First and second order periodic error measurement for non-constant velocity motions

Tony L. Schmitz [a,*], David C. Chu [b], Hyo Soo Kim [a]

[a] Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32608, USA
[b] Agilent Technologies, Inc., 5301 Stevens Creek Blvd, Santa Clara, CA 95051, USA

ARTICLE INFO

ABSTRACT

This paper describes a digital algorithm that can be applied in real time to measure and compensate first and second order periodic error in heterodyne displacement measuring interferometers. Comparisons are made between the new algorithm and the traditional frequency domain measurement approach, where the error signal is Fourier transformed into the frequency domain to identify periodic error magnitudes. Experimental results are provided for both constant velocity and non-constant velocity conditions.

1. Introduction

Differential-path interferometry offers the unique combination of high accuracy, range, and resolution for demanding, non-contact displacement measurement applications, such as position feedback for precision manufacturing equipment (e.g., lithographic stages for semiconductor fabrication) and transducer calibration. A common interferometer choice in these situations is the heterodyne (or two frequency) Michelson-type setup with single, double, or multiple passes of the optical paths. These systems infer changes in the selected optical path length difference by monitoring the optically induced variation in the photodetector current, which is generated from the optical interference signal. The phase measuring electronics convert this photodetector current to displacement by digitizing the phase progression of the photodetector signal. Due to non-ideal performance, mixing between the two heterodyne frequencies may occur, which results in periodic errors superimposed on the desired displacement data.

A preferred feature of any position transducer is linear operation, where the output signal is directly proportional to displacement. The existence of first or second order periodic errors, with amplitudes that vary cyclically with the target position, however, leads to nonlinear performance in heterodyne interferometers. This limits the achievable accuracy to approximately the nanometer level (or higher) depending on the optical setup, even when the interferometer is operated in vacuum to minimize the error associated with variations in the refractive index of air due to uncompensated fluctuations in temperature, pressure, humidity, and composition. Frequency mixing in the interferometer leads to both first order periodic error, which appears as single sideband (SSB) modulation on the data at a spatial frequency of one cycle per displacement fringe, and second order periodic error, with a spatial frequency of two cycles per displacement fringe.

To compensate for these errors, real-time digital error measurement may be applied. Because it requires no changes to the optical configuration, it enables convenient implementation for existing systems. This paper describes experiments used to evaluate the real time, first and second order periodic error measurement-compensation for both constant and non-constant target velocity conditions [2,3] using an extension of a scheme proposed by Chu and Ray [1]. To differentiate from the original approach [1], the

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

extended method is referred to here as the time domain regression, or TDR, algorithm. The results reported in this paper were obtained using a single pass,[1] heterodyne Michelson-type interferometer that enabled: (1) isolation of periodic error as the primary uncertainty source in displacement measuring interferometry and (2) variation of the frequency mixing that leads to periodic error so that the error amplitude could be changed. During constant velocity target motion (achieved using an air bearing stage), the first and second order errors were measured using the TDR algorithm and traditional frequency domain approach [4–6] for validation purposes. In the frequency domain approach, the periodic error magnitudes are determined by computing the discrete Fourier transform of the time domain displacement data. This measurement step requires constant velocity motions. For this reason, direct comparison between the TDR and frequency domain methods could only be achieved under constant velocity conditions. Various frequency-mixing levels were realized by adjustment of the setup optics. Additionally, data was collected during non-constant velocity motions and the time-dependent first and second order error results were computed using the TDR algorithm. These results are reported. An overview of the digital logic, hardware-based TDR approach is also provided.

2. Periodic error background

As noted, the focus of this work is periodic error measurement and compensation for heterodyne Michelson-type interferometers. In these systems, imperfect separation of the two light frequencies into the measurement (moving) and reference (fixed) paths has been shown to produce first and second order periodic errors. 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 a number of influences, including non-orthogonality 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 frequency travels to a fixed target, while a second, single frequency travels to a moving target and is Doppler shifted according to the target motion. Interference of the combined signals yields a perfectly sinusoidal trace with phase that differs from a corresponding 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 unintended spectral content) and leads to periodic error in the measured displacement.

Early investigations of periodic error in heterodyne Michelson interferometers were completed by Fedotova [7], Quenelle [8], and Sutton [9]. Several researchers have subsequently reported studies of periodic error and its reduction [7–36]. Topics that have been addressed include: (1) periodic error measurement under various conditions [e.g. [10–13]]; (2) frequency domain analysis techniques [4–6]; (3) analytical modeling approaches [14–18]; (4) Jones calculus modeling methods [13,19]; (5) reduction of periodic error [e.g. [4,20,21]]; (6) uncertainty evaluation of interferometric displacement measurement [22,23].

3. Digital algorithm

3.1. Periodic error measurement

In measurement (but not in compensation), the periodic error can be modeled as two pure sine/cosine pairs since the fundamental magnitudes of the periodic errors are relatively insensitive to the presence of their second harmonics. However, the instantaneous periodic error amplitudes (i.e., the projection of the sinusoids on the real axis in the complex plane) are materially affected by the harmonics and this must be addressed in compensation. For measurement, the sum of a parabola and two sinusoid pairs are best fitted to a sequence of 320 position values, ϕ_j, expressed in Unit Intervals (UI) or cycles, where 1 UI = 2π radians. Macroscopic motion up to constant acceleration is removed by the parabola. Therefore, excessive jerk (or the time derivative of acceleration) reduces the effectiveness of this approach. A mathematical model for approximating both first and second order error is shown in Eq. (1):

$$x_0 + x_1 \cdot j + x_2 \cdot k + x_c \cdot \cos(2\pi\phi_j) + x_s \cdot \sin(2\pi\phi_j) + x_{2c} \cdot \cos(4\pi\phi_j)$$
$$+ x_{2s} \cdot \sin(4\pi\phi_j) \approx \phi_j, \tag{1}$$

where $x_0, x_1, x_2, x_c, x_s, x_{2c}$ and x_{2s} are seven parameters to be determined by curve fit. The model implies that the (uncompensated) phase data, ϕ_j, are collected uniformly in time using the index j. Uniform motion is not required, however, due to the quadratic variation, k, which accommodates acceleration. Here, j (velocity term for the motion profile) and k (acceleration term) can be elements of any convenient linear and quadratic index, respectively. For example, j may be integers from 1 to 320 and k may be j^2. However, a custom index set with symmetries will be presented to simplify processing.

Eq. (1) includes both the quadratic phase variation ($x_0 + x_1 j + x_2 k$) as well as two sine/cosine pairs, which are spatially coherent to the displacement fringes. The first pair, $x_c \cdot \cos(2\pi\phi_j) + x_s \cdot \sin(2\pi\phi_j)$, represents one cycle per ϕ_j fringe (UI) and the magnitude is a reasonable representation of first order periodic error magnitude. However, the higher frequency pair, $x_{2c} \cdot \cos(4\pi\phi_j) + x_{2s} \cdot (2\pi\phi_j)$, which represents two cycles per ϕ_j fringe, consists of both the second harmonic of first order error and the fundamental of second order error. They are identical in frequency and inseparable in the frequency domain. Frequency confounding aside, error measurement using Eq. (1) form requires a time consuming algorithm to best fit seven unknown parameters simultaneously. Fortunately, the presence of second order error does not significantly affect the measurement and compensation of first order error (though not vice versa). The first order error can therefore be measured separately using the abbreviated curve fit model shown in Eq. (2).[2]

$$x_0 + x_1 \cdot j + x_2 \cdot k + x_c \cdot \cos(2\pi\phi_j) + x_s \cdot \sin(2\pi\phi_j) \approx \phi_j \tag{2}$$

Eq. (2) enables first order periodic error measurement and compensation using the Chu and Ray algorithm [1], which removes both the fundamental and the second harmonic (of first order error). The resulting first order-compensated phase data is labeled as $\bar{\phi}_j$. A second order curve fit model for $\bar{\phi}_j$ is provided in Eq. (3), which includes the sum of a higher frequency sinusoid pair and the macroscale (quadratic) motion.

$$x_0 + x_1 \cdot j + x_2 \cdot k + x_{2c} \cdot \cos(4\pi\bar{\phi}_j) + x_{2s} \cdot \sin(4\pi\bar{\phi}_j) \approx \bar{\phi}_j \tag{3}$$

---

[1] In other words, one cycle of phase shift was obtained for a $\lambda/2$ target motion (λ = 633 nm).

