarization are found in arm I.

After being reflected, the beams pass the splitter once more and only the fraction a of the "wrong" component finally reaches the photodiode, whereas the fraction $(1 - a)$ leaves the splitter at the side. Similarly, only the fraction $(1 - a)$ of the already weakened component E_2 finally reaches the photodiode. Therefore, the effective components in the interferometer arms are described as follows:

$$\text{Arm I}': \quad E_{x1} + aE_2' = E_{01}\sin(2\pi f_1 t + \varphi_{01} + \varphi_1) + E_{02} a^2 \sin(2\pi f_2 t + \varphi_{02} + \varphi_1)$$

$$\text{Arm II}': \quad E_{x2} = E_{02}(1 - a)^2 \sin(2\pi f_2 t + \varphi_{02} + \varphi_2) \quad (16)$$

The following expression is obtained for the AC component produced in diode D_m:

$$I_m \sim I_0 K^* \cos[2\pi(f_1 - f_2)t + \varphi_{0m} + \Delta\varphi - \zeta] \quad (17)$$

$$\zeta = tg^{-1} \frac{a^2 \sin\Delta\varphi}{(1 - a)^2 + a^2 \cos\Delta\varphi} \quad (18)$$

$$K^* = \sqrt{1 + 2a^2(1 - a)^2 \cos\Delta\varphi} \quad (19)$$

For $a^2 \ll 1$ Equation (19) can be simplified:

$$\zeta \approx a^2 \sin\Delta\varphi \quad (20)$$

(Polarizing beam splitters usually have extinction ratios of >300:1 [$a^2 < 3 \times 10^{-3}$] for the inferior polarization direction, whereas the other one often is nearly perfect. The resulting phase error leads to an error in the length measurement of 0.17 nm for corner cube interferometers as shown in *Figure 1a*).

Influence of amplitude variation in coherent transmission

The relative amplitude variation in coherent transmission in an interferometer arm due to the displacement of the measuring reflector may be another important cause of phase errors when there is already incorrect frequency mixing.[8]

In the following it is assumed that in addition to an existing orthogonality error α the amplitude in one interferometer arm is attenuated. This attenuation is described by the amplitude factor k:

$$E_{x1} = E_{01}\sin(2\pi f_1 t + \varphi_{01} + \varphi_1) + E_{02}\sin\alpha\cos(2\pi f_2 t + \varphi_{02} + \varphi_1) \quad (21)$$

$$E_{x2} = kE_{02}\cos\alpha\sin(2\pi f_2 t + \varphi_{02} + \varphi_2)$$

For the AC component of the signal formed in diode D_m, the following expression is obtained:

$$I_m \sim I_0 A^{**} \cos[2\pi(f_1 - f_2)t + \varphi_{0m} + \Delta\varphi - \gamma^*] \quad (22)$$

$$\gamma^* = tg^{-1} \frac{\sin\alpha \sin\Delta\varphi}{k\cos\alpha + \sin\alpha\cos\Delta\varphi} \quad (23)$$

$$A^{**} = \sqrt{k^2\cos^2\alpha + \sin^2\alpha + k\sin 2\alpha\cos\Delta\varphi} \quad (24)$$

For the sake of simplicity, the following is assumed:

$$\alpha \ll 1\ (\alpha < 5^\circ), \qquad 1 > k > 0.5$$

The following is then valid in first approximation:

$$\gamma^* = \frac{\alpha\sin\Delta\varphi}{k} \quad (25)$$

In the case of elliptic polarization α can be replaced by the elliptic polarization angle, ρ, without changing the result.

Unequal amplitude distribution in the arms of the interferometer increases the influence of nonorthogonality on nonlinearity.

In case the amplitude attenuation varies with the displacement of the reflector (e.g., due to varying overlapping of the measurement and reference beam) the nonlinearity depends on the actual displacement.

