Available online at www.sciencedirect.com
SCIENCE DIRECT®
Precision Engineering 27 (2003) 311–322
PRECISION ENGINEERING
www.elsevier.com/locate/precision

An investigation of two unexplored periodic error sources in differential-path interferometry

Tony L. Schmitz [a,*], John F. Beckwith [b,1]

[a] *Department of Mechanical and Aerospace Engineering, University of Florida, 237 MEB, Gainesville, FL 32611, USA*
[b] *Electronics Engineering Technologies Division, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA*

Received 12 March 2001; received in revised form 3 October 2002; accepted 26 February 2003

Abstract

This paper describes two sources of periodic error in differential-path interferometry that have remained largely unexplored: *dynamic periodic error* that is exhibited by heterodyne interferometer systems under high-speed displacements and *intermodulation periodic error* caused by amplifier nonlinearity. *Dynamic periodic error* occurs when the measurement signal, or intended *ac interference* term, and unwanted *dc interference* terms, that exist due to frequency leakage in physical implementations of heterodyne interferometers, are both present within the phase measuring electronics' modulation bandwidth. The situation is similar to the well-documented *pseudo-static periodic error* observed at low slide speeds, where the intended *ac interference*, leakage-induced *ac interference*, and *ac reference* terms all lie within the modulation bandwidth.

The Frequency–Path (F–P) model of the propagation of light from the source to detector for differential-path interferometers is also presented. This model identifies each possible path for each light frequency from the source to detector and predicts the number of interference terms that may be expected at the detector output. We show that, regardless of the interferometer configuration, the behavior of each interference term with respect to optical path changes may be grouped into one of four categories: *optical power*, *ac interference*, *ac reference*, and *dc interference*. The application of the F–P model to the generic description of periodic error in a single pass, Michelson-type heterodyne interferometer is provided.

Keywords: Periodic error; Heterodyne interferometry; Frequency–Path model

1. Introduction

Differential-path interferometry is used extensively in situations requiring accurate displacement measurements. Commercially available configurations for differential-path interferometers include, but are not limited to, single frequency homodyne and dual frequency heterodyne Michelson-type interferometers with single, double, or multiple passes of the optical paths. Typical applications for these measurement systems are lithographic stages for semiconductor fabrication, transducer calibration, and precision cutting and measuring machines.

Regardless of the optical configuration, all differential-path interferometers infer changes in displacement of a selected optical path by monitoring the optically-induced changes in a photodetector current. The phase-measuring electronics converts this photodetector current to displacement by an assumed relationship between corresponding changes in detector current and displacement, where this assumed relationship is defined by some idealized performance of the optical elements. In reality, the system optics do not behave perfectly and small deviations in the optical signal are induced. Furthermore, intermodulation distortions of the multiple interference terms in the nonideal optical signal may be produced during amplification of the photodetector current. Because the final displacement is calculated according to an assumed algorithm in the system electronics, the measurement signal does not exactly reproduce the target motion and errors are generated. In this paper, we are specifically interested in two previously unexplored periodic error sources: *dynamic periodic error* during high-speed displacements and *intermodulation periodic error* due to amplifier nonlinearities. Fundamental descriptions of these errors will be provided.

2. Background

In the literature, investigations of periodic error have generally focused on heterodyne Michelson-type interferometers.

* Corresponding author. Tel.: +1-352-392-8909; fax: +1-352-392-1071.
E-mail address: tschmitz@ufl.edu (T.L. Schmitz).
[1] Retired.

0141-6359/03/$ – see front matter © 2003 Elsevier Science Inc. All rights reserved.
doi:10.1016/S0141-6359(03)00036-9

In heterodyne systems, imperfect separation of the two light frequencies into the measurement and reference paths has been shown to produce first- and second-order periodic errors, or errors of one and two cycles per wavelength of optical path change, respectively. The two heterodyne frequencies are typically carried on collinear, mutually orthogonal, linearly polarized laser beams in a method referred to as polarization-coding; unwanted leakage of the reference frequency into the measurement path and vice versa may occur due to nonorthogonality between the ideally linear beam polarizations, elliptical polarization of the individual beams, imperfect optical components, parasitic reflections from individual optical surfaces, and/or mechanical misalignment between the interferometer elements (laser, polarizing optics, and targets). In a perfect system, a single wavelength would travel to a fixed reference target, while a second, single wavelength traveled to a moving target. Interference of the combined signals would yield a perfectly sinusoidal trace with phase that varied, relative to a reference phase signal, in response to motion of the moving target. However, the inherent frequency leakage in actual implementations produces an interference signal which is not purely sinusoidal (i.e. contains spurious spectral content) and leads to a periodic, or noncumulative, error in the measured displacement.

Hirst [1], Fedotova [2], Quenelle [3], and Sutton [4] performed early investigations of periodic error in heterodyne Michelson interferometers. Subsequent publications identified and described these periodic errors and built on the previous work [5–29]. Specific areas of research have included efforts to measure periodic error under various conditions (e.g. [5–8]), frequency domain analyses [9–11], analytic modeling techniques [12–15], Jones calculus modeling methods [8,16], and reduction of periodic errors (e.g. [9,17,29]). In general, previous publications have relied on complex mathematical models, including the polarization state of each electric field during interaction with individual optical elements, to describe the sources of periodic error. Although technically sound, these analyses tend to complicate the simplicity of the physical reality. As an alternative, we present here a generic description of periodic error in differential-path interferometry. We term this the Frequency–Path model and offer its description in the following sections.

3. Frequency–Path model

The Frequency–Path (F–P) model is comprised of three fundamental components: the light source composed of one or more frequencies, the detector, which serves as a bridge between the interferometer optics and electronics, and the collection of all available paths between the source and detector consistent with the propagation laws of physical optics. In an ideal interferometer, each source frequency could reach the detector by the intended path only and no frequency leakage would occur between paths. However, due to deviations from the ideal, each frequency may in fact reach the detector by other available paths and contribute to the total electric field. The sum of these individual electric field contributions, or Frequency–Path elements, is then ideally squared at the photodetector, in the absence of saturation, to produce the output current.

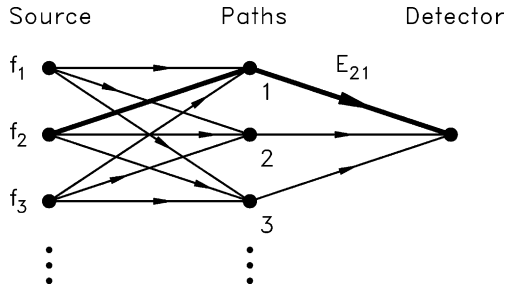

Fig. 1. General Frequency–Path model—some amount of light (possibly zero) from each source frequency (solid circles in left column) can reach the detector (solid circle on right) by each available path (solid circles in middle column). Each Frequency–Path element is identified by the continuous line connecting a particular source frequency to the detector by one of the available paths.

A schematic representation of the general F–P model is shown in Fig. 1. In the figure, the number of source frequencies is represented by an appropriate number of solid circles in the left column. Similarly, the number of paths from the source to detector, including intentional or unintentional multiple passes, is represented by solid circles in the middle column, while the detector is shown on the right. Each continuous line that connects a given source frequency to the detector through one of the available paths identifies an F–P element (note that these lines do not represent actual optical paths, but rather the electric field contribution from light of a particular frequency, identified in the left column, reaching the detector via a path identified in the middle column). Furthermore, each of these F–P elements corresponds to an actual light beam in the physical system traveling independently of any other from the source to the detector where the interference takes place.

In the F–P model notation, the electric field amplitude of each F–P element is designated by the letter E appended with two subscripts. The first subscript denotes the source frequency, while the second lists the path from the source to detector. For example, E_{21}, displayed as a heavy solid line in Fig. 1, gives the electric field amplitude of frequency 2 that reaches the detector via path 1.