[2] In measurement, the periodic error can be modeled as a pure sinusoid without incurring significant error in the magnitude/phase results. However, in compensation the actual shape of the periodic error waveform must be accurate, so the second harmonic is considered. See Section 3.3.
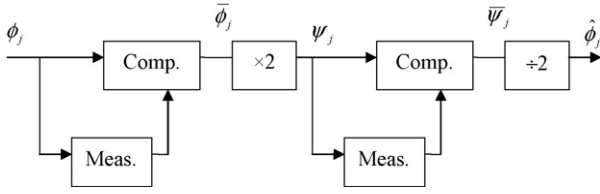
Fig. 1. Signal process flow block diagram for first and second order periodic error real-time measurement and compensation. The measurement and compensation blocks are identical.

The form of Eq. (3) is not readily solvable by the Chu and Ray algorithm. However, by the change of variables: $\psi_j \equiv 2\bar{\phi}_j$, $y_0 \equiv 2x_0$, $y_1 \equiv 2x_1$, $y_2 \equiv 2x_2$, $y_{2c} \equiv 2x_{2c}$, and $y_{2s} \equiv 2x_{2s}$, Eq. (3) becomes Eq. (4).

$$y_0 + y_1 \cdot j + y_2 \cdot k + y_{2c} \cdot \cos(2\pi\psi_j) + y_{2s} \cdot \sin(2\pi\psi_j) = \psi_j \qquad (4)$$

Eq. (4) is identical in form to Eq. (2) and can therefore be solved using the Chu and Ray approach. The solutions, y_{2c} and y_{2s}, are then used to compensate ψ_j, without modification, to form $\bar{\psi}_j$, the first order error-compensated version of ψ_j. Since ψ_j is at the second order spatial frequency, $\bar{\psi}_j$ is therefore equivalent to a second order compensation of $\bar{\phi}_j$. Finally, the desired first and second order error-compensated position phase $\hat{\phi}_j$ is simply half the value of $\bar{\psi}_j$, or $\hat{\phi}_j = \bar{\psi}_j/2$. A block diagram illustrating the process flow is shown in Fig. 1. Even though the measurement-compensation algorithm is performed in two stages for first and second order error, compensation is still completed nearly instantaneously; only the latency is doubled.

3.2. Measurement algorithm for finding periodic error by Eq. (2)

The periodic error measurement algorithm has the same form for both ϕ_j and ψ_j (Eqs. (2) and (4)). Therefore, only ϕ_j (Eq. (2)) is described. The actual algorithm is a curve fit to 320 repetitions of Eq. (2) resulting in the best fit five parameters x_0, x_1, x_2, x_c, and x_s. The parameters x_c and x_s are the complex plane (Cartesian) coordinates of periodic error magnitude and phase. Once identified, they are used to compensate periodic error in the next group (320 points) of data. The other three parameters, x_0, x_1, x_2, are local dummy variables necessary to remove the macroscopic travel trajectory to better reveal the small periodic error. They are not explicitly used in the measurement-compensation algorithm and, therefore, need not be computed. However, for convenience, we use them to remove an underlying parabola to enhance the display of periodic errors (see Figs. 15 and 16). Eq. (2) can be expressed in matrix notation as **MX** = **P**. The 320×5 matrix **M** consists of five columns **I**, **J**, **K**, **C**, **S**. Column **I** consists of 320 '1's. By design, column **J** is a linear vector with odd symmetry, which increases from -159.5 to $+159.5$ in unit steps. Since **J** is incremented in unit steps, its current value, j, is conveniently used as the index for **K**, i.e., $\mathbf{K}^T = [k(-159.5), k(-158.5), \ldots, k(+159.5)]$, where $k(j)$ is the current **K** value indexed by the non-integer j. Column **K** is a U-shaped quadratic function starting at $k(-159.5) = 16,395$. The next value, $k(-158.5)$, is defined by the previous value $k(-159.5)$ incremented by $2j - 1$. Specifically, $k(-158.5) = 16,395 + 2(-158.5) - 1 = 16,077$. Proceeding similarly, the final value $k(+159.5)$ returns to the original value of 16,395. Since j begins negative, k decreases initially, then levels off as j approaches zero and increases as j becomes positive; this establishes the U shape. Column **P** is the uncompensated input vector of 320 data points ϕ_j expressed in UI. Columns **C** and **S** are table lookup values of $\cos(2\pi\phi_j)$ and $\sin(2\pi\phi_j)$, respectively. The five unknown parameters x_0, x_1, x_2, x_c, x_s are elements of the 5×1 vector **X**.

The solution for vector **X** is traditionally computed using the min-squared error regression $\mathbf{X} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{P}$. This implies that the most effective regression operator is $\mathbf{M}^T$. Computation is simplified, however, if a new operator $\mathbf{O}^T$ with a restricted alphabet (1, 0, -1, and -2) is chosen to mimic $\mathbf{M}^T$. The simpler **O** matrix also has five columns: **U**, **L**, **Q**, **E**, and **D**, imitating, respectively, columns **I**, **J**, **K**, **C**, and **S** of **M**. Substitution of **O** for **M** gives the best fit solution shown in Eq. (5). The 5×5 equation set $(\mathbf{O}^T\mathbf{M})\mathbf{X} = \mathbf{O}^T\mathbf{P}$ is shown in Eq. (6).

$$\mathbf{X} = (\mathbf{O}^T\mathbf{M})^{-1}\mathbf{O}^T\mathbf{P} \qquad (5)$$

$$\begin{bmatrix} \mathbf{U}^T\mathbf{I} & \mathbf{U}^T\mathbf{J} & \mathbf{U}^T\mathbf{K} & \mathbf{U}^T\mathbf{C} & \mathbf{U}^T\mathbf{S} \\ \mathbf{L}^T\mathbf{I} & \mathbf{L}^T\mathbf{J} & \mathbf{L}^T\mathbf{K} & \mathbf{L}^T\mathbf{C} & \mathbf{L}^T\mathbf{S} \\ \mathbf{Q}^T\mathbf{I} & \mathbf{Q}^T\mathbf{J} & \mathbf{Q}^T\mathbf{K} & \mathbf{Q}^T\mathbf{C} & \mathbf{Q}^T\mathbf{S} \\ \mathbf{E}^T\mathbf{I} & \mathbf{E}^T\mathbf{J} & \mathbf{E}^T\mathbf{K} & \mathbf{E}^T\mathbf{C} & \mathbf{E}^T\mathbf{S} \\ \mathbf{D}^T\mathbf{I} & \mathbf{D}^T\mathbf{J} & \mathbf{D}^T\mathbf{K} & \mathbf{D}^T\mathbf{C} & \mathbf{D}^T\mathbf{S} \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_c \\ x_s \end{bmatrix} = \begin{bmatrix} \mathbf{U}^T\mathbf{P} \\ \mathbf{L}^T\mathbf{P} \\ \mathbf{Q}^T\mathbf{P} \\ \mathbf{E}^T\mathbf{P} \\ \mathbf{D}^T\mathbf{P} \end{bmatrix} \qquad (6)$$

