2013 American Control Conference (ACC)
Washington, DC, USA, June 17-19, 2013

A Tutorial on Laser Interferometry for Precision Measurements

Russell Loughridge*

Daniel Y. Abramovitch**

Abstract— **Laser interferometers have found wide usage in a variety of precision measurement applications. The ability to gain precise position information with minimal change to the dynamics of the device being measured has a large set of advantages. This allows interferometer systems to be used in feedback loops for precision systems. This paper presents a tutorial on laser interferometers, their use in precision motion feedback systems, the issues faced by such systems, and some of the solutions that have been applied to these issues.**

I. Introduction

Michelson type laser interferometers measure distance by measuring the phase difference between two portions of the same beam, one sent to a reflector at a fixed distance, and one sent to a measurement surface at an unknown distance. When the two signals are recombined in the interferometer, the resulting phase is related to the distance of the reflected surface from the interferometer. As the distance changes, so does the phase of the combined signal. The utility of these methods are that the measurement can be made over long distances while maintaining accuracy. However, as the needed accuracy of the target applications has increased, interferometers have been adjusted to desensitize them to an increasing number of effects.

Most engineers hear about optical interference in college physics and promptly forget about them. However, for the group of engineers and scientists that keep using them, optical interference provides highly precise and yet remote position measurements. The ability to resolve small features, down the the fraction of the wavelength of light (390 to 700 nm [1]), as well as the non-contact nature of the measurements are inherent advantages. Because of this interferometric measurements are used in fields that go from astronomy to oceanography, from chemistry and physics to laser tape measures and laser mice for computers [2].

For this tutorial, we will restrict our discussion to laser interferometers (rather than so called white light interferometers). Furthermore, the interferometers (IFs) we discuss will largely focus on those taking the form of the Michelson interferometer [3], [4], [5], but using a laser light source. While the original Michelson interferometer is not a practical design for reasons that we will discuss, it provides a clean conceptual model with a very understandable idealized behavior. Our approach will be to describe the idealized behavior of the Michelson in detail, and then show how practical issues arise which cause design changes. Still, the net effect of these design changes is to cancel out the non-ideal behavior, so as to restore the instrument's behavior back to the original model.

From a simplistic control design view, an interferometer is simply a complicated, expensive sensor, that provides highly accurate, non-contact measurements of a certain number of axes at a high rate for a certain cost. However, we believe that understanding how precision interferometers operate will give control designers much better insight into how to use these devices and how to configure them for maximum utility.

The rest of this paper will proceed as follows: Section II will introduce the Michelson Interferometer and derive the idealized behavior from the electromagnetic wave equations [6]. Factors that make the original Michelson interferometer impractical and fixes to them are discussed in Section III. In Section IV, we introduce two frequency, or heterodyne interferometry. Finally, in Section V, we show how the interferometer fringes on the optical detectors are turned into position measurements.

With the basic concepts explained, we describe the history of and uses of precision interferometers in Sections VI and VII. We then return to issues with modern interferometers in Section VIII. and cover their use in closed loop servo systems in Section IX. We close with a discussion of turbulence in Section X, one of the more difficult and pervasive problems with interferometry. This will tie in well with the last paper in the tutorial session which will discuss a closed-loop method of combating turbulence using a multi-segment detector and an Extended Kalman Filter [7].

II. The Michelson Interferometer

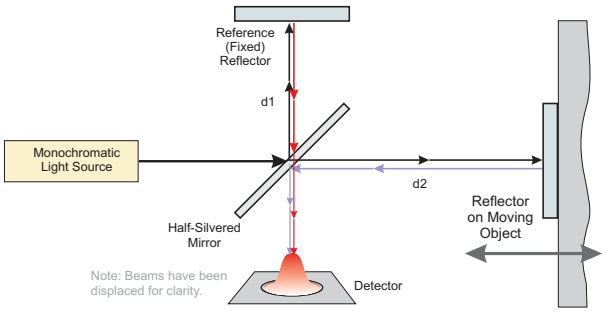

Fig. 1. Basic Michelson Interferometer

The basic Michelson interferometer (Figure 1) uses a half silvered mirror to split a monochromatic light source into two

*Russell Loughridge is with the Agilent Technologies Nano Position and Measurement Division, Cleveland, OH 44060 USA, (847) 944-6262; russell_loughridge@agilent.com)

**Daniel Y. Abramovitch is a principal project engineer in the Molecular Imaging Lab at Agilent Laboratories, 5301 Stevens Creek Blvd., M/S: 4L-IC, Santa Clara, CA 95051 USA, danny@agilent.com

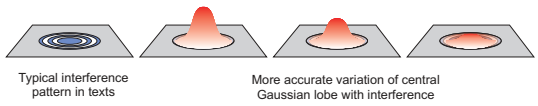

Fig. 2. Effects of interference on detector. On the far left is the typical diagram one sees in books. However, the banding is typically caused by the beam being cropped and not the effect of the interferometry. A better picture comes from the three diagrams on the right, in which the intensity of the central Gaussian lobe is modulated by the interference pattern.

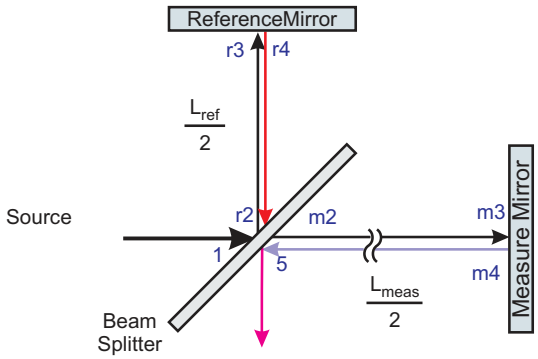

Fig. 3. Some details on the beams of a Michelson interferometer

beams. Each beam reflects off of a mirror, to be recombined at the half-silvered mirror. The recombined beam contains an interference pattern that changes when either of the mirrors move. Keeping one mirror fixed allows one to attribute all of the interference pattern changes to motion of the other mirror.

Many texts show the an interference pattern such as the one in the far left of Figure 2. However, in the absence of the beam being cropped, the detector will see a collimated beam on its center axis. Typically, this is modeled as a Gaussian beam and as the measure mirror moves the intensity of this pulse will vary, as shown in the right three figures of Figure 2. The detector then acts to integrate the intensity of the beam over its spatial extent, and – assuming the integration is faster than the change in the interference pattern – this integrated intensity can be used to measure distance, modulo the wavelength of the laser used.

The Michelson interferometer is one of the most basic models of interferometry available. It is not a practical interferometer, in that there are significant issues with the actual implementation. However, it provides an easy to understand conceptual model for understanding precision measurement interferometers. What we will find is that for every imperfection of the Michelson interferometer, there is a practical fix that expands the range of usefulness of the interferometer. In the course of this tutorial we will go through a "problem-fix" approach. Each of these fixes essentially returns the interferometer back to a more ideal Michelson behavior.

We start our analysis of the Michelson IF equations by looking at Figure 3. For our purposes, the source beam can be considered to originate at position 1, right before contact with the half silvered mirror. At the mirror, half of the beam is reflected to the reference mirror (path r2-r3-r4) where it is reflected back towards the half silvered mirror. At this interface, half of the beam is passed through to position 5, while half reflects back to the source. Meanwhile, the transmitted portion of the beam goes to the measurement mirror (path m2-m3-m4) and reflects back. At the half silvered mirror, half of the measure beam is reflected to position 5, while half passes back to the source. We are concerned with the two beams that meet at position 5 and are imaged on the detector.

A few things are important to understand interference as it is used in our measurements. First, since both the reference beam and the measure beam originate from the same laser, they are coherent with each other. Second, every time a beam goes through a reflection, it undergoes a 180° phase shift. A look at the diagram of Figure 3 indicates that each beam at position 5 has gone through 360° in phase shifts and thus they are still in phase with each other. Third, by the time both beams reach position 5, their amplitude has been reduced to $\frac{1}{4}$ of their original amplitude. If the reflection/transmission is exactly 50/50 and if the mirrors are perfectly aligned, then both beams add through linear superposition and have the same amplitude. Thus, we can attribute the variation at the detector to interference.

The equations for the interference pattern are derived in classic optical texts [5], [8] from application of the vector electromagnetic wave equations [6], [9]. Consider the electric field of the source beam at position 1:

$$E_{z,source}(z,t) = A\cos(kz - \omega t + \phi) \tag{1}$$

where z is the direction of travel, $k = \frac{2\pi}{\lambda}$ is the wave number, λ is the wavelength of the light, and A is the amplitude of the beam. From position 1, the reference beam travels a distance $L_{ref} = 2d_1$ to get back to position 5, while the measure beam travels a distance $L_{meas} = 2d_2$ to get back to position 5. If we consider position 1 to be $z = 0$, then the two beams are thus,

$$E_{ref}(t) = \frac{A}{4}\cos(kL_{ref} - \omega t + \phi) \tag{2}$$

for the reference beam and

$$E_{meas}(t) = \frac{A}{4}\cos(kL_{meas} - \omega t + \phi) \tag{3}$$

for the measure beam. Through linear superposition, the beams add, so that the electric field of the combined beams at position 5 is

$$\begin{aligned} E_{tot}(t) \quad = \quad & \tfrac{A}{4}\left[\cos(kL_{ref} - \omega t + \phi)\right. \\ & \left. + \cos(kL_{meas} - \omega t + \phi)\right]. \end{aligned} \tag{4}$$

A word about notation is useful here. Normally, when one is working with wave equations [6], [9], the equations are set up as vector quantities along some frame of reference. This works very well in analysis of plane equations, point sources, etc. but in our interferometer, the direction of the beams are

switched so often that keeping track of all the vector frames becomes confusing. For this tutorial, we will assume that the source electric field is in the X-Y plane, and the source magnetic field is rotated 90° in that plane. This means that the Poynting vector which describes the energy density is in the Z direction. Every time we go through a reflection, polarizer, or beam splitter, the reference frame is changed, but our signals will end up so that the Poynting vector is normal to the detector plane. For the sake of simplicity, we will leave off the unit vector designations on the equations.