Detection of phase errors

To investigate the nonlinearity of interferometers, it is desirable to obtain signals that allow conclusions to be drawn concerning the prevailing phase errors.

The heterodyne interferometer shown in *Figure 4a* has been extended to comprise two signal outputs. Superposition of the individual components from the two interferometer arms is not only detected in one polarization direction (as is usually the case behind the polarizing filter), but also in the direction orthogonal to it. The conditions on the diodes are represented in *Figure 4b* for the case of nonorthogonality (see *Figure 2*). This description also holds for the case in which the incident beam is elliptic. The necessary splitting was achieved by using a $\lambda/2$ plate

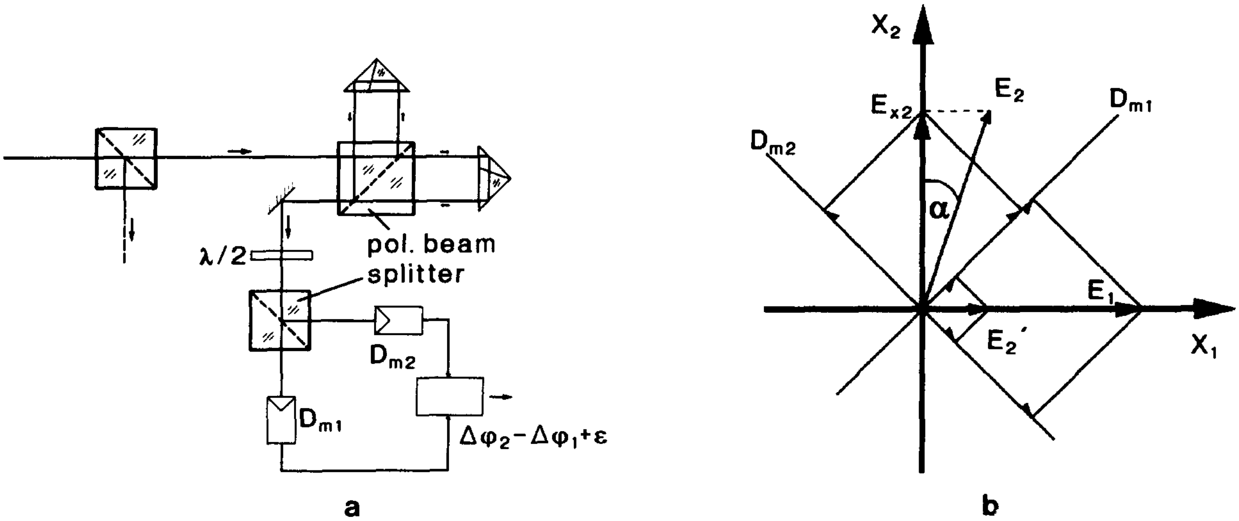


Figure 4 Heterodyne interferometer with detection of nonlinearity. (*a*) Principle. (*b*) Beam components at the photodiode (at the presence of nonorthogonality)

(rotation of the polarization direction by 45°) in combination with another polarizing splitter. One of the two output signals (here assigned to diode D_{m1}) corresponds to that described by Equation (7):

$$I_{m1} \sim I_0 A^* \cos[2\pi(f_1 - f_2)t + \Delta\varphi_1] \quad (7a)$$

$$\Delta\varphi_1 = \varphi_{0m} + \Delta\varphi - \gamma$$

$$\gamma = tg^{-1}\frac{\sin\alpha\sin\Delta\varphi}{\cos\alpha + \sin\alpha\cos\Delta\varphi} \quad (8)$$

$$A^* = \sqrt{1 + \sin 2\alpha\cos\Delta\varphi} \quad (9)$$

The following relation is obtained for the other signal (I_{m2}):