A judicious design of the five column vectors **U**, **L**, **Q**, **E**, and **D** of **O** further simplifies the $\mathbf{O}^T\mathbf{M}$ matrix product. The selected design for the columns of **O**, arranged in ten blocks of 32 identical elements, is: $\mathbf{U}^T = (1, 1, 0, 1, 1, 1, 1, 0, 1, 1)$, $\mathbf{L}^T = (-1, -1, 0, 0, 0, 0, 0, 0, 1, 1)$, $\mathbf{Q}^T = (1, 1, 0, 0, -2, -2, 0, 0, 1, 1)$, $\mathbf{E}^T = (e, e, e, e, e, e, e, e, e, e)$, and $\mathbf{D}^T = (d, d, d, d, d, d, d, d, d, d)$. For example, the first element, 1, in $\mathbf{U}^T$ is repeated 32 times, the second element, 1, is repeated 32 times and so on to give a total vector length of 320. The elements e and d in $\mathbf{E}^T$ and $\mathbf{D}^T$, respectively, are quantized versions of the cosine and sine functions, equaling 1 whenever the function exceeds $\sqrt{2}/2$, -1 when less than $-\sqrt{2}/2$, and 0 otherwise. With this design, the boundaries for e and d are at octant boundaries of the unit circle. Therefore, when phase ϕ_j is expressed in UI (cycles), the most significant three bits of the fractional part of ϕ_j logically determine the octant and therefore the ternary values of e or d, without table lookup.

With this design, nine elements of the $\mathbf{O}^T\mathbf{M}$ matrix become $\mathbf{U}^T\mathbf{I} = 2^8$, $\mathbf{L}^T\mathbf{J} = 2^{14}$, $\mathbf{Q}^T\mathbf{K} = 2^{21}$, and $\mathbf{U}^T\mathbf{J} = \mathbf{U}^T\mathbf{K} = \mathbf{L}^T\mathbf{I} = \mathbf{L}^T\mathbf{K} = \mathbf{Q}^T\mathbf{I} = \mathbf{Q}^T\mathbf{J} = 0$. The 5×5 equation set $(\mathbf{O}^T\mathbf{M})\mathbf{X} = \mathbf{O}^T\mathbf{P}$ is shown in Eq. (7).

$$\begin{bmatrix} 2^8 & 0 & 0 & \mathbf{U}^T\mathbf{C} & \mathbf{U}^T\mathbf{S} \\ 0 & 2^{14} & 0 & \mathbf{L}^T\mathbf{C} & \mathbf{L}^T\mathbf{S} \\ 0 & 0 & 2^{21} & \mathbf{Q}^T\mathbf{C} & \mathbf{Q}^T\mathbf{S} \\ \mathbf{E}^T\mathbf{I} & \mathbf{E}^T\mathbf{J} & \mathbf{E}^T\mathbf{K} & \mathbf{E}^T\mathbf{C} & \mathbf{E}^T\mathbf{S} \\ \mathbf{D}^T\mathbf{I} & \mathbf{D}^T\mathbf{J} & \mathbf{D}^T\mathbf{K} & \mathbf{D}^T\mathbf{C} & \mathbf{D}^T\mathbf{S} \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_c \\ x_s \end{bmatrix} = \begin{bmatrix} \mathbf{U}^T\mathbf{P} \\ \mathbf{L}^T\mathbf{P} \\ \mathbf{Q}^T\mathbf{P} \\ \mathbf{E}^T\mathbf{P} \\ \mathbf{D}^T\mathbf{P} \end{bmatrix} \qquad (7)$$

A time consuming full 5×5 matrix inversion $(\mathbf{O}^T\mathbf{M})^{-1}$ to determine **X** from Eq. (7) can be avoided by the following steps. The top row of Eq. (7) is used to eliminate the first column, taking full advantage of a convenient power-of-two (2^8) divisor. Likewise, the second and third rows eliminate the second and third columns. All three columns can be eliminated simultaneously, and only a 2×2 equation set remains for determining x_c and x_s, which are the required solutions. The periodic error magnitude, V, and phase, θ, are the polar version of x_c and x_s. Therefore, $V = \sqrt{x_c^2 + x_s^2}$ and $\theta = (1/2\pi)\tan^{-1}(x_c/-x_s)$, where the quadrant dependence of the arctangent function must be observed in implementation.

When travel during the sample time for 320 points (1.024 ms for the 312.5 kHz sampling frequency applied here) is much less than one fringe ($\lambda/2$, where λ is the source wavelength for a single pass interferometer), the rank of the 2×2 matrix may be less than 2 and matrix inversion is problematic. We define this condition to be realized when the input ϕ_j dwells within the same quadrant for more than 64 consecutive points.[3] If true, the new measurement result

---

[3] The motion must traverse at least one fringe in 320 points for the algorithm to be effective. For constant velocity, this yields 80 samples in each quadrant. We
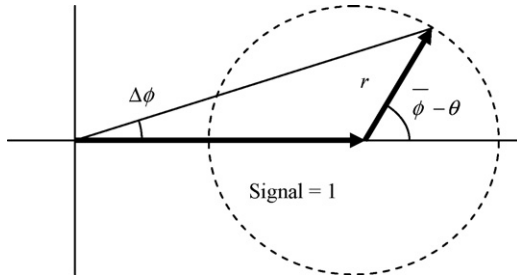
Fig. 2. Vector diagram illustrating the SSB modulation. The ideal signal is represented by the horizontal unity vector. First order periodic error $\Delta\phi$ is produced as a function of the object position $\bar{s} = (\lambda/2)(\bar{\phi} - \theta)$ when a smaller signal of fractional magnitude, r, is added vectorially to the ideal signal. Angle $(\bar{\phi} - \theta)$ is expressed in UI (or cycle) and is cumulative (not modulo 2π). Periodic error is only approximately sinusoidal with position. It is rich in second harmonic, confounding any potential second order perturbation, which has exactly the same frequency.

is rejected. The existing magnitude and phase V and θ are retained for further compensation. This criterion to reject measurement also prevents inappropriate measurements at certain narrow velocity bands where aliasing causes the 2×2 matrix to be singular. For example, at a velocity where phase sampling at 3.2 μs (312.5 kHz) coincides with near one fringe ($\lambda/2$) of travel, measurement is also automatically rejected.