The detector is sensitive to signal intensity, not amplitude, and we can calculate this from the Poynting vector. If we assume that the electric field is in the x direction and the magnetic field is in the y direction, then

$$\begin{aligned} H_{tot}(t) &= \tfrac{A}{4}\sqrt{\tfrac{\epsilon}{\mu}}\left[\cos(kL_{ref}-\omega t+\phi)\right. \\ &\quad \left.+\cos(kL_{meas}-\omega t+\phi)\right]. \end{aligned} \tag{5}$$

We now have two choices to simplify this: proceed with trigonometric identities or switch gears to saying that Equations 4 and 5 are the real parts of a complex exponential notation. For pedagogical purposes, we will plug through the trigonometric equations here. With the polarizations we have assumed, the Poynting vector, $\mathcal{P}(t)$, will be in the direction normal to the detector with

$$\begin{aligned} \mathcal{P}_{tot}(t) &= \vec{E}_{tot}(t)\times\vec{H}_{tot}(t) \\ &= \frac{A^2}{16}\sqrt{\frac{\epsilon}{\mu}}\left[\cos\alpha+\cos\beta\right]^2 \qquad (6) \\ &= \frac{A^2}{4}\sqrt{\frac{\epsilon}{\mu}}\left[\cos^2\alpha+\cos^2\beta+2\cos\alpha\cos\beta\right] \qquad (7) \end{aligned}$$

where $\alpha = kL_{ref}-\omega t+\phi$ and $\beta = kL_{meas}-\omega t+\phi$. With this and some trigonometric identities, we end up with

$$\cos^2\alpha = \frac{1+\cos 2(kL_{ref}-\omega t+\phi)}{2}, \tag{8}$$

$$\cos^2\beta = \frac{1+\cos 2(kL_{meas}-\omega t+\phi)}{2}, \text{ and} \tag{9}$$

$$\begin{aligned} 2\cos\alpha\cos\beta &= \cos\left(k(L_{meas}+L_{ref})-2\omega t+2\phi\right) \\ &\quad \times\cos\left(k(L_{meas}-L_{ref})\right). \end{aligned} \tag{10}$$

Putting these all together,

$$\begin{aligned} \mathcal{P}_{tot}(t) &= \frac{A^2}{32}\sqrt{\frac{\epsilon}{\mu}}\left[1+\cos 2(kL_{ref}-\omega t+\phi)\right. \\ &\quad +1+\cos 2(kL_{meas}-\omega t+\phi) \\ &\quad +2\cos\left(k(L_{meas}+L_{ref})-2\omega t+2\phi\right) \\ &\quad \left.+\,2\cos\left(k(L_{meas}-L_{ref})\right)\right] \end{aligned} \tag{11}$$

If we average over an integer number of periods, $T = \frac{1}{f} = \frac{2\pi}{\omega}$ then the time varying portion integrates out, leaving only the DC portion:

$$\mathcal{P}_{tot,avg} = \tfrac{A^2}{16}\sqrt{\tfrac{\epsilon}{\mu}}\left[1+\cos\left(k(L_{meas}-L_{ref})\right)\right] \tag{12}$$

As a practical matter, the laser frequency is so much faster than the integration time of our detector that we are always getting the "DC portion". Thus, the relationship that is most commonly used for this type of interferometer is that for the intensity:

$$I \sim K\left[1+\cos\left(k(L_{meas}-L_{ref})\right)\right] \; W/m^2 \tag{13}$$

This is often rewritten in terms of the wavelength, λ, as

$$I \sim K\left[1+\cos\left(\tfrac{2\pi}{\lambda}\left(L_{meas}-L_{ref}\right)\right)\right] \; W/m^2 \tag{14}$$

This provides the power density at the detector in $Watts/m^2$. The detector integrates the energy density (intensity) over the detector surface. Thus, it is not the pattern on the surface that matters so much as the amount of intensity on that surface. Equation 13 gives the density at a given point. In fact for a highly collimated beam, the distribution is likely Gaussian and effect of a change in L_{meas} is to cause the height of this Gaussian distribution to rise and fall. If all other variables are held constant, one can measure a change in distance by counting the passing of these light and dark times. From Equation 13 we see that we are still missing an ability to discern direction of motion. This is fixed in a single frequency IF by splitting the the beam and adding a phase delay to one portion, thus allowing for in-phase and quadrature demodulation (IQ).

III. Issues Affecting Michelson Interferometers

Michelson interferometers are easy to understand but face some very practical issues. The first is that the half silvered mirror works through amplitude splitting and this causes beam power to be lost. Looking at Figures 1 and 3, we see that both the beam from the reference reflector and the beam from the moving reflector will be split again, resulting in only half of the amplitude of each showing up on the detector. The rest the light will be sent back into the laser, causing interference with the laser.

Furthermore, the Michelson interferometer is very sensitive to mis-alignment and producing an exact 50-50 amplitude split is difficult.

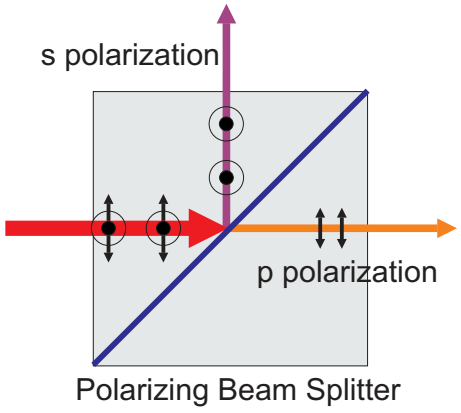

Fig. 4. A polarizing beam splitter (PBS) allows light in one polarization (P) to pass through unhindered while it reflects light in the orthogonal (S) polarization.

To get around this, most Michelson style interferometers work with multiple light polarizations and a polarizing beam

splitter (PBS), diagrammed in Figure 4. A polarizing beam splitter allows light in one polarization (P) to pass through unhindered while it reflects light in the orthogonal (S) polarization. Aligned with a $45°$ angle to the incoming beam, it results in the beam being split into two directions, each of a single linear polarization. The great advantage of using multiple polarizations and a PBS is that unlike the half silvered mirror, there is little loss in optical power.

With the proper control of the polarization states of the source light, one could split it into two components and combine them with minimal loss at the equivalent of position 5 in Figure 3, except that in place of the half silvered mirror is the PBS. The key then is controlling the polarization of each of the split beams so that they are in matching polarizations when they arrive. This is done with a combination of mirrors, cube corners (to be described below), and wave plates.

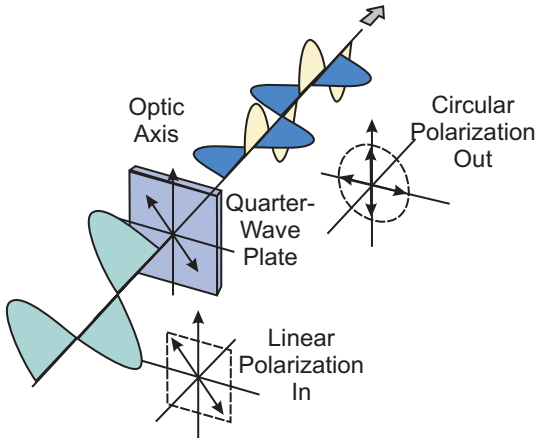

Fig. 5. Wave plates are used to change the polarization state of light. (Recreated from [10].)

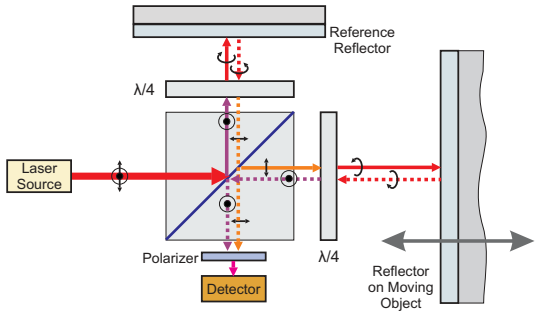

Fig. 6. Interferometer with Polarizing Beam Splitter (PBS) and quarter wave plates.

Wave plates are optical components that slow one of the polarization states of the incident light, relative to the other. They are typically made of birefringent crystal, where the birefringence is characterized by the difference in the indices of refraction of the two axes of the crystal, Δn. If light enters with equal polarization in two axes, one of those axes will have a phase delay given by [11]:

$$\Gamma = \frac{2\pi \Delta n L}{\lambda_0} \tag{15}$$

where λ_0 is the wavelength of the light in vacuum. By choosing the length, L, of the crystal, a chosen phase delay may be imparted. Typical values for L result in a quarter wave plate, which results in a $90°$ phase lag in one of the polarization states (relative to what it was when it entered the wave plate) and a half wave plate, which retards one polarization state by $180°$. As such, the quarter wave plate, often designated in a diagram by $\frac{\lambda}{4}$, can be used to transform linearly polarized light to circularly polarized light and vice-versa. In fact, a typical optical path sequence that is seen in interferometers is that light with S polarization passes through a quarter wave plate (giving it circular polarization), is reflected (reversing the direction of the circular polarization), and passes back through the quarter wave plate (giving it P polarization). These polarization changes allows portions of the beam to either pass directly through a PBS or be diverted by it.

One more needed component is a polarizer, which selects out a particular polarization of light. One can think of any polarization vector as being the resultant of two other vectors. For example, vertical, S, polarization can be viewed as the resultant of equal polarization vectors at $\pm 45°$. Likewise, horizontal, P, polarization can be viewed as the resultant of equal polarization vectors at $45°$ and at $135°$. Sending these two beams through a polarizer that selected out polarizations at $45°$ would result in the $45°$ components of the S and P light being selected out. Polarizers will often be seen right before the detectors in an interferometer configuration to select out the components that can be interfered with each other.