3.1. Interference terms

The squaring action of the photodetector creates $n(n+1)/2$ distinct interference terms from the sum of n active F–P elements. In a fully leaking interferometer, for example, n is equal to the product of the number of frequencies and paths. Each of the interference terms is a product of two F–P elements and it is the superposition of these multiple interference terms in the photodetector current that is converted to displacement by the system electronics. This would appear

to generate a complex picture of the photodetector current frequency content. However, these interference terms may be grouped by optical path change dependency into only four categories.

3.1.1. Equal frequency subscripts/equal path subscripts

If the frequency and path subscripts of the two F–P elements of an interference term are both equal, then it corresponds to energy from the same source frequency reaching the detector via the same path, e.g. $E_{21}E_{21}$. This type of interference term is called an *optical power* term and contributes a constant intensity to the photodetector current, independent of optical path changes. There are n terms of this type, one corresponding to each active F–P element. A particular *optical power* term may in principle, and sometimes in practice, be isolated in the detector by blocking all paths and extinguishing all frequencies, except the frequency and path of interest.

3.1.2. Unequal frequency subscripts/equal path subscripts

This type of interference term corresponds to energy from different source frequencies reaching the detector by the same path, e.g. $E_{21}E_{11}$. It produces a time-harmonic contribution to the photodetector current with a frequency equal to the difference in the optical frequencies of the mode pair involved. Terms of this type are called *ac reference* terms with phase that varies by one full cycle over the synthetic wavelength, or the distance defined by the difference in wave numbers (λ^{-1}) between the source frequencies in question, i.e. ${\lambda_2}^{-1} - {\lambda_1}^{-1}$ for $E_{21}E_{11}$. The *ac reference* terms can be isolated in the detector current by blocking all paths but the one of interest. They may then be observed at the frequency $f_1 - f_2$, in this case, using a spectrum analyzer or scope.

3.1.3. Equal frequency subscripts/unequal path subscripts

A *dc interference* term is produced when energy from one source frequency reaches the detector by two different paths, e.g. $E_{21}E_{22}$. When the optical path is not changing, *dc interference* terms contribute a dc level to the detector current. This contribution becomes harmonic in time, due to the Doppler effect, when the optical path changes at a constant speed. A specific *dc interference* term may be observed with a spectrum analyzer by first suppressing all other source frequencies and then noting the signal that is Doppler-shifted from zero frequency during constant velocity target motion. There is also a spatial interference between every pair of *dc interference* terms, which is again periodic over a distance corresponding to one synthetic wavelength of the two source frequencies involved.

3.1.4. Unequal frequency subscripts/unequal path subscripts

This type of interference term is established when energy from one source frequency reaches the detector by one path, while energy from a second frequency reaches the detector via another path, e.g. $E_{21}E_{12}$. When the target is at rest, *ac interference* terms produce a time-harmonic variation in the detector current at the difference frequency (alternately referred to as the split, reference, or beat frequency) of the modes involved. During target motion, the *ac interference* term is Doppler-shifted either up or down from the beat frequency, depending on the direction of travel, thus enabling their visual isolation with the aid of a spectrum analyzer.

The fundamental concept of the F–P model is that a formal treatment of the vector character and polarization state of the individual electric fields, as well as the nature and degree of imperfections in the interferometer optics and system misalignments, is not necessary. The reader will note that they were ignored in the preceding interference term descriptions. These items have no effect on the way the detector current varies with changes in the optical path lengths; they only serve to partially determine the electric field coefficients of the individual F–P elements. All interferometers with the same nonzero F–P elements have comparable terms present in the photodetector current and, additionally, the variation with changes in optical path of these terms is one of only four types.

The situation is analogous to defining a class of functions by specifying only the frequencies, but not the amplitudes present in their Fourier expansions. Knowledge of these frequencies alone allows the discussion of many common properties of all functions in the class but the exact values of any individual function requires the additional knowledge of its expansion coefficients. In this case, the expressions for these coefficients are complex functions of the polarization of the light as it leaves the source and of the transmission, reflection, and polarization-altering properties of the optical elements along the path to the detector which, in general, are not known for a particular interferometer configuration and can change from one setup to another.

3.2. F–P model example

As an example of the F–P model, consider a heterodyne, Michelson-type, single pass interferometer. In the ideal case, there are two source frequencies, f_1 and f_2, and only one possible path for each frequency from the source to the detector, i.e. two F–P elements. One F–P element represents the propagation of the first frequency along its intended path to the reference target, while the other F–P element represents the propagation of the second frequency along its intended path to the moving target. The two F–P elements give a total of $(2)(2 + 1)/2 = 3$ interference terms: two *optical power* terms and the desired *ac interference* term. In reality, there are two source frequencies and two possible paths for each frequency from the source to detector (assuming multiple extraneous reflections are neglected) due to imperfect frequency separation at the polarizing beam splitter. In other words, both frequencies can travel to either the reference or moving target in a fully leaking system. This gives four F–P elements and 10 distinct interference terms. Fig. 2 shows: (a) the F–P model for a heterodyne, Michelson-type, single pass

interferometer, (b) the actual optical configuration for this interferometer, and (c) the four F–P elements as identified from the actual optical configuration.

Path 1, which ideally contains light of frequency f_1 (expressed in Hz) only (F–P element E_{11}) propagates two signals due to frequency leakage:

$$E_{11}\cos(\omega_1 t - k_1 x_1 + \phi_{11}) \quad \text{and}$$
$$E_{21}\cos(\omega_2 t - k_2 x_1 + \phi_{21}),$$

where the φ_{ij} are the initial phases of the corresponding signals (E_{ij}), x_1 represents motion of the target in path 1, k_1 and k_2 are the propagation constants (equal to $2\pi/\lambda_1$ and $2\pi/\lambda_2$, respectively), and $\omega_{1,2} = 2\pi f_{1,2}$ (rad/s). Similarly, path 2, ideally composed of frequency f_2 light only (F–P element E_{22}), also contains two signals:

$$E_{22}\cos(\omega_2 t - k_2 x_2 + \phi_{22}) \quad \text{and}$$
$$E_{12}\cos(\omega_1 t - k_1 x_2 + \phi_{12}),$$

where the parameter definitions are analogous. The photodetector current is obtained by squaring the sum of the four F–P element electric fields, consisting of two intended and two leakage-induced signals, as shown in Eq. (1):

$$\begin{aligned}
I \propto E^2 &= (E_{11}\cos(\omega_1 t - k_1 x_1 + \phi_{11}) + E_{21}\cos \\
&\times(\omega_2 t - k_2 x_1 + \phi_{21}) + E_{22}\cos(\omega_2 t - k_2 x_2 + \phi_{22}) \\
&+ E_{12}\cos(\omega_1 t - k_1 x_2 + \phi_{12}))^2 \\
&= \tfrac{1}{2}E_{11}^2[\cos(2\omega_1 t - 2k_1 x_1 + 2\phi_{11}) + 1] + \tfrac{1}{2}E_{21}^2 \\
&\times[\cos(2\omega_2 t - 2k_2 x_1 + 2\phi_{21}) + 1] + \tfrac{1}{2}E_{22}^2 \\
&\times[\cos(2\omega_2 t - 2k_2 x_2 + 2\phi_{22} + 1)] + \tfrac{1}{2}E_{12}^2 \\
&\times[\cos(2\omega_1 t - 2k_1 x_2 + 2\phi_{12}) + 1] + E_{22}E_{11} \\
&\times[\cos(\omega_2 t + \omega_1 t - k_2 x_2 - k_1 x_1 + \phi_{22} + \phi_{11}) \\
&+ \cos(\Delta\omega t - k_2 x_2 + k_1 x_1 + \phi_{22} - \phi_{11})] + E_{21}E_{12} \\
&\times[\cos(\omega_2 t + \omega_1 t - k_1 x_2 - k_2 x_1 + \phi_{21} + \phi_{12}) \\
&+ \cos(\Delta\omega t + k_1 x_2 - k_2 x_1 + \phi_{21} - \phi_{12})] + E_{22}E_{12} \\
&\times[\cos(\omega_2 t + \omega_1 t - k_2 x_2 - k_1 x_2 + \phi_{22} + \phi_{12}) \\
&+ \cos(\Delta\omega t - k_2 x_2 + k_1 x_2 + \phi_{22} - \phi_{12})] + E_{21}E_{11} \\
&\times[\cos(\omega_2 t + \omega_1 t - k_2 x_1 - k_1 x_1 + \phi_{21} + \phi_{11}) \\
&+ \cos(\Delta\omega t - k_2 x_1 + k_1 x_1 + \phi_{21} - \phi_{11})] + E_{11}E_{12} \\
&\times[\cos(2\omega_1 t - k_1 x_2 - k_1 x_1 + \phi_{11} + \phi_{12}) \\
&+ \cos(k_1 x_2 - k_1 x_1 + \phi_{11} - \phi_{12})] + E_{21}E_{22} \\
&\times[\cos(2\omega_2 t - k_2 x_2 - k_2 x_1 + \phi_{21} + \phi_{22}) \\
&+ \cos(k_2 x_2 - k_2 x_1 + \phi_{21} - \phi_{22})], \qquad (1)
\end{aligned}$$