3.3. Periodic error compensation

The two parameters V and θ are used to compensate the next 320 points of raw data ϕ_j to produce the "ideal" data estimate $\bar{\phi}_j$ as shown in Eq. (8). This simple one term compensation removes both the first and second harmonics of first order error. The derivation for Eq. (8) is provided in the following paragraphs.

$$\bar{\phi}_j = \phi_j + V \sin(2\pi(\phi_j - \theta)) \tag{8}$$

The SSB nature of the perturbation is taken into consideration during compensation. Fig. 2 graphically illustrates the relationship between the first order periodic error, $\Delta s = (\lambda/2)\Delta\phi$, and the ideal position of the object, $\bar{s} = (\lambda/2)(\bar{\phi} - \theta)$. It is more convenient to handle periodic error and displacement in the angular forms $\Delta\phi$ and $\bar{\phi}$. These angle parameters are expressed in UI (cycle), not radians, and are cumulative, i.e., not modulo 2π. Fig. 2 shows a periodic error $\Delta\phi$ is produced when a perturbing vector of fractional size r is vectorially added to the unity signal vector. The angle $(\bar{\phi} - \theta)$ of the perturbing vector is the ideal position of the moving target expressed in cumulative UI. Simple trigonometry shows that the relationship is:

$$\Delta\phi(\bar{\phi}) = \left(\frac{1}{2\pi}\right) \tan^{-1} \left[\frac{r \sin(-2\pi(\bar{\phi} - \theta))}{1 + r \cos(-2\pi(\bar{\phi} - \theta))} \right]. \tag{9}$$

For small and moderate $\Delta\phi$, some approximations can be made. Eq. (9) becomes:

$$\Delta\phi(\bar{\phi}) \approx \frac{r/2\pi \sin(-2\pi(\bar{\phi} - \theta))}{1 + 2\pi(r/2\pi)\cos(-2\pi(\bar{\phi} - \theta))} = \frac{V \sin(-2\pi(\bar{\phi} - \theta))}{1 + 2V\pi \cos(-2\pi(\bar{\phi} - \theta))}$$

or

$$\Delta\phi(\bar{\phi}) \approx -V \sin(-2\pi(\bar{\phi} - \theta)) + V^2\pi \sin(-4\pi(\bar{\phi} - \theta)), \tag{10}$$

where $V = r/2\pi$ and θ are the magnitude and phase offset of the periodic error, both expressed in UI. Eq. (10) shows a potentially

significant second harmonic component to the periodic error. It is independent of the presence of any second order periodic error, which has the same frequency. The second harmonic is included in compensation.

For compensation convenience, the periodic error $\Delta\phi$ should be expressed as a function of the readily available uncompensated phase ϕ_j, rather than the ideal phase $\bar{\phi}$ (recall that the phase is determined from position by dividing by $\lambda/2$). From calculus, we have the relationship:

$$\Delta\phi(x - \Delta x) \approx \Delta\phi(x) - \Delta'\phi(x) \cdot \Delta x. \tag{11}$$

Substituting ϕ_j for x and $(\phi_j - \bar{\phi})$ for Δx, the calculus equation becomes $(\phi_j - \bar{\phi}) \equiv \Delta\phi(\bar{\phi}) \approx \Delta\phi(\phi_j) - \Delta'\phi(\phi_j) \cdot (\phi_j - \bar{\phi})$. Solving for $(\phi_j - \bar{\phi}) = \Delta\phi(\phi_j)/1 + \Delta'\phi(\phi_j)$, the compensation $(\phi_j - \bar{\phi})$ becomes a function of the uncompensated position ϕ_j and not the ideal position $\bar{\phi}$, as desired. If only one dominant term $-V \sin(2\pi(\phi_j - \theta))$ is differentiated, and all harmonics higher than two are ignored, the compensation simplifies to $(\phi_j - \bar{\phi}) = -V \sin(2\pi(\phi_j - \theta))$. The compensation has only one sine term with a spatial period of $\lambda/2$ for first order error. Its argument, however, is perturbed by periodic error in such a way that it almost exactly compensates for the SSB perturbation of the first order periodic error, including its second harmonic.

3.4. Hardware implementation

In hardware, the vectors **I**, **J**, **K**, **U**, **L**, **Q** are most conveniently implemented with six shift registers with lengths of 320 to appear sequentially for 320 clock cycles. Because they are well formulated, they may also be synthesized more elegantly with sequential logic, using less hardware. Vectors **C** and **S** are created by addressing cosine and sine tables, respectively, with the most significant 12 bits of the fractional part of input ϕ_j. In vectors **E** and **D**, the boundaries of elements e and d occur at exact octant values of ϕ_j. Therefore the most significant 3 bits of the fractional part of ϕ_j determine the (ternary) values of e or d. The logic design is straightforward and is omitted here.

With all the vectors in place, the 21 dot products shown in Eq. (7) can be "computed" simultaneously using 21 accumulators. In general, the **O** matrix vectors (**U**, **L**, **Q**, **E**, **D**) control the Clock-Enable and Polarity terminals while the **M** matrix elements appear at the Data terminals of the accumulators. For example, the dot product $\mathbf{E}^T\mathbf{C}$ is constructed as follows. The 12 most significant bits of the input phase ϕ_j address a cosine table and the result appears on the Data terminal of an initially empty accumulator (labeled EC, for example). The most significant 3 bits of the fractional part of ϕ_j determine the value of e. If $e = 0$, the accumulator clock is disabled. If $e = 1$, then the Polarity is set to "add", and if $e = -1$, the Polarity is set to "subtract". After 320 clocks, the content of EC contains the dot-product $\mathbf{E}^T\mathbf{C}$. The other 20 dot products are similarly synthesized. After 320 clock cycles, all elements of $\mathbf{O}^T\mathbf{M}$ and $\mathbf{O}^T\mathbf{P}$ are obtained without any computation. For the "−2" value in the **Q** operator, a one bit left shift of the data is required.

In compensation, the term $V \sin 2\pi(\phi_j - \theta)$ is produced by subtracting θ from 320 incoming ϕ_j values. The differences address a sine table and multiply each result by V. These compensation terms are added to the raw ϕ_j to produce the compensated sequence $\bar{\phi}$.

4. Experimental setup

A photograph and schematic of the setup are provided in Fig. 3. The orthogonal, linearly polarized beams with a frequency difference of approximately 3.65 MHz (helium–neon laser source with a

---

apply a safety margin by reducing this value to 64 (1.25 fringes). The minimum measurement velocity for these tests is 1.25 fringes, or 1.25(633)/2 nm, per 1.024 ms (0.39 mm/s).
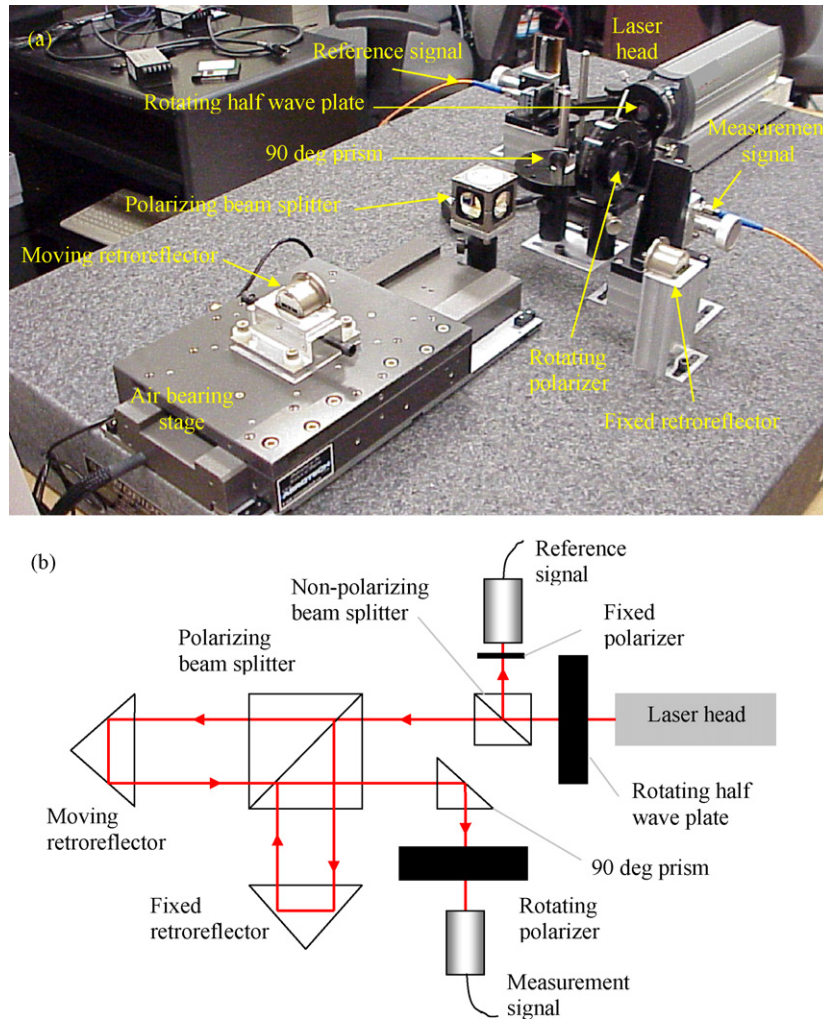
Fig. 3. (a) Photograph of single pass, heterodyne interferometer experimental setup. (b) Schematic of setup.