$$I_{m2} \sim I_0 B^* \cos[2\pi(f_1 - f_2)t + \Delta\varphi_2] \quad (26)$$

$$\Delta\varphi_2 = \varphi_{0m} + \Delta\varphi + \gamma' + \pi$$

$$\gamma' = tg^{-1}\frac{\sin\alpha\sin\Delta\varphi}{\cos\alpha - \sin\alpha\cos\Delta\varphi} \quad (27)$$

$$B^* = \sqrt{1 - \sin 2\alpha\cos\Delta\varphi} \quad (28)$$

The phase difference of the two signals (Equations [7a] and [26]) is described by the following expression:

$$\Delta\varphi_2 - \Delta\varphi_1 = \gamma' + \gamma + \pi + \varepsilon \quad (29)$$

where ε is an additional constant phase shift that is due to the electro-optical arrangement and the phase measurement. It is constant and does not affect displacement measurements. With the aid of Equations (8) and (27) the phase error can be described as follows:

$$\gamma + \gamma' = \sin^{-1}\frac{\sin 2\alpha\sin\Delta\varphi}{\sqrt{1 - \sin^2 2\alpha\cos^2\Delta\varphi}} \quad (30)$$

To simplify matters, the following is assumed:

$$\alpha \ll 1\ (\alpha < 5°)$$

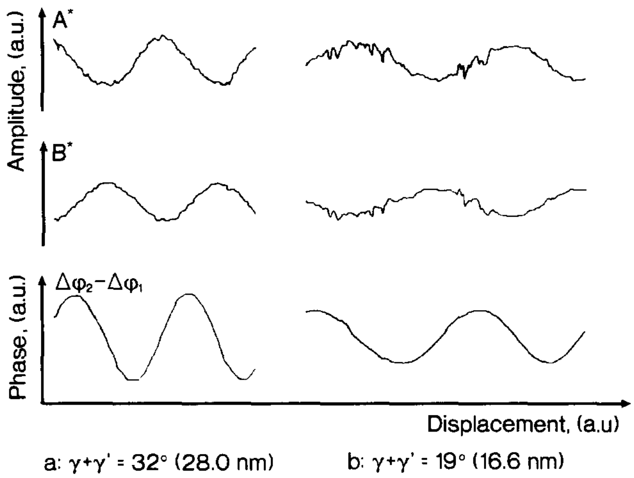
The following is then valid in first approximation:

$$2\Delta = \gamma + \gamma' \approx 2\alpha\sin\Delta\varphi \quad (31)$$

Δ corresponds to the phase error due to nonorthogonality. The maximum error is approximately 2α. It is a direct measure of the interferometer's nonlinearity.

The measurement of the phase difference of the two output signals may serve to detect the causes of nonlinearity and eliminate them if possible. Phase errors due to incomplete separation of the polarization components cannot be detected with this arrangement. This contribution can be determined by separately measuring the extinction ratio and by using Equation (17). This amount is usually small. If the cause is known (nonorthogonality of the polarization components of the laser beam, ellipticity, etc.), quantitative statements about its parameters also can be made. (The test of an interferometer with a two-mode laser, for example, furnished a maximum value of the phase difference [Δ_{max}] of about 15°. Additional investigations revealed elliptic polarization, which was eliminated by using a $\lambda/4$ and a $\lambda/2$ plate. The adjustment could at once be checked on the basis of the result of the phase measurement. The remaining phase error was 0.5°, the corresponding nonlinearity in the displacement measurement 0.4 nm.) *Figure 5* shows the results of phase and amplitude measurements of two different interferometers. It is clearly seen that the signal-to-noise ratio of the phase measurement is greater than that of the amplitude measurements.

A means is thus available that allows essential components of the phase error, and consequently the nonlinearity of an interferometer, to be measured without comparison with a reference interferometer, the particular advantage being that a phase measurement forms the basis and not amplitude measurements, which are less accurate.