where $\Delta\omega = \omega_2 - \omega_1$, and $\omega_{1,2} = 2\pi f_{1,2}$

Eq. (1) gives a generic expression for the detector current in the case of a four F–P element system. For the interferometer considered here, we may apply several simplifications to this equation. First, path 1 may be arbitrarily designated as the reference leg and path 2 as the measurement leg. Therefore, x_1 is ideally constant and may be set equal to zero. Second, for a relatively small split frequency between the two heterodyne signals (typically 20 MHz or less in commercially available systems), the propagation constants, k_1 and k_2, are nearly equal and a single value, k, may be substituted for each. Third, we may neglect the initial phase of the signals. Finally, due to limited detector bandwidth, those terms which oscillate at twice the optical frequency, or terms which contain $2\omega_1 t$, $2\omega_2 t$, or $\omega_1 t + \omega_2 t$, may be ignored. After the application of these simplifications, the detector current may be rewritten as the sum of the four types of interference terms described previously; see Eq. (2):

$$\begin{aligned}
I \propto E^2 = &\underbrace{\left(\frac{E_{11}^2}{2} + \frac{E_{21}^2}{2} + \frac{E_{22}^2}{2} + \frac{E_{12}^2}{2}\right)}_{\text{optical power}} \\
&+\underbrace{E_{22}E_{11}\cos(\Delta\omega t - kx_2)}_{\text{ac interference (intended)}} \\
&+\underbrace{E_{21}E_{12}\cos(\Delta\omega t + kx_2)}_{\text{ac interference (leakage-induced)}}+\underbrace{E_{22}E_{12}\cos(\Delta\omega t)}_{\text{ac reference (path 1)}} \\
&+\underbrace{E_{21}E_{11}\cos(\Delta\omega t)}_{\text{ac reference (path2)}} \quad +\underbrace{E_{11}E_{12}\cos(kx_2)}_{\text{dc interference (frequency 1)}} \\
&+\underbrace{E_{21}E_{22}\cos(kx_2)}_{\text{dc interference (frequency 2)}}. \qquad (2)
\end{aligned}$$

Eq. (2) lists all 10 distinct interference terms present in a fully leaking two frequency interferometer. The intended *ac interference* term is defined by the interference of the E_{22} and E_{11} terms, as expected (see Fig. 3a). It represents the signal of choice in commercial heterodyne systems. A second *ac interference* term is obtained due to interference between the leakage terms E_{21} and E_{12}, as shown in Fig. 3b. This term, generally referred to as the spatial second-harmonic nonlinearity in the literature (e.g. [10]), carries a Doppler phase shift, due to target motion, of equal amplitude, but opposite sign relative to the intended ac term.

The *ac reference* terms, also referred to as first-harmonic nonlinearities in the literature (e.g. [10]), occur due to interference between the *intended and leakage terms of different frequencies that exist in a single path* of the interferometer (Fig. 3c and d). As noted previously, these terms exhibit one full cycle of phase shift over a distance defined by the synthetic wavelength ($1/\lambda_2 - 1/\lambda_1$). For a 20 MHz split frequency, for example, the required displacement is approximately 7.5 m for a single pass interferometer with a fold factor of two. A common application of this high phase resolution is calibration of the phase measuring electronics' linearity. In this method, the full phase cycle displacement defined by the frequency split is commanded on a long travel stage, the actual displacement is recorded using a separate interferometer, and the measured phase for the system electronics to be calibrated is compared to the independently measured displacement [30].

Two *dc interference* terms also exist because *each source frequency is found in each path* (see Fig. 3e and f). For a commercial single frequency, or homodyne, system, the corresponding *dc interference* term is the selected measurement

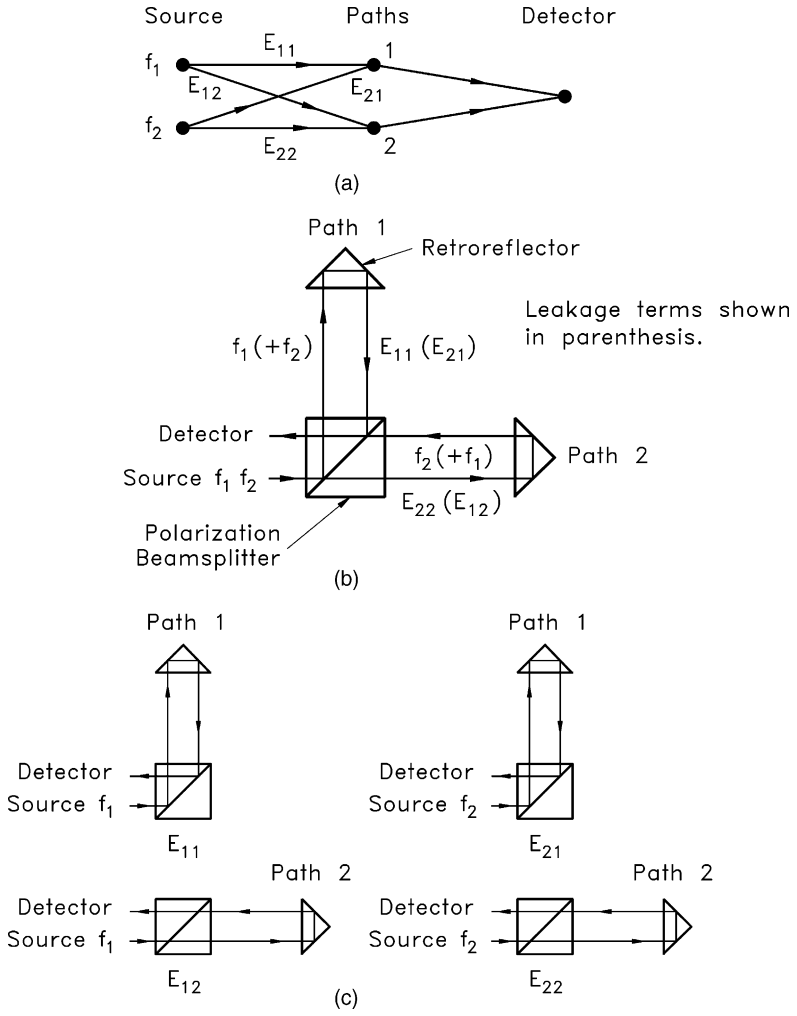

Fig. 2. Heterodyne Michelson-type interferometer: (a) F–P model with four F–P elements for this configuration; (b) physical representation of optics showing F–P elements; and (c) identification of physical origin of each F–P element.

signal ($E_{11}E_{12}$ for frequency f_1). These terms exhibit a positive Doppler phase shift of equal amplitude to the *ac interference* terms regardless of target motion direction, but do not include the $\Delta\omega t$ phase offset seen in the *ac interference* terms, i.e. zero frequency shift when the moving target is at rest. Also, in an analogous manner to the *ac reference* terms, the *dc interference* terms interfere with one another over the synthetic wavelength.