Zeeman split) first pass through a half wave plate (HWP). Rotation of the half wave plate enables variation in the apparent angular alignment (about the beam axis) between the polarization axes and polarizing beam splitter; deviations in this alignment lead to frequency mixing in the interferometer. The light is then incident on a non-polarizing beam splitter (80% transmission) that directs a portion of the beam to a fiber optic pickup after passing through a fixed angle sheet polarizer (oriented at 45° to the nominal laser orthogonal polarizations). The pickup is mounted on a two rotational degree-of-freedom flexure, which enables efficient coupling of the light into the multi-mode fiber optic. This signal is used as the phase reference in the measurement electronics.

The remainder of the light continues to the polarizing beam splitter where it is (ideally) separated into its two frequency components that travel separately to the moving and fixed retroreflectors. In this design, motion of the moving retroreflector is achieved using an air bearing stage and linear (direct) drive.[4] After the beams are recombined in the polarizing beam splitter, they are directed by a

90° prism through a polarizer with a variable rotation angle. Finally, the light is launched into a fiber optic pickup. This serves as the measurement signal in the measurement electronics (0.3 nm resolution for the single pass configuration).

The intent of the setup design was to minimize other well-known error contributors [22–25] and set various periodic error magnitudes for each order. To isolate periodic error, the setup was constructed with zero dead path difference (i.e., the distance between the polarization beam splitter and the moving retroreflector was equal to the distance between the polarization beam splitter and the fixed retroreflector at initialization) and small Abbe offset (25 mm). The measurement time (~100 ms) and motion excursions were kept small to minimize the contribution of air refractive index variations due to the environmental changes [24]. Additionally, careful alignment of the air bearing stage axis with the optical axis resulted in small beam shear.

5. Experimental results

In this section, the analysis procedures are described and experimental results provided for two primary scenarios: (1) constant velocity tests with variable half wave plate and polarizer orientations and (2) non-constant velocity tests with fixed polarizer and half wave plate angles.

---

[4] It was verified that the drive did not contribute frequency content to the stage motion in the same range as the periodic errors. However, if it did, then the measurement-compensation algorithm would attempt to remove any content that matched the expected periodic error form, independent of its source.
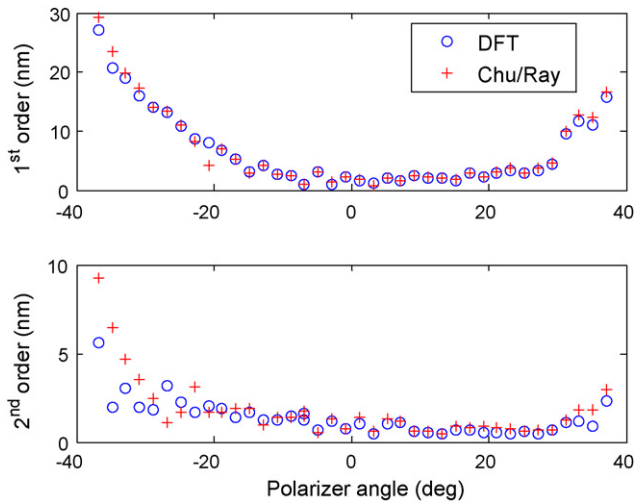
Fig. 4. Error comparison for frequency (DFT) and TDR (modified Chu/Ray) approaches with variable polarizer angle.

Fig. 5. Standard deviation values for each polarizer angle from TDR (modified Chu/Ray) approach (92 periodic error data sets are available for each 96 ms measurement period).

5.1. Variable polarizer angle

In these tests the half wave plate angle was held fixed at its nominal orientation and the polarizer angle was systematically varied about its nominal value (normalized to $0°$ in Fig. 4). The velocity for the air bearing stage (and moving retroreflector) was 100 mm/min. The measurement and reference signal data from the interferometer were sampled at 312.5 kHz by the phase measuring hardware (30×10^3 samples over 96 ms); displacement was then determined from these signals. For the single pass helium–neon interferometer setup used here, first order error repeats approximately every $633/2 = 316.5$ nm, while second order completes a full cycle in $633/4 = 158.3$ nm. To identify the first and second order error amplitudes for the various polarizer angles under constant velocity conditions, two methods were compared. First, the discrete Fourier transform (DFT) of the error was computed [4–6] after performing a least squares linear regression to subtract the best fit line from the displacement data. Second, the TDR algorithm was applied. Note that the DFT approach provides one set of periodic error values for each constant velocity displacement record, while the TDR algorithm provides 92 sets for each record (i.e., the error is updated every 1.024 ms over the 96 ms record, excluding the first 1.024 ms latency period). Both the mean value and standard deviation for each test are reported for the TDR approach.

The first and second order periodic errors for the two methods with polarizer angle variation are displayed in Fig. 4. The top panel shows the first order error comparison, while the bottom panel displays the second order error. The dramatic increase in first order error with polarizer angle misalignment agrees with the behavior modeled, for example, by Cosijns et al. [14]. The additional increase in second order error indicates other imperfections in the system alignment were also present (such as the nominal half wave plate orientation). Fig. 5 shows the standard deviation, σ, in the 92 first and second order periodic error values from the TDR algorithm for each polarization angle. It is observed that the standard deviations for the dominant first order error are essentially independent of the polarizer angle, while the second order error standard deviation values follow the trends in first and second order growth at extreme misalignments. This is explored further in Fig. 11, where it is shown that second order measurement accuracy is improved if first order error is first compensated.
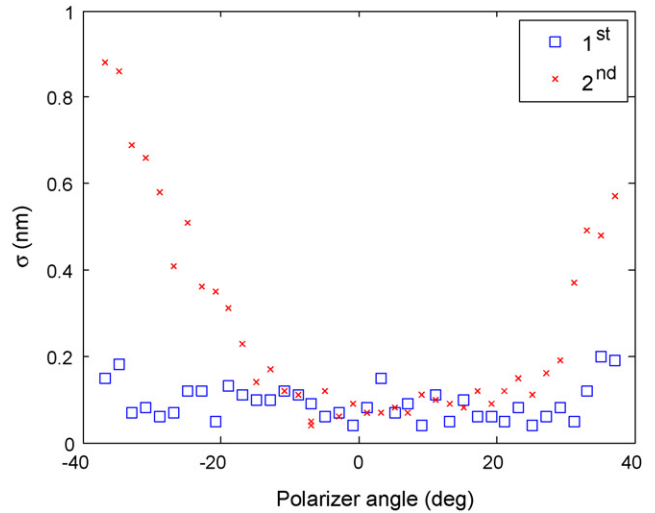
5.2. Variable half wave plate angle

In this case the half wave plate angle was varied while the polarizer angle was held fixed at its nominal orientation. The stage velocity was again 100 mm/min. The first and second order periodic errors identified by the two methods with half wave plate angle variation are displayed in Fig. 6. The reader may note that a θ half wave plate angle change results in a 2θ linear polarization rotation. Again, the sharp increase in second order error with half wave plate misalignment agrees with the Cosijns et al. analysis [14] and the corresponding first order error growth suggests other setup misalignments. Fig. 7 shows the standard deviation in the 92 TDR periodic error values for each half wave plate angle. In this dominant second order error case, both standard deviation profiles track the periodic error growth.