$\gamma+\gamma'$ — detected non-linearity
A^* — amplitude of signal I_{m1}
B^* — amplitude of signal I_{m2}
$\Delta\varphi_2-\Delta\varphi_1$ — phase difference of signals I_{m1} and I_{m2}

Figure 5 Detection of nonlinearities of two different interferometers

Compensation of phase errors

Elimination of phase errors by means of adjustment is not always possible. Examples of this are the nonorthogonality of the polarization direction of laser radiation or the influence of a glass–fiber connection between laser and interferometer (mainly, a mixture of ellipticity and nonorthogonality). These errors can be compensated for by using the measurement and evaluation method suggested in the following.

For the sake of simplicity the phase error due to nonorthogonality is chosen to be the only one. However, the method works equally in the case of ellipticity, amplitude variation, and mixtures of these errors. The phase error due to incomplete separation in the polarizing beam splitter is not compensated.

As shown in *Figure 6*, the differences $\Delta\varphi_1$ and $\Delta\varphi_2$ of the reference signal and the phases of the two output signals are measured. The arithmetical mean value of these two phase values is described by the following:

$$\Sigma = \tfrac{1}{2}(\Delta\varphi_1 + \Delta\varphi_2 + \vartheta) = \tfrac{1}{2}(\Delta\varphi - \gamma + \Delta\varphi + \gamma' + \vartheta)$$
$$= \Delta\varphi + \tfrac{1}{2}(\gamma' - \gamma + \vartheta) \qquad (32)$$

ϑ is an additional, constant phase shift that is due to the electro-optical arrangement and the phase measurement. It is constant and does not affect displacement measurements.

From Equations (8) and (27) follows:

$$\gamma - \gamma' = \sin^{-1}\frac{\sin^2\alpha\,\sin 2\Delta\varphi}{\sqrt{1-\sin^2 2\alpha\,\cos^2\Delta\varphi}} \qquad (33)$$

The mean value Σ contains the phase $\Delta\varphi$, which is proportional to the true displacement in the interferometer, the residual error $(\gamma' - \gamma)$, and a constant phase shift, ϑ.

In *Figure 7*, the maximum residual error is shown as a function of the maximum uncompensated error. The measuring method described here opens up the possibility of compensating first-order phase errors "on-line." This measuring method allows a considerable improvement in linearity to be expected.

Experiments

The arrangement of *Figure 8* is used to verify the detection and compensation method. The cube corner of the interferometer to be examined is fastened near the pivot of a lever actuated by a piezoelectric element. The long end of the arm carries the cube corner of a second interferometer. The zero crossings of the phase output of this interferometer generate equidistant