Finally, the *optical power* terms are considered. They contribute a dc, or zero frequency, offset to the photodetector current regardless of optical path changes. In general, they do not appear near the intended *ac interference* term in the frequency domain and, therefore, do not significantly affect the measured displacement. In the case of a homodyne interferometer, however, the intended *dc interference* term exists within the same passband for low or zero velocity target motion and variations in the *optical power* terms cause measurement errors due to the inability of the frequency-selective electronic circuitry to separate the two. This is a fundamental weakness of traditional homodyne interferometers and led to the development of (1) heterodyne interferometry; and (2) more sophisticated optical configurations for homodyne systems.

Fig. 4 shows a graphical representation of the interference terms that result from combinations of the four F–P elements in a heterodyne Michelson-type interferometer. In this figure, the solid circles signify the F–P elements and the diameters of these circles indicate their relative amplitudes, i.e. it is expected that the intended E_{11} and E_{22} terms will be larger than the E_{12} and E_{21} leakage terms, the horizontal lines denote *dc interference* terms, the vertical lines represent *ac reference* terms, and the diagonal lines designate *ac interference* terms. The *optical power* interference terms are represented by loops originating from and terminating at each F–P element since *optical power* terms are the product of an F–P element with itself.

Each of these interference terms may be seen in the frequency spectrum of the detector current. When the measurement path target is at rest, the intended *ac interference* term, leakage-induced *ac interference* term, and *ac reference* terms are found at the beat frequency, which is equal to

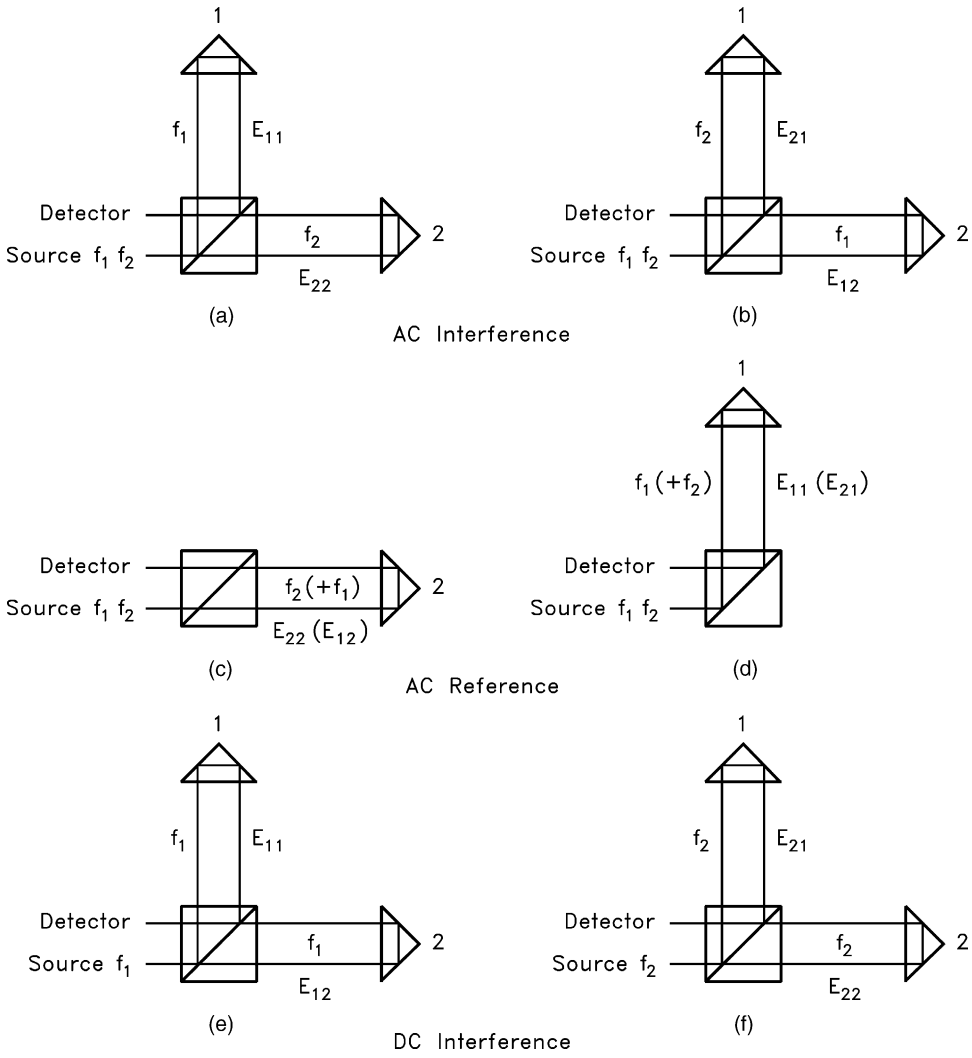

Fig. 3. Physical representation of single pass, Michelson-type heterodyne interferometer showing combinations of F–P elements that lead to *ac interference* (a) and (b); *ac reference* (c) and (d), and *dc interference* (e) and (f), terms (*optical power* terms are not shown).

$\Delta\omega/(2\pi)$, while the *dc interference* and *optical power* terms are located at zero frequency. During a constant velocity target motion, the intended *ac interference* term is either up or down frequency shifted, depending on the moving target direction, by the scalar amount Δf_d, which is proportional to the target velocity (a Δf_d down-shift in the beat frequency is seen in Fig. 6). The leakage-induced *ac interference* term is shifted by the same amount in the opposite direction (up-shifted by Δf_d in this case). The *ac reference* terms remain at the original beat frequency, except for the slight phase shift introduced by the periodicity of the interference between the two terms over the synthetic wavelength. The *dc interference* terms are upshifted from zero by Δf_d and the *optical power* terms remain at zero frequency. Figs. 5 and 6 display a representation of the spectral content for the static and low constant velocity cases. Fig. 7 shows the spectrum for a higher velocity translation of the moving target in the same direction. Recall that the relative amplitude of the individual signals and corresponding periodic error is defined by the amount of leakage that occurs due to deviations of the physical realization of the heterodyne interferometer from the ideal.[2] An example calculation of the periodic error for a

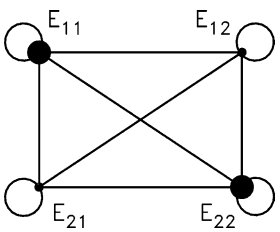

Fig. 4. Graphical representation of interference terms resulting from combinations of F–P elements in heterodyne Michelson-type interferometer, where the solid circles signify the F–P elements, horizontal lines denote *dc interference* terms, vertical lines represent *ac reference* terms, diagonal lines designate *ac interference* terms, and loops identify the *optical power* terms.

[2] The authors also acknowledge that a misalignment between the mechanical and interferometric axes (i.e. cosine error) will lead to a variation in

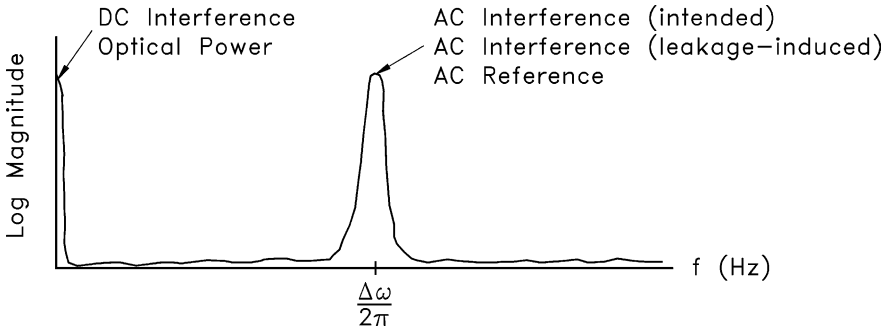

Fig. 5. Heterodyne interferometer frequency spectrum for the static case.