5.3. Constant velocity comparison

To aid in comparing the frequency and time domain analysis results, the difference between the first and second order periodic
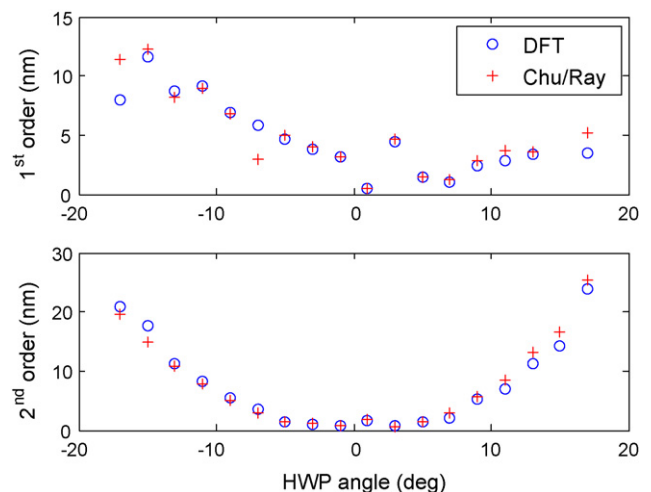

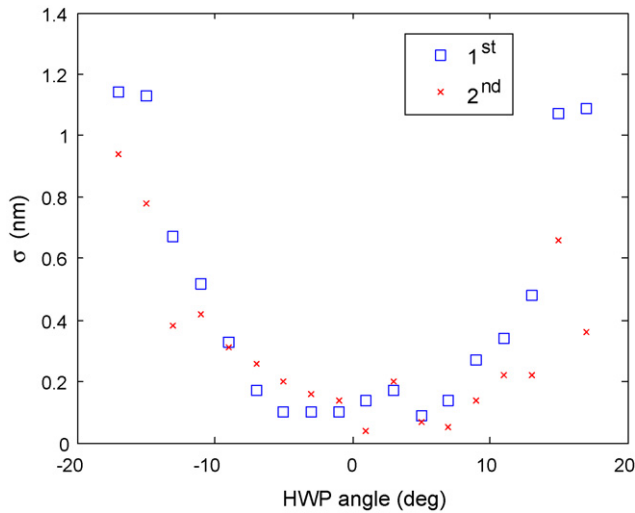**Fig. 6.** Error comparison for frequency (DFT) and TDR (modified Chu/Ray) approaches with variable HWP angle.

Fig. 7. Standard deviation values for each half wave plate angle from TDR (modified Chu/Ray) approach.

Fig. 9. Difference between Fourier and TDR (modified Chu/Ray) periodic error levels for realistic misalignment ranges.

errors for the variable polarizer and half wave plate angle tests was calculated. The results are shown in Fig. 8. It is seen that larger deviations occur for the extreme misalignment conditions. To better represent actual setups (which would not generally suffer from such large misalignments), Fig. 9 presents results for polarizer misalignment within 10° of the nominal and half wave plate misalignment within 5° of nominal (±10° linear polarization misalignment). It is observed that the agreement between the two methods is within the interferometer resolution (0.3 nm) and no obvious bias exists (i.e., the first and second order error differences are both positive and negative). Additionally, from Figs. 5 and 7, it is seen that there is no trend in the σ values, which are at the 0.2 nm level or below, over these ranges.

5.4. Non-constant velocity

Here, data was collected during non-constant velocity profiles. In this case, the frequency domain approach cannot be applied because the periodic error measurement step (Fourier transform of time domain motion data) requires signals with constant time
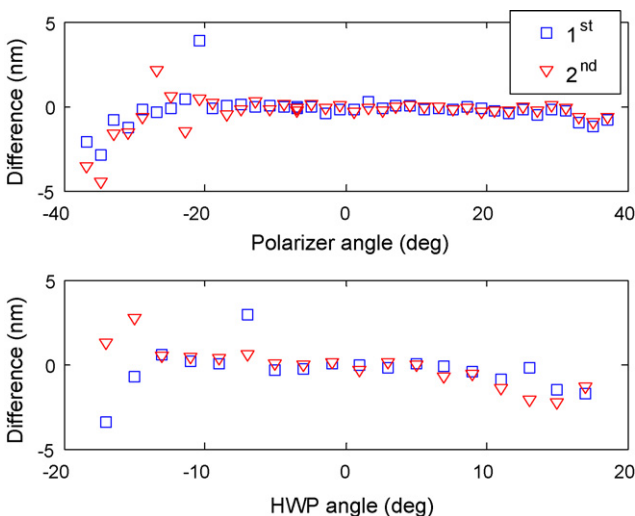
periods, which is not true for non-constant velocity motions (i.e., although the periodic error spatial period is constant, the time period is not).

An example motion profile for non-constant velocity conditions is provided in Fig. 10. The top panel shows the overall motion, x, for the 30×10^3 samples, while the bottom panel shows the velocity, v (obtained from numerical differentiation and polynomial smoothing). Note the direction reversal, which occurs near 41 ms. Fig. 11 shows the displacement (top panel) and periodic error (bottom panel), which was visually isolated by subtracting a least squares fit polynomial, x_{fit}, for the 32–50 ms range about the motion reversal (the low frequency mean drift is due to an imperfect polynomial fit). The variation in the periodic error time period with velocity is observed.

Fig. 12 shows the first and second order periodic error levels which were measured using the time domain algorithm, again for the 32–50 ms time span. As noted, a new value is available every 1.024 ms, although updates cannot be calculated for very low velocities (as described previously). The value remains constant if no update is available (38–41 ms). Fig. 12 shows traces for: (1) first order error measurement from the uncompensated signal (large

Fig. 8. Difference between Fourier and TDR (modified Chu/Ray) periodic error levels for (top) variable polarizer angle and (bottom) HWP angle.
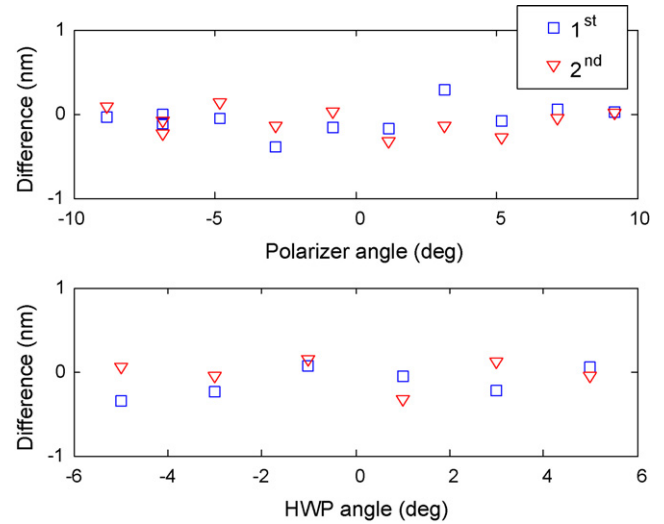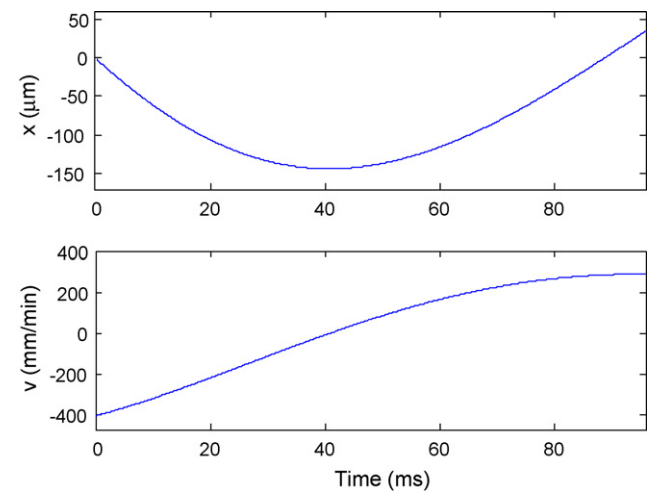