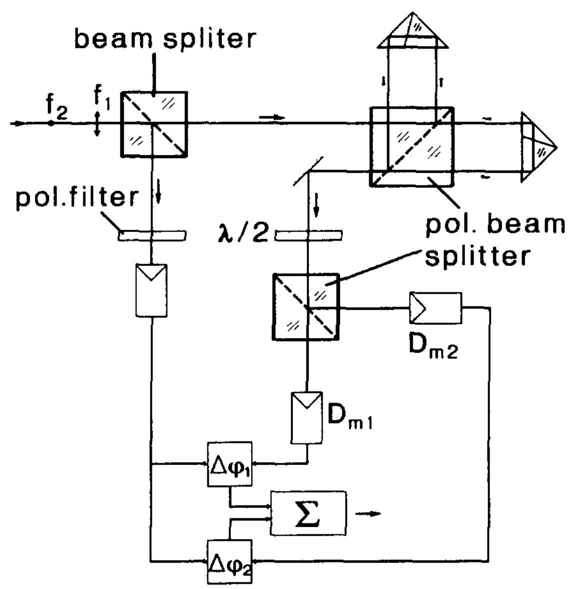


Figure 6 Heterodyne interferometer with compensation of nonlinearity

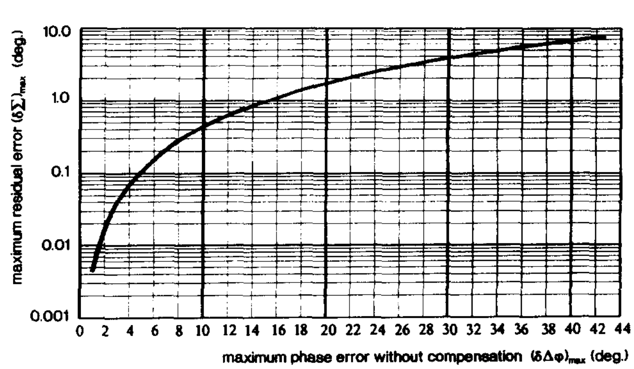


Figure 7 Calculated maximum residual error $(\delta\Sigma)_{max}$ after compensation versus maximum phase error without compensation $(\delta\Delta\varphi)_{max}$

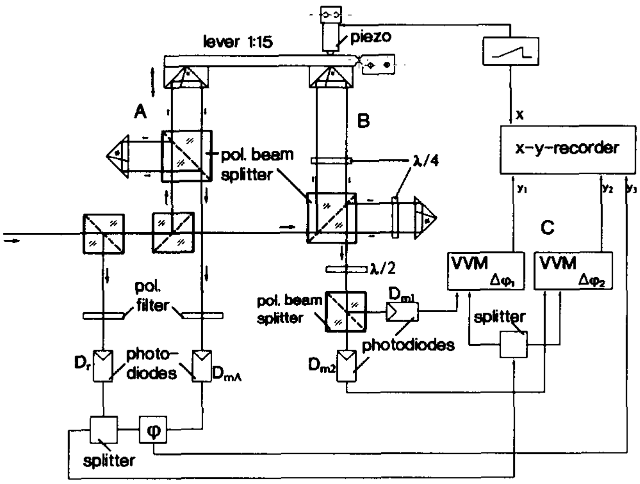
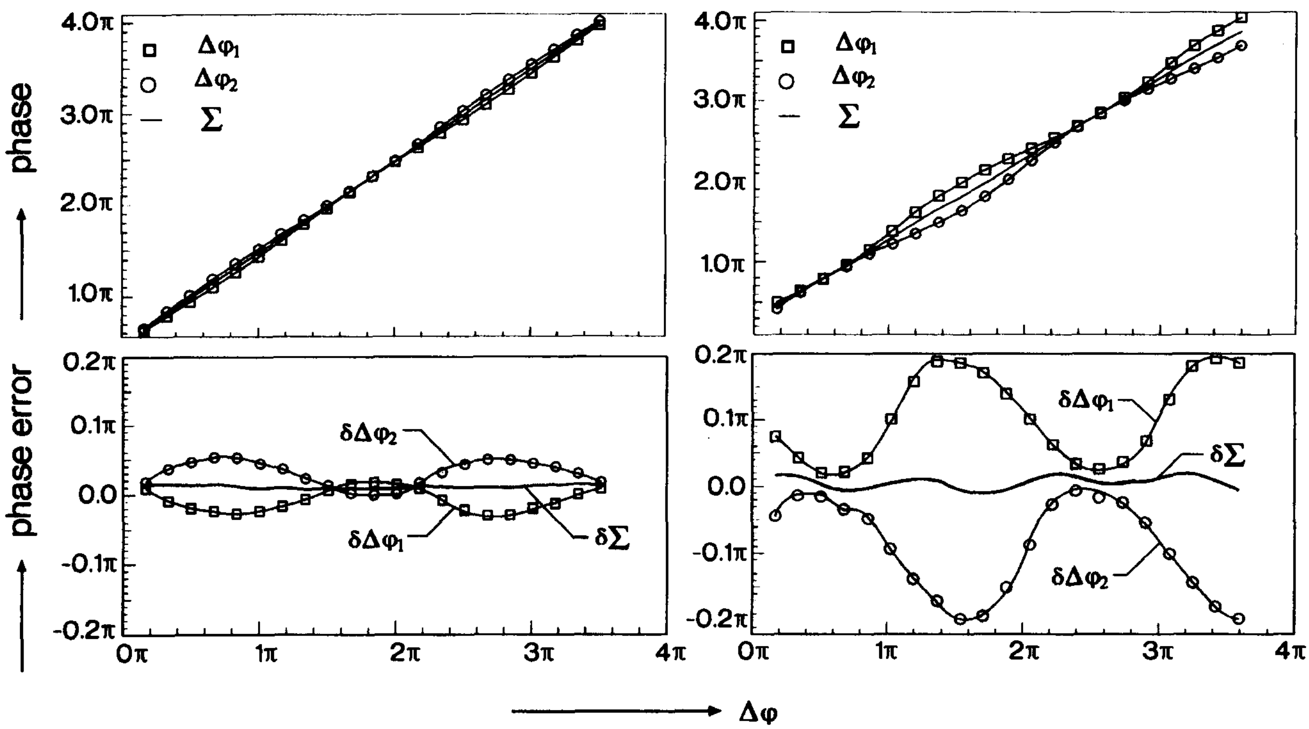