Figure 6 shows an interferometer configuration using a polarized light source, a polarizing beam splitter, quarter wave plates, and plane mirror reflectors. Say we have a source beam with equal components of polarization in S and P. Upon striking the interface of a PBS, the S polarized component will be diverted to reference mirror, while the P polarized component will pass on to the measure mirror. To pass the return beams to a location distinct from the laser source, the beams exiting the PBS are passed through a quarter wave plate resulting in a change of polarization from linear to circular. Reflection off of the plane mirror results in a reversal of the direction of circular polarization and both beams pass through the quarter wave plate again. The net result is that the reference beam returns to the PBS with a P polarization while the measurement beam returns to the PBS with an S polarization. The properties of the PBS cause both of these beams to combine and exit the PBS in a new direction. Both measure and return beams are then passed through a polarizer, which selects out the common linear polarization between the two. This is what the detector sees.

This is all an improvement, but we have not solved the angle sensitivity problem. One solution that can be used when the measure mirror is only moving along a single axis is to use a corner cube or retroreflector. A two dimensional

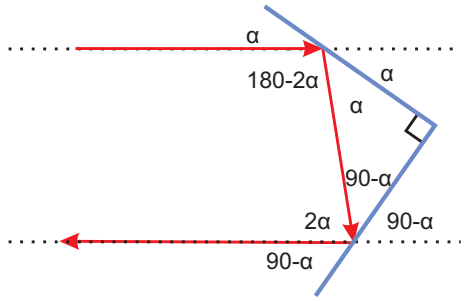

Fig. 7. Planar corner cube (retroreflectors). Actual retroreflectors are built in 3 dimensions, but the concept can be diagrammed well in two.

diagram is shown in Figure 7.

The corner cube/retro-reflector has the unique property that it returns light back along an axis parallel to the axis of the incident beam, independent of the angle of the retroreflector. This can be seen in the simple geometry of Figure 7. Ideally, the mirrors on our two dimensional cube corner would be at $\pm 45°$ from the incoming beam direction. A variation in the alignment of the cube corner would not result in a different direction for the reflected beam, but would result in a beam that was translated a small amount with respect to its original axis. As long as this translation is small relative to the size of the detector, it is not an issue. However, unlike a plane mirror reflector, when the cube corner translates side to side (or up or down), the return beam translates twice as far in the same direction. This characteristic reduces the allowable lateral motion of the reflector to about 1/4 of the beam's diameter (about ± 1.5 mm for a 6 mm beam diameter) so there is always about 50% overlap with the reference beam (the actual required overlap depends on the receiver's sensitivity). For typical precision systems, this straightness of travel requirement is usually not an issue.

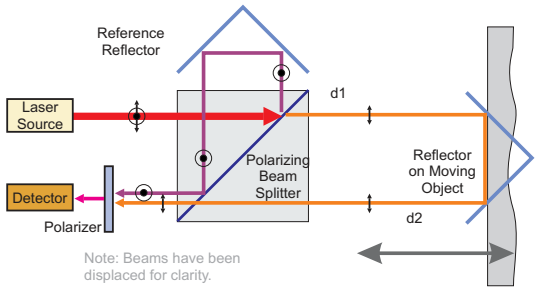

Fig. 8. Interferometer with Polarizing Beam Splitter (PBS) and corner cubes.

The corner cube also allows the reflected beam to be rerouted to a different portion of the PBS. This offset in the beams (known as *Offset Beam Interferometry*) can be used to avoid sending return laser beams back into the laser source. The example in Figure 8 uses two corner cubes to route the beam to a different location so that the detector and the laser source do not have to be coincident. The reference beam with S polarization is reflected to the reference reflector and back to the PBS with the same S polarization. The measure beam with P polarization passes through the PBS and returns with the same polarization and passes through the PBS. In order to combine the two return beams, we need to select out the common polarizations of the two beams with a polarizer.

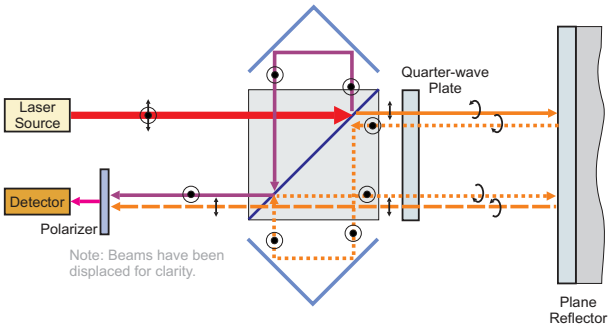

Fig. 9. Laser beam paths in a two-pass plane mirror interferometer (PMI)

As noted, corner cubes are restricted to one direction of movement. To allow the measurement surface to move in multiple directions, a Plane Mirror Interferometer (PMI) must be used. In order to provide angular insensitivity, the measure beam is offset using a corner cube and passed back to the measure mirror again. Thus, to first order, variations due to angles of the measure beam are undone by the second pass.

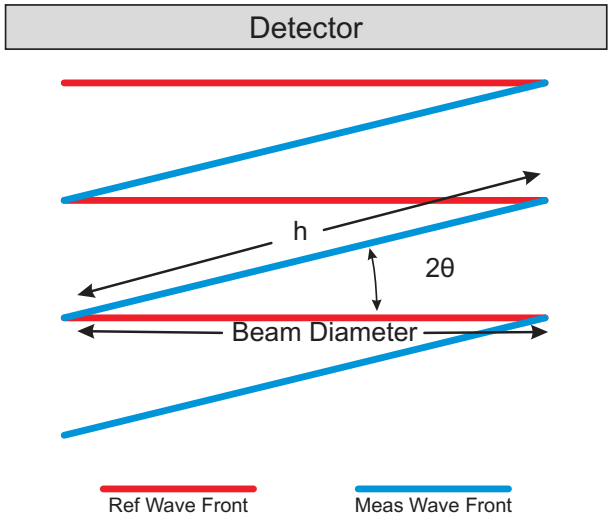

Fig. 10. Angle limit for single pass plane mirror reflector

Introduction of the Plane Mirror Interferometer significantly increased the allowable angular range for the moving mirror. With single pass interferometers, the allowable angular range is limited by Equation 16. The PMI (Figure 9)

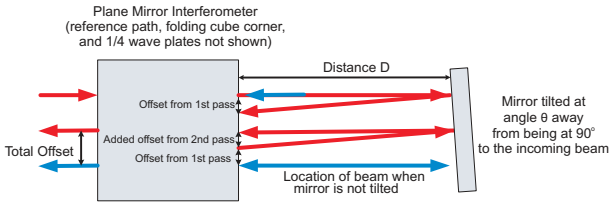

Fig. 11. Angle limit for two pass plane mirror reflector

eliminates this restriction by automatically correcting any wavefront tilt caused by angular movement of the mirror. Any tilt from the first reflection off the mirror is untilted by the second reflection because the cube corner on the interferometer swaps the beam side to side and top to bottom.

$$AngularRange \sim \frac{1}{2}\tan^{-1}\frac{3\lambda}{4 \times BeamDiam} \tag{16}$$

This is derived from the fact that the wavefront tilt must be less than λ over the beam diameter so a wavefront of the measurement beam does not cross more than 1 wavefront of reference beam. Using $\approx \frac{3}{4}\lambda$ is a reasonable limit to set.

Thus:

$$h\sin(2\theta) \leq \frac{3}{4}\lambda \tag{17}$$

$$h\cos(2\theta) = BeamDiam \tag{18}$$

$$h = \frac{BeamDiam}{\cos(2\theta)} \tag{19}$$

$$\tan(2\theta) \leq \frac{3}{4}\frac{\lambda}{BeamDiam} \tag{20}$$

There still is a restriction on the mirror's angular range, but it is much less restrictive than a single pass configuration. For example, with a 6 mm beam, a single pass IF has an angular range of approximately 40 μRadians. With the same beam a PMI would have an angular range of about 7.5 mRadians when the distance to the measure mirror, $L_{meas}/2$ = 100 mm, and about 1.5 mRadians when $L_{meas}/2$ = 500 mm.

$$AngularRange \sim \frac{1}{2}\sin^{-1}\frac{BeamDiam}{4D} \tag{21}$$

This is derived from knowing that beams should overlap 50% to get good signal (note high sensitivity receivers allow operation with less than this, but general rule of thumb is to have at least this much). Drawing out a ray trace diagram, one can see that the final meas path output beam is offset from its original position by $2 \times D \times \sin(2\theta)$, where D is the distance between the mirror and the interferometer and theta is the difference between the mirror's angle relative to the incoming beam and 90 degrees. (yes there are some additional offsets caused by the beam path within the interferometer, so this equation changes for small values of D, but for most situations, D is large and the other term can be ignored). So the requirement is:

$$2 \times D \times \sin(2\theta) \leq \frac{BeamDiam}{2} \tag{22}$$

There are two additional consequences of having the second pass in a PMI. One, the resolution of the measurement is double that of a single pass interferometer, as any motion of the mirror now causes two times the phase shift in the measurement beam. Two, the maximum allowable velocity is cut in half for two frequency interferometers, which is discussed in Section IV and velocity requirements in Section VIII-E below.

IV. Two Frequency Interferometry

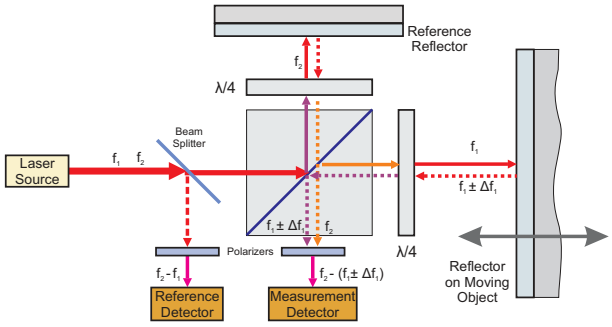

Fig. 12. Two frequency (heterodyne) Michelson interferometer configuration.