Fig. 10. Displacement (top) and velocity (bottom) profiles for non-constant velocity test.
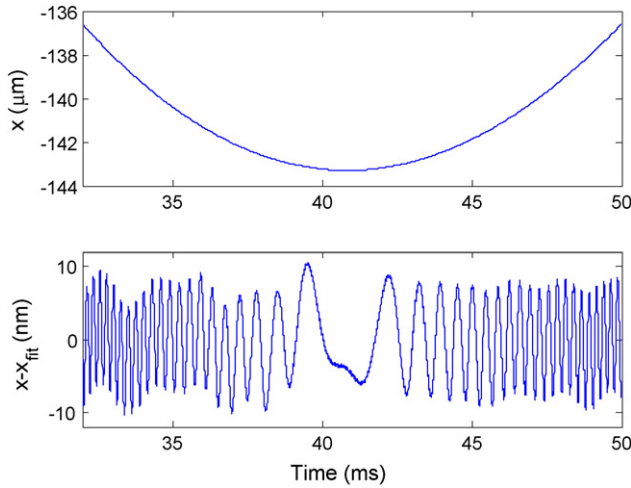
Fig. 11. Displacement (top) and periodic error (bottom) near motion reversal.

Fig. 13. Results of compensation using the measured periodic error shown in Fig. 12.

Fig. 14. Measured periodic error magnitudes for signal with small first order error (1.6 nm) and small second order error (0.9 nm). The second harmonic of the first order error is negligible and does not interfere with the second order error measurement. The results for second order measurement without removing the first order error (filled circles) are identical to the second order measurement when compensating the first order error (unfilled circles).

filled diamond); (2) the compensation result after removing first order error only (unfilled diamond); (3) the second harmonic of the first order error (small filled diamond); (4) second order error measurement using the uncompensated signal (filled circle); (5) second order error measurement using the first order-compensated signal (unfilled circle). It is seen that accurate second order measurement requires that the first order error be removed by compensation when the second harmonic of the first order error is similar in magnitude to the second order error. Using the measured errors (second order error from first order-compensated data in this case), the periodic error waveform can be generated. Given the error waveform, the position can be compensated by subtracting the waveform from the position signal. Fig. 13 shows the uncompensated (dotted line) and first and second order-compensated (solid line) positions. Again, a least squares polynomial fit was used to visually isolate the uncompensated and compensated periodic errors and the low frequency drift is caused by the imperfect fitting for display. The source of the low frequency content could generically be vibration or refractive index variation, for example. In any case, the algorithm only removes what it considers periodic error (in accordance with the curve fit model). Other errors are not compensated and appear as residuals. Additionally, for very large periodic errors, the removal algorithm is only approximate in that it ignores all harmonics above two (in each order).
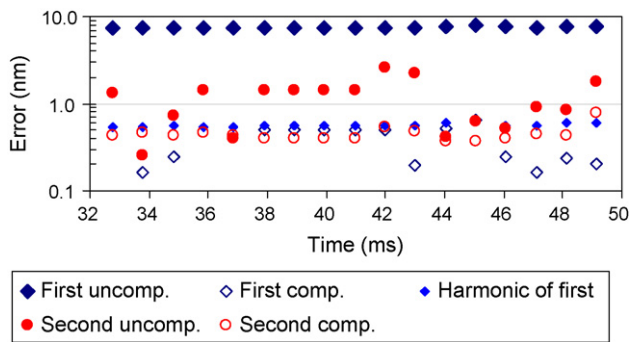
5.5. Compensation examples

Fig. 14 shows measurement results for a constant velocity signal with small first and second order error magnitudes (1.6 and 0.9 nm, respectively). Fig. 15 displays the time domain compensation

Fig. 12. Measured periodic error magnitudes for Fig. 11 signal with large first order error (7.5 nm) and small second order error (0.4 nm). The second harmonic of the first order error (0.6 nm) is larger than the second order error. An attempt to measure the second order error without first removing the first order error results in erratic and inaccurate results (filled circles). Improved results are obtained if the first order error is first compensated (unfilled circles). During direction reversal (38–41 ms), no parameters are updated.
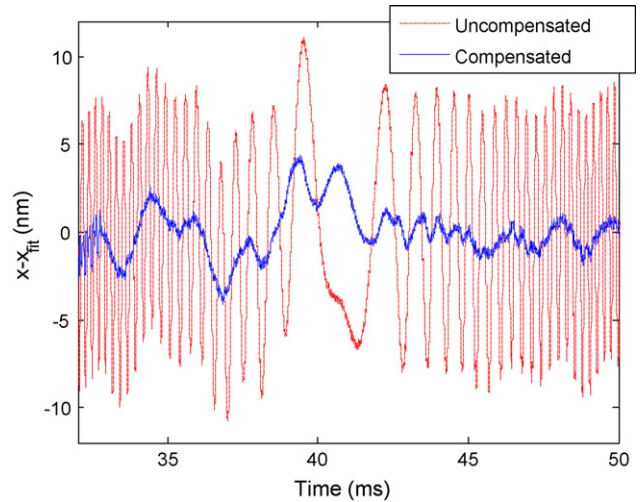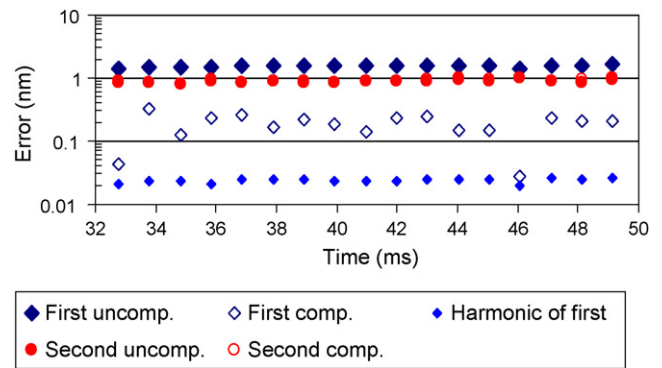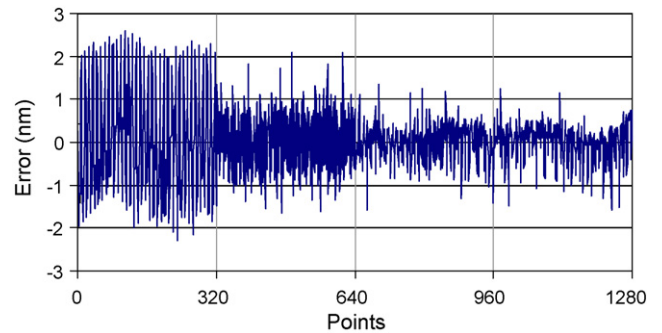