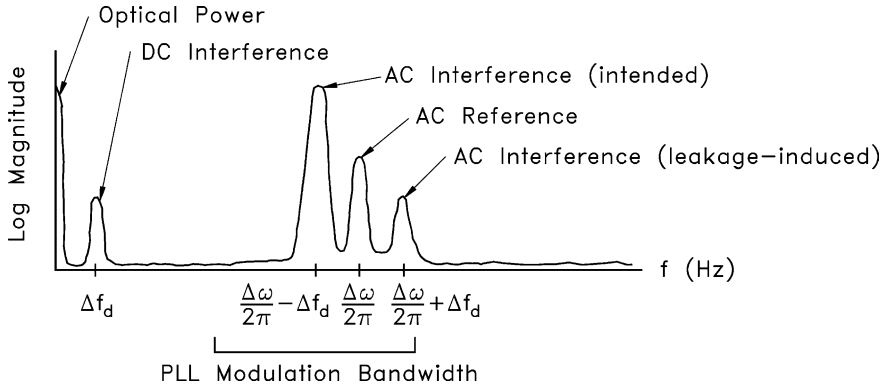

Fig. 6. Heterodyne interferometer frequency spectrum for low velocity target motion.

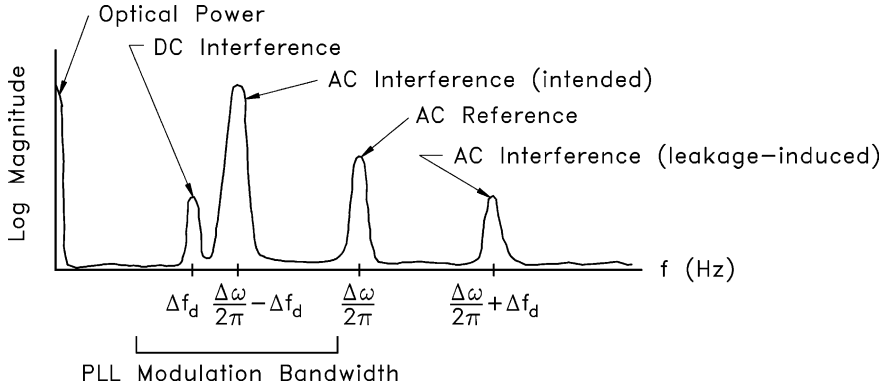

Fig. 7. Heterodyne interferometer frequency spectrum for high velocity target motion.

particular system from spectrum analyzer data, as well as the associated measurement uncertainty, is given in Appendix A.

3.3. *Dynamic periodic error*

The phase measuring electronics for heterodyne interferometers may be considered roughly analogous to phase-modulated radio receivers. The intended *ac interference* term has a carrier frequency equal to the heterodyne beat frequency. The modulation of this carrier frequency is defined as the total phase deviation (or the time integral of the cumulative frequency shift) of the measurement signal with respect to the phase of the reference signal[3] and contains the displacement information. A phase-locked loop (PLL) is used to track this modulation because the peak value is a very large amount of phase, equal to the number of wavelengths of optical path change, and, unlike a radio, causes the instantaneous frequency of the modulated carrier signal to change by large amounts (3.16 MHz/(m s) for a single pass helium–neon, or He–Ne, heterodyne system) under

measurement signal intensity due to a change in overlap (or shear) between the reference and measurement beams.

[3] The reference signal is the direct interference of the two heterodyne frequencies prior to their traveling to the interferometer optics. It can be derived, for example, inside the laser head and made available as a separate input to the measurement electronics.

high velocity target motions. The ability of the PLL to track rapid fluctuations in phase, i.e. rapid phase deviations from the instantaneous frequency, is inherently limited by the PLL filter to approximately 100 kHz, termed the modulation bandwidth. Although all the leakage-induced periodic error terms are always present in the photodiode current, they only pass through the phase measuring electronics when their frequency is within this modulation bandwidth of the intended *ac interference* term instantaneous frequency.

During low velocity motions, the instantaneous frequency of the intended *ac interference* term remains close to the reference frequency and the leakage-induced *ac interference* and *ac reference* terms dominate the periodic error because the *dc interference* and *optical power* terms are outside the PLL modulation bandwidth (see Fig. 6). Under high-speed motions, this may no longer be the case.[4] As the velocity and corresponding Doppler frequency shift increase, the leakage-induced *ac interference* and *ac reference* terms move outside the PLL passband causing an attenuation or extinction of the *pseudo-static periodic error*. For high velocities (demonstrated in Fig. 7), the *dc interference* terms now begin to approach the intended *ac interference* signal. Once both the *dc interference* and intended *ac interference* terms are within the PLL modulation bandwidth, an analogous *dynamic periodic error* is now seen. Unlike the pseudo-static case, where an overlap of the intended and leakage-induced signals occurs only at rest, the dynamic situation gives an overlap of the intended *ac* and *dc interference* terms at a finite velocity, corresponding to a Doppler shift of half the beat frequency. This velocity, v_{dynamic} (in m/s), may be calculated according to Eq. (3):

$$v_{\text{dynamic}} = \frac{\Delta\omega}{2\pi}\frac{\lambda}{2\text{FF}}, \tag{3}$$

where $\Delta\omega$ is the beat frequency (in rad/s), λ is the nominal laser wavelength (in m), and FF is the fold factor (2 for single pass, 4 for double pass, etc.). If we consider a double pass He–Ne interferometer with a frequency split of 3 MHz, this critical velocity is approximately 0.2374 m/s, which is well within the range of today's high-speed/high acceleration linear stages. For a 20 MHz beat frequency He–Ne double pass interferometer, the critical velocity is increased to 1.5825 m/s.

It should also be noted that there is a directional dependence associated with the dynamic periodic error. For a high-speed target motion in the opposite direction as shown in Figs. 6 and 7, the intended *ac* and *dc interference* terms are up-shifted in frequency by the same amount, the leakage-induced *ac interference* term is now down-shifted and will overlap with the *dc interference* signals at v_{dynamic}, and the *ac reference* terms essentially remain fixed at the beat frequency. Therefore, no unwanted interference terms lie within the PLL modulation bandwidth and both the *pseudo-static* and *dynamic periodic error* contributions are attenuated. For example, with a 100 kHz PLL modulation bandwidth, a target velocity greater than 0.01583 m/s effectively filters the *pseudo-static* and *dynamic periodic errors* from the measured displacement for a single pass interferometer.

[4] Bobroff briefly noted this effect in references [6] and [18]. However, at the time of publication, high velocity displacement measurements were not being actively explored and this periodic error has remained neglected.

3.4. Intermodulation periodic error

As noted previously, the photodetector functions as a bridge between the interferometer optics and electronics. The optical components act on one or more source frequencies to produce an electric field composed of various F–P elements. The photodetector then converts the optical signal to an electrical current composed of various interference terms, each of which may be described by one of four categories of optical path change dependence. The action of the interferometer electronics is to first amplify the photodetector current, and then extract the displacement information from this amplified signal according to some assumed relationship between current variations and target motion. We are concerned here with the possible nonlinear behavior of the electronic amplifier and its effect on periodic error [30].

Two-tone, second and third-order intermodulation testing is commonly carried out by electrical engineers/scientists on RF/IF signal processing components. These measurements provide a means for quantifying the nonlinear behavior of the device under test. Two tones, or frequencies, are applied to the input of an amplifier, for example, and the output frequency spectrum is analyzed. Rather than observing only the linearly amplified input signals, the frequency signature of the output includes the input signals, as well as higher-order harmonics of the sum of the input terms (i.e. the Taylor series expansion of the input sum). A typical frequency response for RF/IF intermodulation (IMD) testing is shown in Fig. 8, where the input signals (at frequencies f_1 and f_2) and second- and third-order products are identified.