All the modifications to the Michelson interferometer discussed thus far essentially are designed to desensitize the interferometer to non-ideal behavior and restore the accuracy of Equation 14. However, even when things are properly aligned, the interferometers described so far operate in the baseband. They use a single frequency of light, also known as homodyne interferometry, and the "difference" between measure and reference only shows up as a baseband phase and Equation 14 is a variation away from DC. DC detection is slow and suffers from $\frac{1}{f}$ and other noise in the detectors, mainly signal intensity variations (due to air turbulence or accumulated contaminants on mirror and optic surfaces) being indistinguishable from position changes.

Borrowing from the world of radio communications, it is more advantageous if the interference shows up at some intermediate frequency. To achieve this, modern IF measurement systems typically operate with multiple wavelengths [12], where the interference pattern is not a baseband signal, but in fact an AC signal, as diagrammed in Figure 12. Thus, distance becomes a measurement of the difference between two signals, one of which (known as the measurement signal) is modulated by the moving object, while the other (known as the reference signal) is generally fixed. The reference signal is usually composed of the difference between the two frequencies before one of them has been

modulated, but can also be another modulated signal to create a differential measurement between two moving mirrors.

In the interferometer described in [12], a single laser is made to produce two frequencies of light in opposite circular polarizations. In the case of a HeNe laser, this is done by applying a magnetic field to the laser cavity which splits the emission of Ne into two lines (known as the Zeeman Effect [5]). We will consider f_1 to be the measurement frequency and f_2 to be the reference frequency. As the two frequency beam comes out, it is split into a portion that goes to the IF and one that returns back to the measurement system. The return portion is split into a portion to control the laser power and one that images on an optical sensor. This sensor, the Reference Sensor then sees

$$\begin{aligned} E_{ref1}(t) &= A\cos(k_1 L_{ref1} - \omega_1 t + \phi) \\ &\quad + A\cos(k_2 L_{ref1} - \omega_2 t + \phi) \end{aligned} \tag{23}$$

Again, we calculate the intensity by using the Poynting vector to get the variation of intensity at the first detector: Putting these all together,

$$\begin{aligned} \mathcal{P}_{ref1}(t) &= \vec{E}_{ref1}(t) \times \vec{H}_{ref1}(t) = \\ &\frac{A^2}{2}\sqrt{\frac{\epsilon}{\mu}}\left[1 + \frac{\cos 2(k_1 L_{ref1} - \omega_1 t + \phi)}{2}\right. \\ &+ \cos\left((k_1 - k_2)L_{ref1} - (\omega_1 - \omega_2)t\right) \\ &+ \cos\left((k_1 + k_2)L_{ref1} - (\omega_1 + \omega_2)t + 2\phi\right) \\ &+ \left.\frac{\cos 2(k_2 L_{ref1} - \omega_2 t + \phi)}{2}\right] \end{aligned} \tag{24}$$

Again, considering that the higher frequency terms are filtered off, we are left with the baseband and difference frequency term:

$$\begin{aligned} \mathcal{P}_{ref1,LP}(t) &\approx \frac{A^2}{2}\sqrt{\frac{\epsilon}{\mu}}[1+ \\ &\cos\left((k_1 - k_2)L_{ref1} - (\omega_1 - \omega_2)t\right))] \end{aligned} \tag{25}$$

This will provide our detection frequency to compare to the signal from the IF. Similarly to above, the signal at the IF will see:

$$\begin{aligned} E_{IF}(t) &= A\cos(k_1 L_{meas} - \omega_1 t + \phi) \\ &\quad + A\cos(k_2 L_{ref2} - \omega_2 t + \phi) \end{aligned} \tag{26}$$

and so our Poynting vector calculation yields:

$$\begin{aligned} \mathcal{P}_{IF}(t) &= \frac{A^2}{2}\sqrt{\frac{\epsilon}{\mu}}\left[1 + \frac{\cos 2(k_1 L_{meas} - \omega_1 t + \phi)}{2}\right. \\ &+ \cos\left(k_1 L_{meas} - k_2 L_{ref2} - (\omega_1 - \omega_2)t\right) \\ &+ \cos\left(k_1 L_{meas} + k_2 L_{ref2} - (\omega_1 + \omega_2)t + 2\phi\right) \\ &+ \left.\frac{\cos 2(k_2 L_{ref2} - \omega_2 t + \phi)}{2}\right] \end{aligned} \tag{27}$$

and once again, the assumptions that only the baseband and difference frequencies are detected result in

$$\begin{aligned} \mathcal{P}_{IF,LP}(t) &\approx \frac{A^2}{2}\sqrt{\frac{\epsilon}{\mu}}[1+ \\ &\cos\left(k_1 L_{meas} - k_2 L_{ref2} - (\omega_1 - \omega_2)t\right)] \end{aligned} \tag{28}$$

Our first signal from Equation 25 is a sinusoidally varying signal locked to the same source as the signal at our interferometer signal and so the AC portion provides a mixing signal to demodulate the IF signal in Equation 28. This would result in a baseband signal that varied as $\cos\left(k_1 L_{meas} - k_2 L_{ref2}\right)$, and since k_1 and k_2 are known, the change in L_{meas} can be determined.

However, another way lends itself to high speed signals. If the measurement mirror is moving, then that movement will appear as a Doppler Shift in $\omega_1 = 2\pi f_1$, so that $\omega_1 \Longrightarrow \omega_1 + \Delta\omega_1$ and Equation 28 becomes:

$$\begin{aligned} \mathcal{P}_{IF,LP}(t) &\approx \frac{A^2}{2}\sqrt{\frac{\epsilon}{\mu}}[1+ \\ &\cos\left(k_1 L_{meas} - k_2 L_{ref2} - (\omega_1 + \Delta\omega_1 - \omega_2)t\right)]. \end{aligned} \tag{29}$$

One way to measure movement then is to count zero crossings of the $\cos\left((k_1 - k_2)L_{ref1} - (\omega_1 - \omega_2)t\right)$ versus those of the $\cos\left(k_1 L_{meas} - k_2 L_{ref2} - (\omega_1 + \Delta\omega_1 - \omega_2)t\right)$ signal. The two counters will run a the same rate if there is no movement of the measure surface. If the measure surface is moving towards the IF, the Doppler Shift will cause a positive $\Delta\omega_1$, causing the second counter to count faster. If the measure surface is moving away from the IF, there will be a negative $\Delta\omega_1$, causing the second counter to count move slowly than the reference.

One may ask why one would count zero crossings of two AC signals rather than simply demodulate the two. The answer is simply that it is often much easier to build accurate high speed zero crossing counters, then high speed demodulators. Furthermore, zero crossing detectors are sensitive to timing variations, but not amplitude.

V. Position Signal Generation

In previous sections, we have discussed how the interference pattern is generated and techniques used to make this signal immune to various forms of imprecision. In this section, we will show how the signal arriving at the detector is turned into a position signal.

We start with the reminder that all position measurements with an interferometer are relative. The fringes give a change in position from some starting position. This is analogous to trying to measure position from velocity measurements: one must assume a starting position.

IF measurements rely on knowing the wavelength of light, λ, and the wavelength being stable. This is why commercial interferometers did not emerge until lasers were invented. It is important to know that as pressure, temperature, humidity,and gas composition change, so does λ. Thus, an IF system is making measurements with a somewhat elastic ruler.

While we are interested in a distance measurement, we are really measuring an optical path length (OPL). To properly translate this optical path back to distance, we must be able to monitor and compensate for environmental changes. We will see such compensation methods in the sections that follow, but for right now we will start with the basic equation for

interferometer distance measurement (IFM)

$$Position = \text{FC}_{meas} \cdot \text{WCN} \frac{\lambda_{vac}}{\text{OFF}} \tag{30}$$

where

- FC_{meas} = Fringe Counts or their equivalent produced by more modern resolution extension methods,
- WCN = Wavelength Compensation Number, described below,
- OFF = Optics Fold Factor = $\{2|4|8\}$ depending on the optics type, and

The Optics Fold Factor represents the number of times that the beam has traversed the distance to or from the measurement surface. A higher fold factor increases the effective resolution because every distance change is multiplied up by the fold factor. From this equation, one can see that the uncompensated optical resolution, R_{optics} of the setup is

$$\text{R}_{optics} = \frac{\lambda_{vac}}{\text{OFF}} \tag{31}$$

However, this base resolution is usually enhanced by further subdividing the signal with an Electronics Resolution Extension (ERX) factor. So after adding in this term, Equation 30 becomes

$$Position = \text{FC}_{meas} \cdot \text{WCN} \frac{\lambda_{vac}}{\text{OFF} \cdot \text{ERX}} \tag{32}$$

and the resolution of a Fringe Count (FC) becomes

$$\text{R}_{FC} = \frac{\lambda_{vac}}{\text{OFF} \cdot \text{ERX}}. \tag{33}$$

Since the wavelength of light from a particular laser is specified in a vacuum, λ_{vac}, we need the correction term, WCN, for operation in air, where

$$WCN = \frac{\lambda_{air}}{\lambda_{vac}}. \tag{34}$$

Initial systems provided two methods of entering in the WCN value. One was a weather station with air temperature, pressure, and humidity sensors along with a hardwired computing engine. The other was a set of four thumb-wheel switches. Either one was used to obtain the last four digits of the value 0.999xxxx, giving 0.1 ppm resolution to the value. The system's user manual contained pages of tables that listed the last four digits of the WCN value for the expected range of pressure, temperature, and humidity. Section VIII-G discusses how one converts these environmental parameters into changes in λ_{air}. However, for this section, it is enough to realize that generating a distance measurement from our interference pattern on our optical detectors is all about counting the passing of cosine waves, which all have a period related to the laser wavelength, λ.