Figure 8 Experimental setup for the detection and compensation of nonlinearities. A, Reference interferometer; zero crossings of the difference phase are taken as equidistant markings. B, Interferometer with complete phase measurement under inspection. C, Evaluation (VVM, vector voltmeter)

subdivisions for the displacement in the interferometer to be investigated. According to Equations 8, 13, and 18, the phase error is periodic with $\Delta\varphi$. Thus, the zero crossings are not influenced by potential nonlinearities. The phase cycle of the interferometer to be investigated is provided with 15 equidistant marks, and the phase shifts $\Delta\varphi_1$ and $\Delta\varphi_2$ of the interferometer in question are measured at these marks.

Figure 9 shows typical measurement results of two different interferometer setups. The measured phase values, mean values, and linearity errors have been plotted as a function of the true displacement. Both interferometers were noncommercial "typical corner cube interferometers." The very large nonlinearity of interferometer b is due to ellipticity introduced by a glass–fiber link between laser and interferometer. It should be noted that well-known commercial interferometers show nonlinearities normally well below 10° phase. The measurement results show the following:

- the traces of the two phase outputs $\Delta\varphi_1$ and $\Delta\varphi_2$ show sinusoidal deviations from linearity (according to Equations [7a] and [26]);



$(\delta\Delta\varphi_1)_{max}$ = 8.5° (7.7nm)
$(\delta\Delta\varphi_2)_{max}$ = 9.6° (8.6nm)
$(\delta\Sigma)_{max}$ = 1.3° (1.2nm)

a

$(\Delta\varphi_1)_{max}$ = 31.2° (27.1nm)
$(\Delta\varphi_2)_{max}$ = 34.0° (29.6nm)
$(\delta\Sigma)_{max}$ = 5.1° (4.4nm)

b

Figure 9 The compensation of phase errors of two different interferometer setups (*a* and *b*); phase differences $\Delta\varphi_1$, $\Delta\varphi_2$ and the arithmetic mean value Σ, and phase errors (nonlinearity) $\delta\Delta\varphi_1$, $\delta\Delta\varphi_2$ and arithmetic mean value $\delta\Sigma$ versus the true displacement $\Delta\varphi$; corresponding length measurement errors

- the deviations from linearity are opposite in phase (according to Equations [7a] and [26]) (i.e., the measurement of the phase difference leads to the sum of the deviations, according to Equation [30]);
- the residual nonlinearity $\delta\Sigma$ decreases with decreasing phase errors (according to *Figure 7*);
- the residual nonlinearity $\delta\Sigma$ shows a doubly periodic dependence on the displacement $\Delta\varphi$ (according to Equation [33]);
- the calculated values for the residual nonlinearities both are approximately 1° smaller than the measured ones (0.4° and 4.1° compared with 1.3° and 5.1°, respectively); this error contribution, which is independent of the measured nonlinearity, can completely be attributed to the combined phase measurement error of the phase-measuring devices, which was found to be 1°.