As shown in Fig. 8, if two sinusoidal signals of frequencies f_1 and f_2 ($f_1 > f_2$) are applied to a nonideal amplifier, the nonlinear squaring action produces signals at frequencies $2f_1$, $2f_2$, $f_1 + f_2$, and $f_1 - f_2$ (see Eq. (4)). The cube of the sum of the input signals gives terms at $f_1, f_2, 3f_1, 3f_2, 2f_1 + f_2, 2f_2 + f_1, 2f_1 - f_2$, and $2f_2 - f_1$ (see Eq. (5)). The reader should note that the latter two terms occur at frequencies near the original input signals. It is the sum of the original linearly amplified signals and the second-, third-, and higher-order harmonics, with appropriate amplification coefficients applied, that gives the final output of the nonlinear amplifier:

$$\begin{aligned}(\cos(2\pi f_1)t + \cos(2\pi f_2)t)^2 &= 1 + \tfrac{1}{2}(\cos2(2\pi f_1)t \\ &+ \cos2(2\pi f_2)) + \cos((2\pi f_1) + (2\pi f_2))t \\ &+ \cos((2\pi f_1) - (2\pi f_2))t,\end{aligned} \tag{4}$$

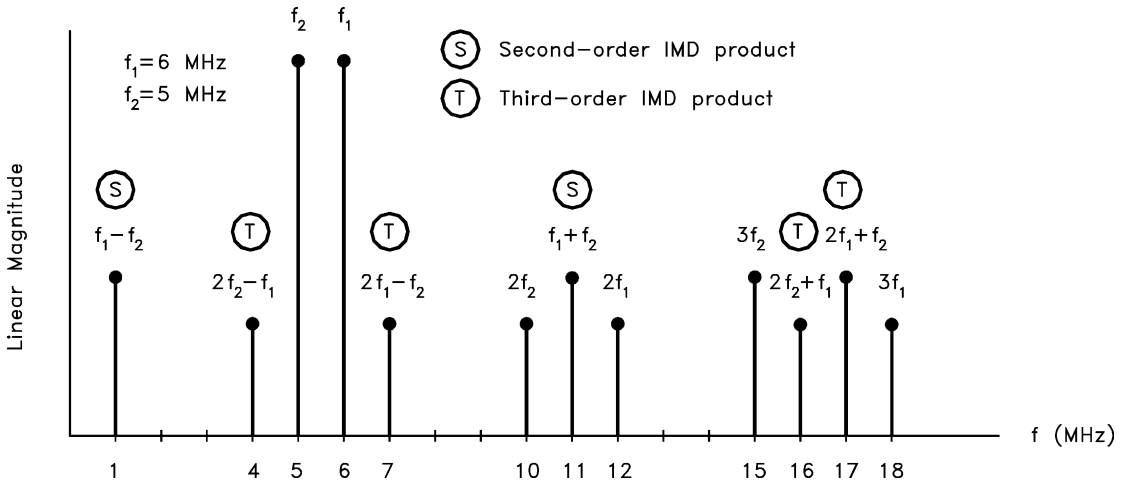

Fig. 8. Intermodulation testing frequency content [31].

$$(\cos(2\pi f_1)t + \cos(2\pi f_2)t)^3 = \tfrac{9}{4}(\cos(2\pi f_1)t + \cos(2\pi f_2)t) + \tfrac{1}{4}(\cos 3(2\pi f_1)t + \cos 3(2\pi f_2)t) + \tfrac{3}{4}(\cos(2(2\pi f_1) + (2\pi f_2))t + \cos(2(2\pi f_1) - (2\pi f_2))t) + \tfrac{3}{4}(\cos(2(2\pi f_2) + (2\pi f_1))t + \cos(2(2\pi f_2) - (2\pi f_1))t). \quad (5)$$

For differential-path interferometers, we have seen that the photodetector current, which is subsequently amplified, is composed of several interference terms of four different types. We will next consider the effect of the nonlinear amplifier on the sum of these *optical power*, *ac interference*, *ac reference*, and *dc interference* terms. The interference terms found in a single pass, Michelson-type heterodyne interferometer were previously derived in Eq. (2). We will continue with this example and determine the new terms introduced by the Taylor series expansion of the sum of these signals.

As shown in Eq. (2), the 10 distinct interference terms provide five separate frequency/phase relationships: *dc*, $\Delta\omega t - kx_2$, kx_2, $\Delta\omega t + kx_2$, and $\Delta\omega t$. The frequency-based amplifier nonlinearity can therefore be expressed as the addition of the collection of these five sinusoids and the square and cube of the sum of the five terms (neglecting higher orders). Eq. (6) gives the square of the sum; the cubic expansion is shown in Eq. (7). In each of these equations, unit coefficients have been applied to the input sinusoids for simplicity. In reality, however, their relative magnitudes will depend on the physical setup of the interferometer as discussed in Section 3.1:

$$(1 + \cos(\Delta\omega t - kx_2) + \cos(kx_2) + \cos(\Delta\omega t + kx_2) + \cos(\Delta\omega t))^2 = 3 + 3\cos(\Delta\omega t - kx_2) + 4\cos(kx_2) + 3\cos(\Delta\omega t + kx_2) + 4\cos(\Delta\omega t) + \cos(\Delta\omega t - 2kx_2) + \cos(\Delta\omega t + 2kx_2) + \tfrac{1}{2}\cos(2\Delta\omega t - 2kx_2) + \tfrac{3}{2}\cos(2kx_2) + \tfrac{1}{2}\cos(2\Delta\omega t + 2kx_2) + \tfrac{3}{2}\cos(2\Delta\omega t) + \cos(2\Delta\omega t - kx_2) + \cos(2\Delta\omega t + kx_2), \quad (6)$$

$$(1 + \cos(\Delta\omega t - kx_2) + \cos(kx_2) + \cos(\Delta\omega t + kx_2) + \cos(\Delta\omega t))^3 = 10 + \tfrac{51}{4}\cos(\Delta\omega t - kx_2) + \tfrac{63}{4}\cos(kx_2) + \tfrac{51}{4}\cos(\Delta\omega t + kx_2) + \tfrac{63}{4}\cos(\Delta\omega t) + \tfrac{3}{2}\cos(\Delta\omega t - 3kx_2) + \tfrac{3}{2}\cos(\Delta\omega t + 3kx_2) + 6\cos(\Delta\omega t - 2kx_2) + 6\cos(\Delta\omega t + 2kx_2) + 3\cos(2\Delta\omega t - 2kx_2) + \tfrac{15}{2}\cos(2kx_2) + 3\cos(2\Delta\omega t + 2kx_2) + \tfrac{15}{2}\cos(2\Delta\omega t) + \tfrac{3}{4}\cos(2\Delta\omega t - 3kx_2) + \tfrac{3}{4}\cos(2\Delta\omega t + 3kx_2) + 6\cos(2\Delta\omega t - kx_2) + 6\cos(2\Delta\omega t + kx_2) + \tfrac{1}{4}\cos(3\Delta\omega t - 3kx_2) + \tfrac{7}{4}\cos(3kx_2) + \tfrac{1}{4}\cos(3\Delta\omega t + 3kx_2) + \tfrac{7}{4}\cos(3\Delta\omega t) + \tfrac{3}{4}\cos(3\Delta\omega t - 2kx_2) + \tfrac{3}{4}\cos(3\Delta\omega t + 2kx_2) + \tfrac{3}{2}\cos(3\Delta\omega t - kx_2) + \tfrac{3}{2}\cos(3\Delta\omega t + kx_2). \quad (7)$$

These expressions contain specific subgroups of terms. In Eq. (6), two subgroups are formed due to squaring of the photodetector current: (1) the original interference terms with additional sidebands at a frequency separation of twice the Doppler shift (i.e. $\Delta\omega t - 2kx_2$ and $\Delta\omega t + 2kx_2$), and (2) a comparable set of signals centered around two times the original beat frequency ($2\Delta\omega t$). The frequency domain representation of these interference terms is shown in Fig. 9, where v_2 is the time rate of change of x_2. In a similar manner, Eq. (7), the cubic expansion generates three distinct frequency subgroups: (1) the input terms with additional sidebands at frequency separations of two and three times the Doppler shift, (2) comparable frequency content at two times the beat frequency, and (3) similar terms at three times the beat frequency. Fig. 10 displays the frequency content of Eq. (7).