The typical Helium Neon (HeNe) laser used in these applications has a wavelength of $\lambda = 632.8$ nm. Using double pass optics (OFF = 4) makes the base optical resolution 632.8/4 = 158.2 nm. The initial systems counted zero crossings as shown in Figure 13, which divided this base optical resolution number in half (ERX = 2), resulting in a resolution of 79.1 nm. Some systems used PLLs to frequency

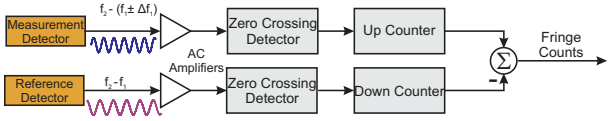

Fig. 13. Generating distance from AC frequency differences.

multiply both reference and measurement signals to further subdivide the base optical resolution to obtain resolutions as fine as 5.3 nm (λ/120) [13]. However, this technique required precise PLL circuits with extremely wide dynamic ranges (3 to 90 MHz) and thus was impractical for increased resolution extension. This might have been solved with high speed digital PLLs, but for many years the processing speed in DSP chips was not sufficient for this task. Furthermore, using both edges of an oscillatory signal can cause errors if the spacing between leading and trailing edges is not uniform, so as higher accuracy was required, systems moved to using leading edges only.

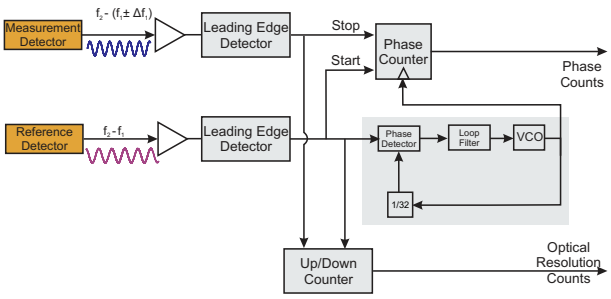

Fig. 14. Upmultiplying the reference signal to increase delay computation accuracy.

To interpolate between optical resolution counts in these systems, a frequency multiplied reference signal ($\times 32$) was used to measure the delay between the reference and measurement signals (Figure 14) and the phase angle [14]. With this method, the result used to compute the phase angle, ϕ was given by

$$\phi_{deg} = \frac{360 \cdot Delay}{32} \tag{35}$$

With requirements for even finer resolution, higher update rates, and the availability of high speed ASICs, this technique evolved into using a very high frequency clock to rapidly ($\geq$ 10 MHz) measure both the period and current phase angle of both the reference and measure signals and digitally process the resulting four data values to calculate the instantaneous phase angle and velocity, while still using a basic up/down counter to accumulate optical resolution fringe count data. This resolution extension method, shown in Figure 15, allows rapid (20 MHz) update rates and 0.1545 nm resolution (ERX = 1024) with plane mirror optics.

With the advent of modern Field Programmable Gate Arrays (FPGAs) the above methods have given way to demodulating the phase as one might have done with an

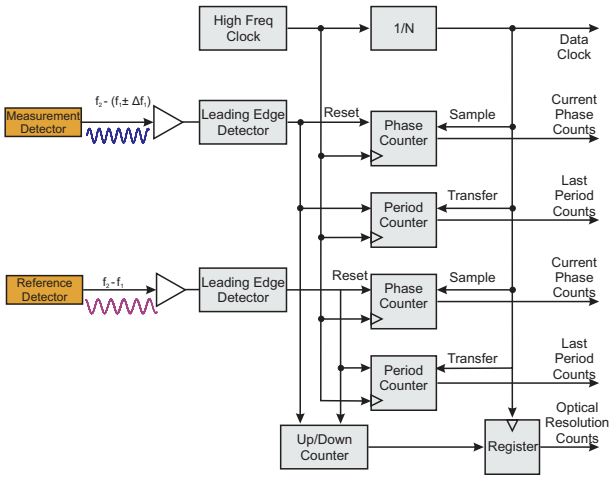

Fig. 15. High speed and high resolution phase measurement.

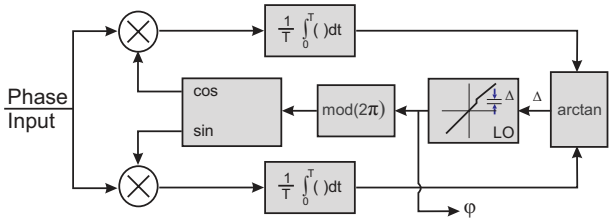

Fig. 16. Phase generation from interference pattern input.

analog PLL, but now entirely digital and at full speed. In particular, one can use a digital Costas loop like scheme shown in Figure 16 [15], [16], [17]. In this case, the incoming phase signal is mixed with sine and cosine signals from a local oscillator. The mixed signals are integrated over an integer number of periods generating signals proportional to the sine and cosine of the phase. These, in turn, can be fed to an arctangent block which backs out the phase difference, Δ, between the LO phase, ϕ, and the phase input. This phase difference is fed into the LO to adjust the phase and the rate of change of the phase. The computed phase can then be digitally differenced from the phase of the reference signal. In fact, with a single reference, phases from many axes can be differenced simultaneously by the FPGA system at high speed. The fractional portion of the difference gives the sub-wavelength accuracy of the measurement, so 10 fractional bits would have an accuracy of $\lambda/4/2^{10} = 0.15449$ nm.

VI. Precision Interferometry Through the Years

Although Michelson invented the interferometer in 1887, it took almost 80 years before displacement interferometers became commercially available. Since then the development has progressed, going from 15.8 nm resolution to 0.15 nm resolution over a period of 40 years. Table I lists a couple of the scientific developments along with the significant commercial developments from HP/Agilent during this timeframe. The table captures some of the key specifications of interest to servo control system designers. Its no surprise that the basic trend has been to higher laser power (for more axes), higher resolutions, higher velocities, faster update rates, shorter delays, and less timing uncertainty. All but the higher power and velocity trends help improve closed loop servo system performance.

VII. Applications of Precision Interferometry

Over the years there have been a few large users of commercial interferometer systems along with numerous small users with specialized measurement needs. The following are some of these uses:

- Servo track writers for hard disk drives were a major business for many years. The servowriters were used to write the dedicated and sectored servo information [18]. Eventually, the rise in areal density of the drives, coupled with the lowering of costs, pushed the idea of self servowriting to fruition largely ending this application around the turn of the millennium [19].
- Wafer steppers are used to move Silicon wafers for imaging with device patterns. As the control capability has improved, these have been replaced by wafer scanners, where the wafers are moved continuously during the imaging process [20].
- Wafer inspection systems examine wafers after they have been imaged to verify the work.
- LCD steppers are conceptually similar to wafer steppers, but have looser accuracy requirements and move much larger surfaces.
- Calibrating other measurement devices, as to their angle, straightness, squareness, etc.
- A variety of other metrology applications abound: calibration of machine tools, coordinate-measuring machines (CMMs), manufacturing and calibration of encoder scales and PZT devices, any distance measurement over long distances where running cables or wires would be prohibitive or dangerous, specialized custom measurement machines.

A. Example: Multi-Axis Precision Measurement Configurations

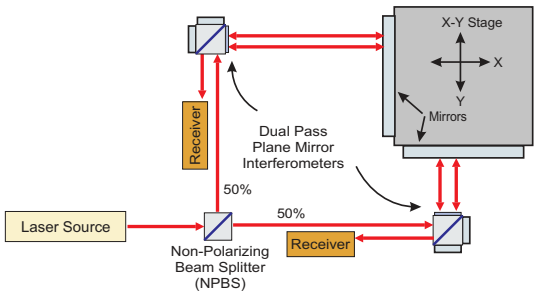

Fig. 17. Two-axis plane mirror interferometer configuration

Intro Year	Model	Linear Optics Resolution (nm) Normal	Extended	Data Rate (kHz) Internal Update	External Output	Output Data Age (μs) Fixed	Variable	Contribution
1887	A.A. Michelson's Interferometer							Basic Measurement Concept
1960	Bell Labs' HeNe Laser							Practical light source for longer measurements
1964	1st Commercial Displacement Interferometer (Airborne Instruments)							Proof of commercial viability for product
1966	HP 5505A	158	15.8	0.01	na	na	na	2 frequency Interferometry
1975	HP 10760A	158	15.8	1800	5 SW		~200	Multi-axis, higher resolution and update rate
1977e	HP 10764B /10762A	158	10.5	≤22500	≤22500			Higher Resolution and update/output rate
1986	HP 5507A	9.9	na	1800	1800—20A	1.7	0.3	Higher Resolution, Lower cost, smaller form factor, higher reliability
1988	HP 5517B							Higher stage velocity (0.5 m/s with Linear Optics)
1991	HP 5517C							Higher stage velocity (0.7 m/s with Linear Optics)
1991	multi-axis interferometers introduced							Reduced axis to axis cosine error, smaller footprint for pitch/yaw measurements
1994	HP 10889B	4.9	na	10000	10000 — 20 A	1.2	0.01	Reduced data age variation, higher update rate
1994	HP 10897A	1.2	na	10000	10000	1.2	0.001	Higher resolution, reduced data age uncertainty
1996	Agilent 5517D							Higher stage velocity (1.0 m/s with Linear Optics)
2002	Agilent N1231A	0.62	na	20000	4000	0.52	0.025	Higher Resolution, Higher Data rate, Multi-Axis Electronics
2004	Agilent N1231B	0.31	na	20000	20000	0.8	0.001	Flexible hardware interface
2004	Agilent 5517FL							Higher stage velocity (2.1 m/s with Linear Optics)
2006	fully integrated multi-axis optics introduced							Reduced Optics Non-Linearity, Higher optical efficiency
2006	Agilent N1211A							Higher stage velocity (4.1 m/s with Linear Optics), Fiber delivery to vacuum
2007	Agilent N1225A	0.31	na	10000	10000	3.15		Integrated high sensitivity receivers, non-linearity correction
2007	Agilent 5517GL							Higher power for high stage velocity (2.2 m/s with Linear Optics)

TABLE I

INTERFEROMETRY THROUGH THE YEARS. NOTE THAT RESOLUTION IS TWO TIMES SMALLER WHEN USING PLANE MIRROR OPTICS. SW STANDS FOR SOFTWARE ONLY UPDATES. "A" STANDS FOR ±10 V ANALOG OUTPUT FOR SERVO APPLICATIONS.