Fig. 15. The effect of periodic error compensation in the time domain is demonstrated for the small first and second order error signal. Four segments of 1.024 ms (320 points each) are shown. Due to latency, the first segment is uncompensated. The second segment is only first order error-compensated. The subsequent segments are compensated in both orders.
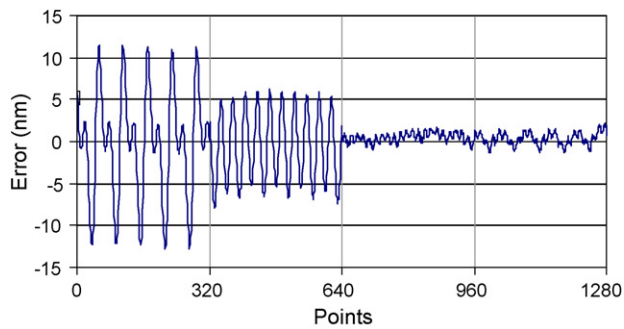
Fig. 16. Example of effect of compensation algorithm on a signal with large periodic error in both first and second orders (6.8 and 5.9 nm, respectively). In the first 1.024 ms segment (320 points) no compensation is applied. In the second segment, first order only is compensated. In the remaining segments both orders are compensated.

result for Fig. 14 signal. The effectiveness is observed in spite of noise present. Fig. 16 shows the effect of compensation in the time domain for a constant velocity signal with large first and second order errors (6.8 and 5.9 nm, respectively). Significant error reduction is again seen.

Although the experimental results presented here were for a single pass system, the approach can also be implemented on multiple pass configurations. In fact, unwanted reflections in a two pass system can lead to periodic error with half the spatial frequency of first order error. This "half order error" can be removed by an additional stage, which precedes the first order error compensation (as shown in Fig. 1).

6. Conclusions

The tests reported in this study validated the periodic error levels computed using an extension of the digital, time domain Chu and Ray algorithm [1] (referred to here as the time domain regression, or TDR, algorithm) by comparison with the well-known frequency domain approach [4–6]. Further, it was demonstrated that the TDR approach was capable of continuous estimates of the first and second order periodic error for non-constant velocity motion. The ability to diagnose periodic error at non-constant velocities and reconstruct the wave shape for error compensation, without modifications to the optical setup, makes the algorithm particularly well-suited to real-time implementation for existing precision motion systems.

References

[1] Chu D, Ray A. Nonlinearity measurement and correction of metrology data from an interferometer system. In: Proceedings of the 4th euspen international conference. 2004. p. 300–1.
[2] Schmitz T, Chu D, Houck III L. First order periodic error correction: validation for constant and non-constant velocities with variable error magnitudes. Measurement Science and Technology 2006;17:3195–203.
[3] Schmitz T, Houck III L, Chu D, Kalem L. Bench-top setup for validation of real time, digital periodic error correction. Precision Engineering 2006;30:306–13.
[4] Patterson S, Beckwith J. Reduction of systematic errors in heterodyne interferometric displacement measurement. In: Proceedings of the 8th international precision engineering seminar (IPES). 1995. p. 101–4.
[5] Badami V, Patterson S. A frequency domain method for the measurement of nonlinearity in heterodyne interferometry. Precision Engineering 2000;24(1):41–9.
[6] 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. 1997. p. 153–6.
[7] Fedotova G. Analysis of the measurement error of the parameters of mechanical vibrations. Measurement Techniques 1980;23(7):577–80.
[8] Quenelle R. Nonlinearity in interferometric measurements. Hewlett–Packard Journal 1983;34(4):10.
[9] Sutton C. Nonlinearity in length measurements using heterodyne laser Michelson interferometry. Journal of Physics E: Scientific Instrumentation 1987;20:1290–2.
[10] Barash V, Fedotova G. Heterodyne interferometer to measure vibration parameters. Measurement Techniques 1984;27(7):50–1.
[11] Bobroff N. Residual errors in laser interferometry from air turbulence and nonlinearity. Applied Optics 1987;26(13):2676–82.
[12] Rosenbluth A, Bobroff N. Optical sources of nonlinearity in heterodyne interferometers. Precision Engineering 1990;12(1):7–11.
[13] Stone J, Howard L. A simple technique for observing periodic nonlinearities in Michelson interferometers. Precision Engineering 1998;22(4):220–32.
[14] Cosijns S, Haitjema H, Schellekens P. Modeling and verifying nonlinearities in heterodyne displacement interferometry. Precision Engineering 2002;26:448–55.
[15] Wu C, Deslattes R. Analytical modeling of the periodic nonlinearity in heterodyne interferometry. Applied Optics 1998;37(28):6696–700.
[16] Wu C, Su C. Nonlinearity in measurements of length by optical interferometry. Measurement Science and Technology 1996;7:62–8.
[17] Hou W, Wilkening G. Investigation and compensation of the nonlinearity of heterodyne interferometers. Precision Engineering 1992;14(2):91–8.
[18] Hou W, Zhao X. Drift of nonlinearity in the heterodyne interferometer. Precision Engineering 1994;16(1):25–35.
[19] Howard L, Stone J. Computer modeling of heterodyne interferometer errors. Precision Engineering 1995;12(1):143–6.
[20] Tanaka M, Yamagami T, Nakayama K. Linear interpolation of periodic error in a heterodyne laser interferometer at subnanometer levels. IEEE Transactions on Instrumentation and Measurement 1989;38(2):552–4.
[21] Wu C, Lawall J, Deslattes R. Heterodyne interferometer with subatomic periodic nonlinearity. Applied Optics 1999;38(19):4089–94.
[22] Schmitz T, Evans C, Davies A, Estler WT. Displacement uncertainty in interferometric radius measurements. Annals of the CIRP 2002;51(1):451–4.
[23] Schmitz T, Kim HS. Monte Carlo evaluation of periodic error uncertainty. Precision Engineering 2007;31(3):251–9.
[24] Bobroff N. Recent advances in displacement measuring interferometry. Measurement Science and Technology 1993;4:907–26.
[25] Steinmetz C. Sub-micron position measurement and control on precision machine tools with laser interferometry. Precision Engineering 1990;12(1):12–24.
[26] Cretin B, Xie W, Wang S, Hauden D. Heterodyne interferometers: practical limitations and improvements. Optics Communications 1988;65(3):157–62.
[27] Petru F, Cip O. Problems regarding linearity of data of a laser interferometer with a single-frequency laser. Precision Engineering 1999;23(1):39–50.
[28] Augustyn W, Davis P. An analysis of polarization mixing errors in distance measuring interferometers. Journal of Vacuum Science and Technology B 1990;8(6):2032–6.
[29] Xie Y, Yu Y. Zeeman laser interferometer errors for high precision measurements. Applied Optics 1992;31(7):881–4.
[30] De Freitas J, Player M. Importance of rotational beam alignment in the generation of second harmonic errors in laser heterodyne interferometry. Measurement Science and Technology 1993;4:1173–6.
[31] De Freitas J, Player M. Polarization effects in heterodyne interferometry. Journal of Modern Optics 1995;42(9):1875–99.
[32] De Freitas J. Analysis of laser source birefringence and dichroism on nonlinearity in heterodyne interferometry. Measurement Science and Technology 1997;8:1356–9.
[33] Li B, Liang J. Effects of polarization mixing on the dual-wavelength heterodyne interferometer. Applied Optics 1997;36(16):3668–72.
[34] Park B, Eom T, Chung M. Polarization properties of cube-corner retroreflectors and their effects on signal strength and nonlinearity in heterodyne interferometers. Applied Optics 1996;35(22):4372–80.
[35] Oldham N, Kramar J, Hetrick P, Teague E. Electronic limitations in phase meters for heterodyne interferometry. Precision Engineering 1993;15(3):173–9.
[36] Schmitz T, Beckwith J. An investigation of two unexplored periodic error sources in differential-path interferometry. Precision Engineering 2002;27(3):311–22.