In each case, additional sidebands are located near the intended *ac interference* term and lead to higher-order periodic errors. Furthermore, the second- and third-order *dc interference* terms, $\cos(2kx_2)$ and $\cos(3kx_2)$, respectively, lead to multiple dynamic periodic errors at velocities defined by Eq. (8), where j is the order of the *dc interference* term. Note that

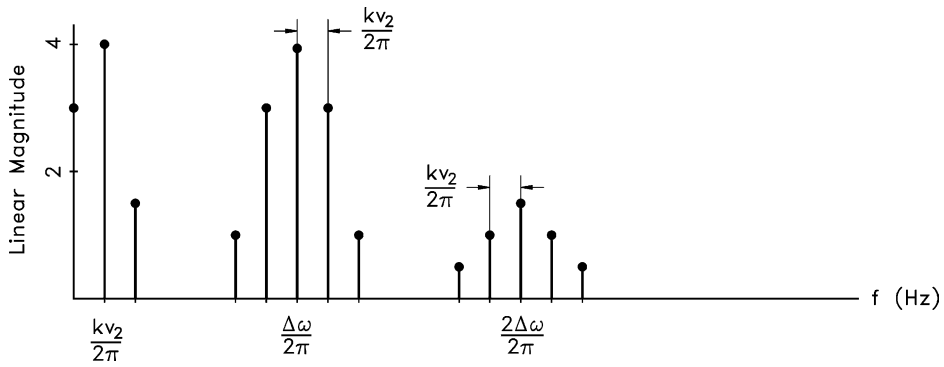

Fig. 9. Squared sum frequency content.

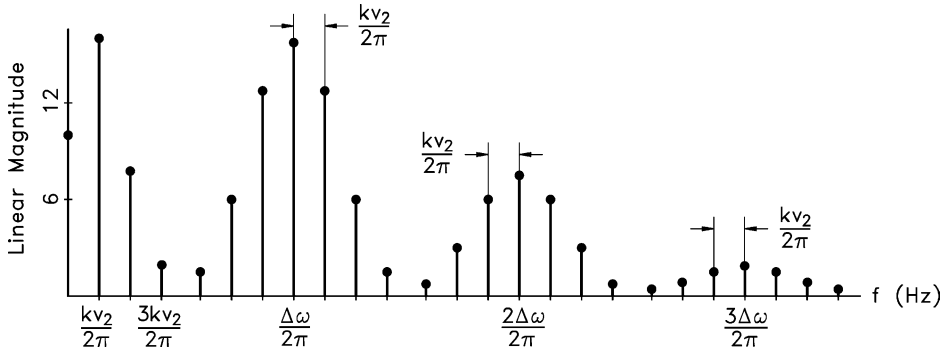

Fig. 10. Cubed sum frequency content.

this expression is equivalent to Eq. (3) when j is equal to one:

$$v_{\text{dynamic}} = \frac{\Delta\omega}{2\pi}\frac{\lambda}{(j+1)\,\text{FF}}. \tag{8}$$

Although multiple Doppler shift terms can lead to increased periodic error, they are not without merit. If frequency-selective signal processing methods could be utilized to isolate these terms (i.e. $\Delta\omega t - 2kx_2$, $\Delta\omega t + 2kx_2$, $\Delta\omega t + 3kx_2$, $\Delta\omega t - 3kx_2$, $2kx_2$, or $3kx_2$), a two or three times increase in the interferometer resolution could be realized with no change in optical configuration. For example, the resolution obtained using a double pass, heterodyne Michelson-type interferometer could be achieved with the combination of a single pass setup and use of the twice Doppler shifted *ac interference* term, $\Delta\omega t - 2kx_2$.

Alternately, the user could choose to generate a synthetic beat frequency of two or three times the original beat frequency, i.e. $2\Delta\omega t$ or $3\Delta\omega t$. The second- or third-order *ac interference* term, $2\Delta\omega t - 2kx_2$ or $3\Delta\omega t - 3kx_2$, could then be compared to this reference signal and used to determine the target motion with a two or three times increase in resolution, respectively. The existence of second- and third-order, or single, double, and triple Doppler shifted, *ac reference* and *dc interference* terms, however, would still lead to unwanted periodic errors. In either case, these methods would take advantage of the amplifier nonlinearity to extend the system resolution without physical changes to the interferometer.

4. Conclusions

In this paper, we describe a *dynamic periodic error*, differentiated from the *pseudo-static periodic errors* previously explored in the literature, that is exhibited by heterodyne interferometer systems under high-speed displacements. This error occurs when the intended *ac interference* and unwanted *dc interference* terms, that exist due to frequency leakage in physical implementations of heterodyne interferometers, are both present within the phase measuring electronics' modulation bandwidth. Applications that may exhibit this dynamic periodic error include any linear positioning device where displacement knowledge is required during translation, especially during high velocity motions. This error may have particular significance for the semiconductor industry where research is being conducted regarding the production of constant velocity lithography stages. These stages would attempt to reduce manufacturing time by exposing the wafers during high-speed motion, rather than relying on the step-and-repeat method currently in use.

We also provide a description of *intermodulation periodic error* caused by amplifier nonlinearity. It is shown that squaring and cubing the sum of the interference terms found in the detector current leads to sidebands of the input signals at frequency separations of two and three times the Doppler shift, as well as comparable frequency content at two and three times the beat frequency. These multiple

Doppler shifted terms contribute higher-order *psuedo-static* and *dynamic* periodic errors. However, with proper signal conditioning, they may also lead to interferometer resolution extension.

Finally, we provide an intuitive Frequency–Path model of the propagation of light from the source to detector in differential-path interferometers. This model identifies each possible path for each light frequency from the source to detector and predicts the number of interference terms that may be expected at the detector output. We show that, regardless of the interferometer configuration, the behavior of each interference term with respect to optical path changes may be grouped into one of four categories: *optical power*, *ac interference*, *ac reference*, and *dc interference*. A description of a single pass, Michelson-type heterodyne interferometer is given, but the analysis extends to other cases as well. For example, it is beneficial in many instances to carry the two heterodyne frequencies from the laser head to interferometer using single mode, polarization maintaining (SMPM) fiber. This situation is analogous to the heterodyne system described in the text because there are again two paths, defined by the birefringence in the optical fiber in this case, and two source frequencies, except that we cannot assume zero motion in either path due to the existence of varying mechanical and/or thermal deformations of the fiber. Additionally, the combination of the SMPM launch with the heterodyne interferometer yields a new four path, two frequency system with eight Frequency–Path elements and increases the number of distinct interference terms from 10 (for the four Frequency–Path element system) to 36.