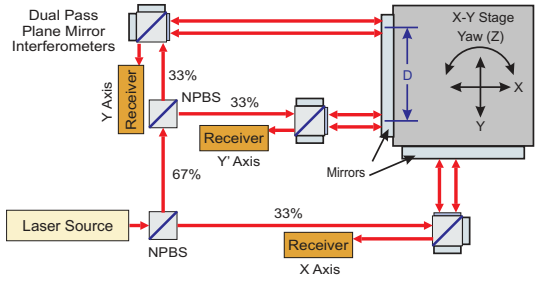

Fig. 18. Yaw measurement of x-y stage with discrete interferometers

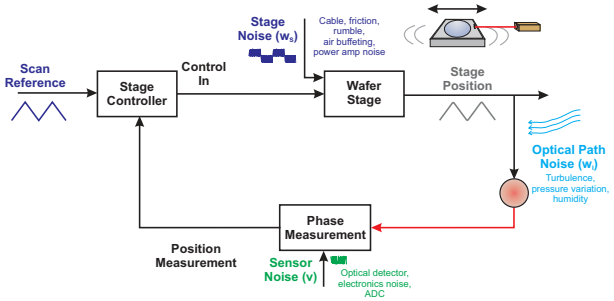

Fig. 19. Using an interferometer in a stage position feedback loop. This diagram shows a SISO application, but IFs are used to measure 6 or more degrees of freedom.

One of the great benefits of precision interferometry for position measurement is that because the measurements are done at a distance, multiple axes can be measured with the same system, by splitting the laser beam and directing it off of different surfaces and back to multiple receivers. This can be seen in the two-axis configuration shown in Figure 17, where a single beam is split and directed at polarizing beam splitter based interferometers. Each of these beams is reflected off of a planar mirrors on the side of a moving stage, resulting in position measurements for the x and y axes. In Figure 18, this system is augmented further

by splitting the beam three ways, allowing for differential measurements in the y direction that can yield a measurement of the stage yaw.

While these two diagrams are by no means exhaustive, they do provide some insight into the design freedom given by IF measurements on moving objects. Considerations have to be made with respect to the:

- system setup to be employed, depending upon needed accuracy and axes to be measured,
- degrees of freedom, basically how many axes will be measured,
- speed and accuracy requirements, how fast those results need to be provided. Calibration measurements typically require far less system speed and allow for considerably higher latency than measurements used to close the loop in precision motion systems.

B. Wafer Stage Measurements

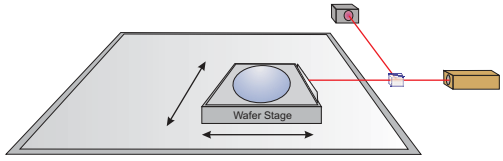

Fig. 20. Wafer stage system measured with interferometer.

The bulk of interferometers systems are used in the IC photolithography industry [20], [21]. Here, very precise machines move an X-Y stage under an optical column. What is critical is the location of the stage relative to the optical column, and the repeatability of this measurement. So these systems use laser interferometers to measure the X and Y positions of the stage and the column, as well as the pitch and yaw of these items. Some even measure the vertical direction of the stage. Figure 20 shows the basic setup without the optical column which would obscure the stage.

VIII. The Modern Interferometer

The improvements to the basic Michelson interferometer described in Section III give the tools to fix a variety of issues, and modern interferometers are built with combinations of these improvements. However, as accuracy and speed requirements increase, these first order fixes become susceptible to other errors. These issues will be delineated in this section, along with designs that compensate for them. In all cases, it is assumed that the interferometers use multiple frequencies and polarizations of light, and that the beam passes through a PBS some number of times to generate interference fringes. We will see that there are a lot of issues remaining in making our distance measurement immune from inaccuracies and noise, as well as some clever designs to overcome these hurdles.

IF measurements depend on knowing the phase difference between two light beams and their wavelength. Anything that affects the apparent phase or wavelength affects the accuracy of the measurement. Thus, despite their utility, there are substantial issues in the use of interferometers for precision measurement systems. These issues must be understood to appreciate the solutions that have been applied already as well as new ones that are being proposed. A non-exhaustive list of sources of errors and other issues [22]:

- Intrinsic: laser wavelength accuracy, measurement resolution, optics non-linearity, computational issues, degrees of freedom, angular range, velocity requirements speed and accuracy requirements.
- Environmental: optical thermal drift, atmospheric compensation, material thermal expansion, turbulence.
- Installation: Deadpath error, cosine error, Abbé error.
- Feedback loop usage: data age, interfacing with control system.

The above issues will be described individually in the following sections.

A. Laser wavelength Accuracy

Michelson [3] used a sodium flame for his monochromatic light source, and although it contained a narrow range of optical frequencies, its stability is not sufficient for modern measurements. With the invention of lasers, and subsequent frequency stabilization methods [23], [24], [25], one now had a light source that was both single frequency and highly collimated, perfectly suited for use in interferometry. Thus for vacuum applications, one would have a very small uncertainty for the wavelength value. A discussion of the atmospheric effects on the wavelength of the measured beam is in Section VIII-G.

B. Optics Non-Linearity

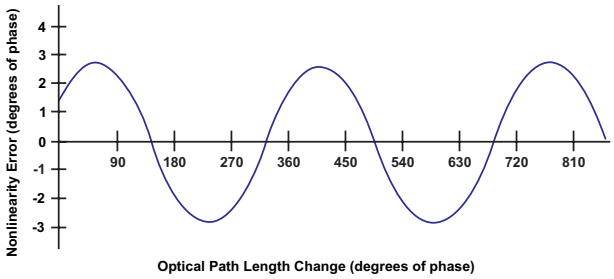

Fig. 21. Worst-case error resulting from imperfect separation of two beam components. Recreated from Chapter 12, Figure 48 of [26].

The discussion so far has been about the intended signals reaching the detector, interfering, and creating a measurement signal, and the phase of that signal being linearly proportional to the motion of the mirror. And, for the most part, this model worked well until the measurement resolution became finer than 4 nm. At this level, the effects of unintended signals reaching the detector started to be seen. Real world interferometers contain additional optical paths due to signal leakage of one polarization into the other polarization's optical path. When these unintended signals reach the detector, they distort the signal from the ideal

[27], [28], [29]. These signals and their effects are covered in [30] (which will be presented later in this session), but in brief, the unintended signals cause periodic non-linearities in the transfer function between the motion of the mirror and the output position value from the laser electronics. The errors do not accumulate, but cycle about zero with a period of one wavelength of optical path length change as illustrated in Figure 21 [26]. To reduce these errors, modern interferometers use higher quality PBSs and wave plates. Electronic methods have also been devised to measure and correct these errors [31].

C. Computational Issues

Modern interferometer systems convert the analog light signals into digital signals and process the resulting data stream with digital electronics. This introduces several issues typically associated with real time computation. In this section, we will discuss the issues that are internal to the interferometer processing itself. A complimentary set of issues involve interfacing the digital electronics of the interferometer with the system that will be making use of it, such as a control system. Those will be addressed in Section IX-B. Here we will discuss

- quantization error,
- computational delay, and
- limited measurement range.

Quantization error is essentially the timing resolution that the system uses to measure the phase change between the measurement and reference signals. For most systems this is sufficiently small to not be a real issue, but as IC line widths continue to decrease, even 0.15 nm resolution of modern electronics will not be fine enough for the position feedback to the photo-lithography equipment's servo systems. But other applications without these tight resolution requirements will have more resolution than they need.

The computational delay causes the most complications for servo systems. As laser IF systems have improved to compensate for optical issues and improve resolution, the required computations have increased. Faster electronics help offset this, but the overall computation times have gone up in the latest systems. Fortunately, what matters most is the variation in these computation times. Fixed delays (generally referred to as Data Age) can often be compensated by adjusting the position values by the velocity times the delay. However, the variable portion of the data age can not be corrected, and it shows up as measurement uncertainty. Efforts are made to minimize this delay uncertainty.

With the high resolution represented by the LSB of a digital position word, it takes a lot of bits to keep track of a stage of moderate size (32 bits allows only +/- 331 mm at $\lambda/4096$ resolution). Early systems provided 28 bits of position information, which was sufficient for most systems then due to the lower resolution (0 to 21 m at $\lambda/8$ resolution). Modern laser systems provide 32 to 36 bits of position information, which still covers a wide range when all bits are used (+/- 5.3 m at $\lambda/4096$ resolution). So designers must balance resolution, word width, dynamic range, processing power, and computation time to match their system's performance requirements (although with the relatively low cost 64-bit computing platforms available today, this is less of an issue).

D. Degrees of Freedom

There are two aspects to this topic:

1) For a given axis, how many degress of freedom does the mirror move?
2) For an overall system, how many degrees of freedom does the system measure?

Ideally, the reflector for each interferometer axis only moves in the direction being measured, or intentionally allowed for in the case of plane mirrors. However this is rarely the case, and cube corners move laterally to the beam, or plane mirrors pitch and yaw. In both of these cases, the result is lateral displacement of the return beam, which causes reduced signal strength to the detector and/or added cosine error. The WOW interferometer discussed below implements a solution to address the first impact of this unwanted motion. Adding multiple measurement axes to the system either with additional interferometers or through the use of multi-axis interferometers, addresses the second impact.

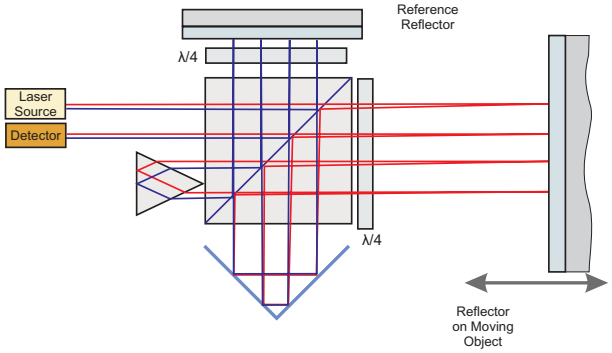

Fig. 22. A WOW interferometer uses four passes to cancel the effects of walkoff as well as angle due to stage rotation. The

Although the PMI improved the measurement mirror's angular range, there is still a need for additional angular range at longer distances. The addition of a prism at the output of the interferometer sends the beam back through the optics where any walkoff that occurred in one direction now occurs in the other direction. In a With Out Walk-Off (WOW) interferometer [32], the result is a final output beam that does not move when the measurement mirror pitches or yaws and with twice the resolution as an added bonus. The same concept can be applied to an interferometer using cube corner reflectors to allow greater lateral movement of the cube corner while maintaining the tolerance for large angular variation of this reflector. With a WOW interferometer, the angular range is limited by the size of the optics instead of the beam diameter for any given measurement distance.