Thus, the conclusion can be drawn that the proposed detection and compensation methods are effective. The limits of the compensation method are given by an inherent systematic residual error and the phase-measuring errors of the phase-measurement devices.

Conclusions

In heterodyne interferometers, nonlinearities in the relation between the output phase signal and the displacement result from the incorrect mixing of frequency components. The investigation of various causes of this mixing has shown that the strongest influence is exerted by the properties of the laser beams used (nonorthogonality of the linear polarization directions, ellipticity). The nonlinearity is periodic, with a complete phase cycle of the output signal. Zero passages of the phase ("interference fringes") are exactly equidistant. The measurement of the phase positions in two orthogonal directions of the output beam provides almost full information on linearity errors. The measurement of the difference of these phases gives information about the quantity of the nonlinearity and can be used to investigate interferometer setups, whereas the absolute measurement (in difference to the reference signal) may serve to compensate the nonlinearity. Experiments on interferometers with different nonlinearities revealed that both the check by means of the difference measurement and the compensation with the aid of the absolute measurement are effective and in good agreement with the theory.

There are several possible arrangements for the detection of the signals in two orthogonal planes (according to *Figure 4*). Either the polarization directions are rotated by means of a $\lambda/2$ plate (*Figures 4a* and *6*), polarizing beam splitter and diodes are rotated by 45° instead, or a normal beam splitter and two polarizing filters in front of the photodiodes are taken.

For the purpose of the detection of nonlinearities, such arrangements can be placed in the output beam of interferometers in the form of separate instruments. Therefore, the detection method even is suitable to check commercial laser interferometer systems. Determination of the occurring maximum nonlinearities requires the measuring reflector to be displaced by at least one interference fringe ($\lambda/2$, in special cases $\lambda/4$). In most cases only small phase errors will occur, and the phase difference measurement can be carried through in a simple way (e.g., by means of a "double balanced mixer").

The compensation method in general cannot be achieved with a separate instrument, as in this case the internal reference for absolute phase measurements must be made use of. For modular commercial multiaxes instruments, instead of the diodes shown in *Figure 8*, the application of receivers of two axes and the calculation of the mean value of the counts of both axes are conceivable.

Acknowledgments

The authors are grateful to X. Zhao for his help with the computation, and F. Reinboth for his help with the experimental setup. Also, the generous support of the Alexander von Humboldt-Stiftung is gratefully acknowledged.

References

1 Sommargren, G. E. "A new measurement system for precision metrology," *Prec Eng* 1987, **9**, 179–184
2 Quenelle, R. C. and Wuerz, L. J. "A new micrometer-controlled laser dimensional measurement and analysis system," *Hewlett Packard Journal* 1983, **34**, 3–13
3 Dorenwendt, K. and Probst, R. "Hochauflösende Interferometrie mit Zweifrequenzlasern," *PTB-Mitt* 1980, **90**, 359–362
4 Reinboth, F. and Wilkening, G. "Optische Phasenschieber für Zweifrequenz-Laser-Interferometrie," *PTB-Mitt* 1983, **93**, 168–174
5 Bobroff, N. "Residual errors in laser interferometry from air turbulence and nonlinearity," *Appl Opt* 1987, **26**, 2676–2682
6 Sutton, C. M. "Non-linearity in length measurement using heterodyne laser Michelson interferometry," *J Phys E* 1987, **20**, 1290–1292
7 Steinmetz, C. R. "Sub-micron position measurement and control on precision machine tools with laser interferometry," *Prec Eng* 1990, **12**, 12–24
8 Rosenbluth, A. E. and Bobroff, N. "Optical sources of non-linearity in heterodyne interferometers," *Prec Eng* 1990, **12**, 7–11