Appendix A. Periodic error calculation and associated uncertainty

Fig. 11 displays an example spectrum for a constant (low) velocity displacement measurement using a Michelson-type, double pass heterodyne interferometer. In the figure, the horizontal scale is 5 kHz per division and the vertical scale is 10 dB per division. Additionally, the left peak is the intended *ac interference* term, the center peak gives the two superimposed *ac reference* terms (at a beat frequency of 20 MHz), and the right peak is the leakage-induced *ac interference* term. Due to its larger amplitude, the contribution from the *ac reference* terms dominates the periodic error in this case. The magnitude of the periodic error, δx, is calculated according to Eq. (A.1), where λ is the source wavelength (633 nm), FF is the fold factor (=4 for a double pass configuration), and ΔdB is the attenuation of the *ac reference* signal with respect to the intended *ac interference* term (14 dB from Fig. 11):

$$\delta x = \frac{\lambda}{\mathrm{FF}}\left(\frac{10^{-\Delta\mathrm{dB}/20}}{2\pi}\right) = \frac{633}{4}\left(\frac{10^{-14/20}}{2\pi}\right) = 5.03\,\mathrm{nm}. \tag{A.1}$$

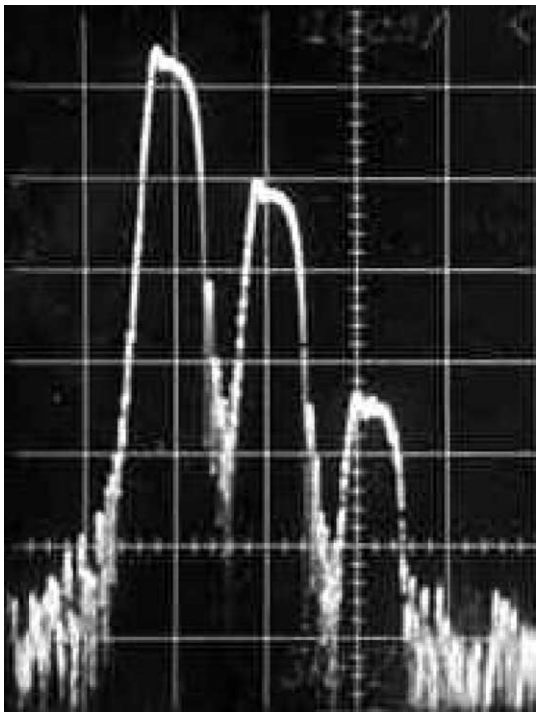

Fig. 11. Example spectral content for double pass, Michelson-type heterodyne interferometer.

The combined standard uncertainty, u_c, in this periodic error value may be determined according to the *law of propagation of uncertainty* [32]. As demonstrated in Eq. (A.2), the variance, u_c^2, in a measured value may be obtained by first calculating the partial derivatives of the functional relationship between the measurand (δx) and input parameters (λ and ΔdB) with respect to each input; then multiplying the partial derivatives by the standard uncertainty in the associated input (zero covariance between input quantities has been assumed); and, finally, summing the squares of each product. For standard uncertainties of 6×10^{-6} nm in wavelength (0.01 ppm) and 0.5 dB in amplitude, the combined standard uncertainty in the periodic error measurement result is 0.58 nm. The reader may note that this value reflects the uncertainty in the periodic error computed from the interference signal alone. Further signal corruption, which could be imposed by intermodulation distortion, for example, is not considered:

$$\begin{aligned} u_c^2(\delta x) &= \left(\frac{\delta f}{\delta\lambda}u(\lambda)\right)^2 + \left(\frac{\delta f}{\delta\,\Delta\mathrm{dB}}u(\Delta\mathrm{dB})\right)^2 \\ &= \left(\frac{10^{-\Delta\mathrm{dB}/20}}{4\pi}u(\lambda)\right)^2 \\ &\quad + \left(\frac{-\lambda\ln(10)10^{-\Delta\mathrm{dB}/20}}{80\pi}u(\Delta\mathrm{dB})\right)^2 = 0.58^2\,\mathrm{nm}^2. \end{aligned} \tag{A.2}$$

References

[1] Hirst G. Personal communication with Beckwith J. March 1980.
[2] Fedotova G. Analysis of the measurement error of the parameters of mechanical vibrations. Meas Tech 1980;23(7):577–80.
[3] Quenelle R. Nonlinearity in interferometric measurements. Hewlett-Packard J 1983;34(4):10.
[4] Sutton C. Nonlinearity in length measurements using heterodyne laser Michelson interferometry. J Phys E: Sci Instrum 1987;20:1290–2.
[5] Barash V, Fedotova G. Heterodyne interferometer to measure vibration parameters. Meas Tech 1984;27(7):50–1.
[6] Bobroff N. Residual errors in laser interferometry from air turbulence and nonlinearity. Appl Opt 1987;26(13):2676–82.
[7] Rosenbluth A, Bobroff N. Optical sources of nonlinearity in heterodyne interferometers. Precision Eng 1990;12(1):7–11.
[8] Stone J, Howard L. A simple technique for observing periodic nonlinearities in Michelson interferometers. Precision Eng 1998;22(4):220–32.
[9] Patterson S, Beckwith J. Reduction of systematic errors in heterodyne interferometric displacement measurement. In: Proceedings of the 8th International Precision Engineering Seminar (IPES), Compiegne, France; 1995. p. 101–4.
[10] Badami V, Patterson S. A frequency domain method for the measurement of nonlinearity in heterodyne interferometry. Precision Eng 2000;24(1):41–9.
[11] Badami V, Patterson S. Investigation of nonlinearity in high-accuracy heterodyne laser interferometry. In: Proceedings of the 12th Annual American Society for Precision Engineering (ASPE) Conference, Norfolk, VA; 1997. p. 153–6.
[12] Wu C, Deslattes R. Analytical modeling of the periodic nonlinearity in heterodyne interferometry. Appl Opt 1998;37(28):6696–700.
[13] Wu C, Su C. Nonlinearity in measurements of length by optical interferometry. Meas Sci Technol 1996;7:62–8.
[14] Hou W, Wilkening G. Investigation and compensation of the nonlinearity of heterodyne interferometers. Precision Eng 1992;14(2):91–8.
[15] Hou W, Zhao X. Drift of nonlinearity in the heterodyne interferometer. Precision Eng 1994;16(1):25–35.
[16] Howard L, Stone J. Computer modeling of heterodyne interferometer errors. Precision Eng 1995;12(1):143–6.
[17] Tanaka M, Yamagami T, Nakayama K. Linear interpolation of periodic error in a heterodyne laser interferometer at subnanometer levels. IEEE Trans Instrum Meas 1989;38(2):552–4.
[18] Bobroff N. Recent advances in displacement measuring interferometry. Meas Sci Technol 1993;4:907–26.
[19] Steinmetz C. Sub-micron position measurement and control on precision machine tools with laser interferometry. Precision Eng 1990;12(1):12–24.
[20] Cretin B, Xie W, Wang S, Hauden D. Heterodyne interferometers: practical limitations and improvements. Opt Commun 1988;65(3):157–62.
[21] Petru F, Cip O. Problems regarding linearity of data of a laser interferometer with a single-frequency laser. Precision Eng 1999;23(1):39–50.
[22] Augustyn W, Davis P. An analysis of polarization mixing errors in distance measuring interferometers. J Vacuum Sci Technol B 1990;8(6):2032–6.
[23] Xie Y, Yu Y. Zeeman laser interferometer errors for high precision measurements. Appl Opt 1992;31(7):881–4.
[24] De Freitas J, Player M. Importance of rotational beam alignment in the generation of second harmonic errors in laser heterodyne interferometry. Meas Sci Technol 1993;4:1173–6.
[25] De Freitas J, Player M. Polarization effects in heterodyne interferometry. J Mod Opt 1995;42(9):1875–99.
[26] De Freitas J. Analysis of laser source birefringence and dichroism on nonlinearity in heterodyne interferometry. Meas Sci Technol 1997;8:1356–9.
[27] Li B, Liang J. Effects of polarization mixing on the dual-wavelength heterodyne interferometer. Appl Opt 1997;36(16):3668–72.
[28] Park B, Eom T, Chung M. Polarization properties of cube-corner retroreflectors and their effects on signal strength and nonlinearity in heterodyne interferometers. Appl Opt 1996;35(22):4372–80.
[29] Wu C, Lawall J, Deslattes R. Heterodyne interferometer with subatomic periodic nonlinearity. Appl Opt 1999;38(19):4089–94.
[30] Oldham N, Kramar J, Hetrick P, Teague E. Electronic limitations in phase meters for heterodyne interferometry. Precision Eng 1993;15(3):173–9.
[31] http://www.analog.com/publications/magazines/Dialogue/Anniversary/13.html.
[32] Taylor B, Kuyatt C. Guidelines for evaluating and expressing the uncertainty of NIST measurement results. NIST Technical Note 1297; 1996.