Figure 18 above shows an example of using two interferometers to measure the yaw of a stage along with the X and Y location. The information obtained from the yaw

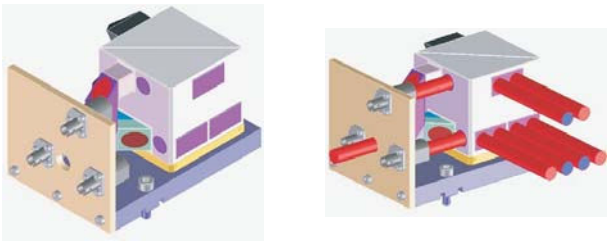

Fig. 23. Agilent Z4399 A Three-Axis Interferometer: On the left is the device. On the right is the device with representation of the laser beams. (Recreated from Figures 242 and 243 of [33].)

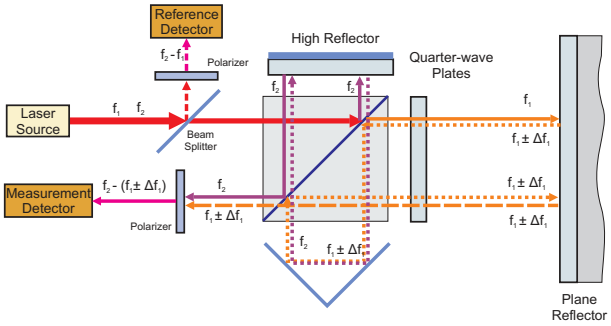

Fig. 24. Laser beam paths in a two frequency balanced path PMI

data can be used to correct for the resulting cosine error changes as well as for the precise location of the point of interest on the stage. This implementation of a yaw measurement has both a plus and some minuses. On the plus side, the large separation between the two measurement points means the angular measurement has a high resolution. On the minus side, the use of two interferometers means that there will be some difference in cosine error between the two measurements, and also that a larger stage mirror will be required. To address both of these minuses, as well as the cost and installation complexity of using multiple interferometers, multi-axis interferometers are now routinely used to measure both pitch and yaw of a stage. And systems with multiple multi-axis interferometers are used to fully monitor all degrees of freedom.

E. Velocity Requirements

For a number of reasons, system velocity requirements have gone up over the years. For DC interferometry, there are no fundamental limits yet to the maximum velocity that the system can track. However for AC interferometry, the split frequency presents a hard speed limit for one direction of travel. Since the basic equation for the Doppler frequency shift is $f_m = f_1 \pm \Delta f_1$ (or $f_m = f_1 \pm 2\Delta f_1$ for two pass IFs), when $\Delta f_1 \geq f_1$, then the detected frequency goes through 0 Hz and becomes "negative". But it is difficult (economically impossible?) to monitor this transition and keep track of the sign of f_m, and thus actual distance, once it has gone to 0.

Section VI shows that the laser head split frequency has increased over the years to accommodate the higher velocity requirements.

Practical experience has shown that using an Acoustic Optic Modulator allows greater control and range of split frequencies than simply using the Zeeman Effect on its own. However, as the split frequency and velocity go up, the reference and measurement signal periods go down. And any fixed timing jitter in the detector circuits becomes a larger percentage of these periods, and thus can limit the resolution and accuracy of the measurement.

F. Optical Thermal Drift

Single pass IFs have balanced paths, so interferometer temperature changes affected both beams almost equally, and thus have minimal impact on the measurement. The initial two pass interferometer shown in Figure 9 does not have balanced paths, so any temperature change causes a greater phase shift to one beam path than the other, and thus a significant measurement error. The initial PMI design has a temperature coefficient of $\approx$ 500 nm / °C due to the measurement beam transversing about twice as much glass as the reference beam. The revised two pass interferometer of Figure 24 adds a quarter waveplate to the reference path and changes the reference mirror from a cube corner to a plane mirror. These changes make both paths identical and reduce the interferometers temperature coefficient to less than 40 nm / °C, a 12.5× reduction. This design change is carried forward with new interferometer design with one goal being to keep the two optical paths as equal as possible so any environmental changes are common mode and thus self canceling. However the realities of interferometer fabrication mean that there will always be some path length differences and thus some measurement change that correlates with temperature .

G. Atmospheric Compensation

Section V introduced the WCN term as the wavelength compensation factor, and Equation 34 defines it as the ratio of λ_{air} to λ_{vac}. If one knows the index of refraction of air, then one can calculate the compensation number as the reciprocal of the index of refraction. In 1966, B. Edlén made numerous measurements of the air pressure, temperature, humidity, and gas composition, and created an equation, the Edlén equation [34], that relates the index of refraction to these measured parameters. In brief, the Edlén Equation [35], [36], [37], [38] describes the variability of the index of refraction of the optical path:

$$n - 1 = K(\lambda)\frac{P}{T} \cdot \frac{1 + \varepsilon P\,(1 - \alpha T)}{1 + \frac{\delta}{T}} \approx K(\lambda)\frac{P}{T} \qquad (36)$$

where $n - 1$ is the deviation of the refractive index away from 1, $K(\lambda)$ is a wavelength dependent scaling factor, and $\rho = \frac{P}{T}$ is the density of air.

The constant K is the bulwark of the Edlén relation which has a wavelength dependence for a standard atmosphere ($P \approx 10^5$ Pa, $T \approx 300$°K) including a correction for

relative humidity. For $\lambda = 633$ nm, the value of $K\frac{10^5}{300}$ is approximately $3 \cdot 10^{-4}$ which means $K \approx 9 \cdot 10^{-7}$ °Kelvin/Pascal at standard temperature and pressure (STP)). Thus a 1mK change in temperature or a 1/3 Pa change in pressure results in a change in index n of 1 ppB. Note that motion of a bluff body (such as one's hand) moving at 1 m/sec can induce one half ($\frac{1}{2}$) Pa change of presser.

Scaling these relations up to larger environmental changes, one can derive some basic rules of thumb that can be used to approximate the change in wavelength due to environmental changes [22]:

- 1 ppm per 1 Deg C change in air temp
- 1 ppm per 2.5 mmHg change in air pressure
- 1 ppm per 80% change in air relative humidity

Relating these back to IF measurements, if continuous updating of a compensation factor is not done, and the environment changes by 0.5 deg C, 0.5 mmHg, and 20% RH, then the measurement will be off by 0.95 ppm, or 0.95 um per meter of distance between the interferometer and stage mirror.

So one should use some form of compensation, and the Edlén equation will provide the index of refraction for the air, and thus a compensation number, based upon quantities that can be measured. Even the early versions of interferometer systems provided a means to either measure these parameters (an indoor weather station) or enter a compensation number that was obtained from a table based on the user's own measurement (or estimation) of them. Use of this compensation method allows automatic (or manual) calibration of the overall atmospheric conditions. However, what this method cannot resolve is turbulence (discussed more in Section X), which is a much more localized effect of air pressure variation in the optical beam path.

H. Material Thermal Expansion Issues

Both the interferometer itself and a physical object being measured will have thermal expansion issues. Ideally, the interferometer and the object being measured can be in a temperature controlled environment, however, much of the time this is not possible. While it is possible to compensate for much of the interferometer's variation by balancing the optical path (see Section VIII-F), the dimension change in the object can be calibrated out by knowing the coefficient of thermal expansion (CTE), α, of the object. Two such applications of distance interferometers are calibration of machine tools and coordinate-measuring machines (CMMs). By standards, the length measurement values have been tied to the standard temperature of 20 °C. When the object to be measured/calibrated is not at 20 °C, the measured value must be recalibrated back to what it would have been at 20 °C. For example, if one were measuring the length of an object, L, then the true length at standard temperature, $T_{nom}(= 20°C)$ could be obtained via

$$L = L_T\left(1 - \alpha(T - T_{nom})\right) \tag{37}$$

The material temperature compensation factor is usually combined with the atmospheric compensation factor used to scale the vacuum wavelength value used in the conversion equation to the wavelength of the laser light in the current air environment.

I. Deadpath Error

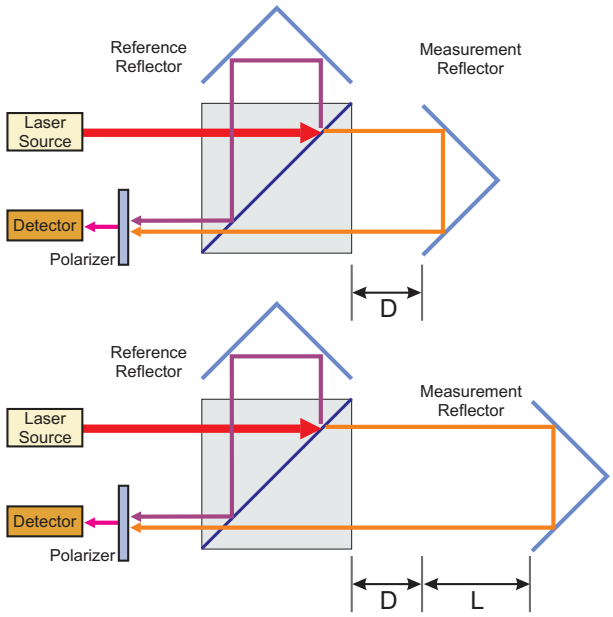

Fig. 25. Deadpath error in an IF measurement. The initial measurement on top includes a distance, D, which is included in the "zero" position. The lower measurement has a total distance of $D + L$, but since the D distance is the "zero" the IF only counts L. Any atmospheric disturbance that affects the optical path in D will contribute to error in the system.

A diagram of Deadpath error [26] is shown in Figure 25. The Deadpath, D, is the part of the optical path that is not part of the measurement, but instead part of the "zero" position. The Deadpath distance should be a part of our IF measurement, so any compensation changes get applied to it as well. To include Deadpath in Equation 32, we both add it in, and subtract it out to end up with

$$Position = (\mathrm{FC}_{meas} + \mathrm{FC}_{Deadpath}) \cdot \mathrm{WCN} \cdot \mathrm{R}_{FC} - D, \tag{38}$$

where

- FC_{meas} = Accumulated fringe counts from the starting position of the measurement and
- $\mathrm{FC}_{Deadpath}$ = Fringe counts that would have been accumulated moving from the zero Deadpath ($D = 0$) position to the actual starting position of the measurement.

To calculate the $\mathrm{FC}_{Deadpath}$ value, we use Equation 32, substituting D for $Position$ and defining WCN_0 as the compensation value at the system reset time. Thus,

$$\mathrm{FC}_{Deadpath} = \frac{D}{\mathrm{WCN}_0 \cdot \mathrm{R}_{FC}}. \tag{39}$$

Substituting (39) into (38), separating terms, and simplifying yields

$$Position = \mathrm{FC}_{meas} \cdot \mathrm{WCNR}_{FC} + D\left(\frac{\mathrm{WCN}}{\mathrm{WCN}_0} - 1\right). \tag{40}$$

From this we can see that the first term of (40) is identical to (32) and the second term is the correction factor for having Deadpath in the system.

To help one determine if Deadpath correction is needed in any given situation, one needs a quick way to estimate the magnitude of this correction term. Section VIII-G presented rules of thumb for obtaining an estimate for a ΔWCN value. Rearranging the Deadpath correction term from Equation 40, using the definition of ΔWCN:

$$\Delta\text{WCN} = \text{WCN} - \text{WCN}_0, \tag{41}$$

and the knowledge that WCN_0 will always be within about 300 ppm of 1, the magnitude of the Deadpath error can be estimated as [26]

$$\text{Error}_{Deadpath} = D \times \Delta\text{WCN} \tag{42}$$

This equation tells us that we want to minimize Deadpath and changes to the compensation number in order to minimize this error source.

J. Cosine Error

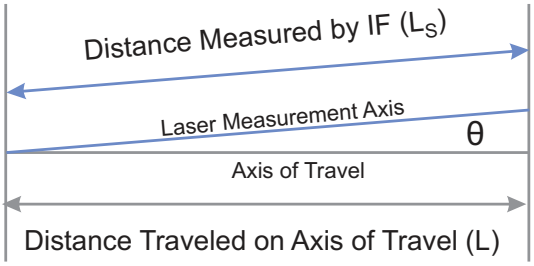

Fig. 26. Diagram of cosine error as diagrammed in [26].

All distance measurements are prone to cosine error, diagrammed in Figure 26, which occurs when the measurement axis is not perfectly parallel with the line connecting the two points that are being measured. The actual length traveled, L, is smaller than the apparent distance traveled, L_S by

$$L = L_S \cos\theta \tag{43}$$

which means that

$$E_{cos} = L - L_S = L_S\left(\cos\theta - 1\right). \tag{44}$$

The cosine error is usually characterized by its magnitude,

$$|E_{cos}| = L_S\left(1 - \cos\theta\right). \tag{45}$$

Proper installation and optical alignment techniques and tools are needed to minimize this error source. Active pitch and yaw measurements (which can be made using multiple beams) can also be used to characterize θ and remove the cosine error in software.

The relevant optics manuals provide detailed alignment techniques to minimize cosine error, generally getting it to less than 0.05 ppm (mis-alignment angle less than 0.35 mRad).

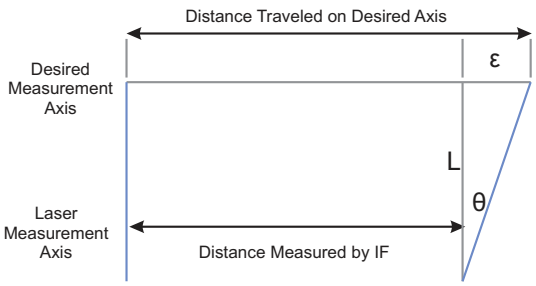

Fig. 27. Diagram of Abbé error as diagrammed in [39].

K. Abbé Error

All distance measurements are also prone to Abbé error [26], [39], shown in Figure 27. Abbé error is caused by a lateral offset, L, between the desired axis of measurement and the actual axis of measurement, as well as a rotation, θ, of the measurement mirror away from being perpendicular from the axis of measurement. The resultant Abbé error, ε is defined as:

$$\varepsilon = L\tan\theta. \tag{46}$$

A general rule of thumb is that Abbé error is approximately 0.1 μm per 20 mm of offset for each arc-second of angular motion [26]. While any type of displacement transducer is susceptible to Abbé error, laser interferometers can minimize this by placing the measurement axis closer to the desired measurement axis or by actively measuring the angle, θ, and compensating for the error in software. Pitch and yaw measurements, which can be accomplished with multiple IF beams, also enable measurement of θ and compensation for this.

IX. Using Interferometer Measurements in Feedback Loops

While precision IF measurements can be used for static measurements, the ability to tie these systems into feedback loops has dramatically raised their utility. Non-contact, multi-dimensional, measurements provide a lot of advantages, and with resolution in the sub-nm range, and sample rates up to 20 MHz, there are few control problems that can outrun the data at sample rate provided. (For example, the Agilent N1231B PCI Three-Axis Board with External Sampling updates position and velocity values for three axes of measurement at 20 MHz, with 0.15 nm resolution, on 32 or 36 bit data words [40].) Control using precision interferometers allow enough measurement precision to push new control methodologies as will be described in [41].

However, as with all control systems, certain issues remain. In particular, data age (or the latency in the interferometer from the time that position is sensed until it is available to the control computer, is described in Section IX-A. Issues of digital interconnect between interferometers and digital control systems are discussed briefly in Section IX-B.

A. Data Age Issues (time delay of measurement)

As system performance requirements have improved, the time required to process the data and generate a position value, as well as the variation in this time, have become more of an issue. When systems are moving slowly or have low servo update rates, then any uncertainty introduced by several microseconds of delay or hundreds of ns variation in this delay are insignificant. But as systems move faster, then 50 ns of timing uncertainty will introduce 50 nm of measurement uncertainty when moving at 1 m/s.

Even with a 20 MHz sample rate, the 50 ns sample period gives opportunity for up to a full sample period of delay in handing off the digital outputs to another processor. To minimize this, the Agilent N1231B "external hardware sample inputs are synchronized to a 160 MHz clock. The circuits then interpolate between two successive internal values so the position read over the PCI bus corresponds to the sample time plus the sample delay time ±4 ns." [26]

B. Interfacing With Control System

Another issue facing use of interferometers is that due to the high resolution, the position value contains many bits. In general, commercial servo control systems are setup to handle A quad B and Sin/Cos inputs, but only a few directly accept large digital words. So some provision must be created to transfer the digital word to the control system. Since there is usually some processor within the control system, it can extend the range of the measurement by adding bits to the left of the transferred position word provided the word size exceeds some minimum number of bits. This word size can be determined from the servo update rate, the system's maximum velocity, and the resolution of the laser measurement system using the following relationship (based on making sure the stage doesn't move more than 1/2 of the dynamic range of the word size in 1 servo period's time so the controller will be able to determine in which direction the motion occurred).

$$Resolution \times 2^{N-1} > \frac{v_{max}}{f_{S,servo}} \tag{47}$$

where N is the number of bits required in the position word, v_{max} is the maximum velocity of the object being measured, and $f_{S,servo}$ is the sample rate of the servo system.

Moreover, beyond the word length requirements, there is the pervasive issue of interfacing two digital systems at high speed. Real time data must be handed off from the interferometer processing boards to the control system with minimal latency and in a synchronous manner. This often requires a lot of custom programming between the interferometer vendor and the customer buying the systems.

X. Turbulence

As discussed all through this article, accurate measurements require an accurate knowledge of the laser wavelength, λ, across the entire length of the measurement path. Basic atmospheric compensation can be accomplished with a weather station that measured pressure, temperature, and humidity as described in Section VIII-G. However, this can only make adjustments for static or overall fluctuations. Turbulence in the measurement beam is more insidious in that it can cause a dynamic change in air pressure across the measurement beam, thereby distorting the wavelength (Figure 28). Two frequency interferometry provides some immunity [12], but as required precision has increased, the minute effects of turbulence have remained. Certainly, turbulence can be minimized in temperature and humidity controlled environments, but movement of objects through any medium, be it liquid or air, causes pressure waves. In high precision, high speed control systems, such as wafer scanners (Figure 29, the turbulence effects can distort the measurement on the order of tens of nanometers, which is too much for the line spacing of current generation computer chips.

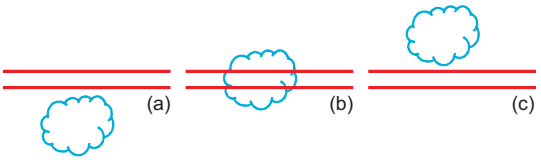

Fig. 28. Turbulence "bubble" crossing interferometer beam.

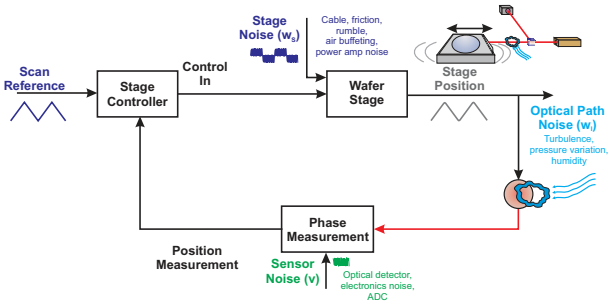

Fig. 29. Conceptual block diagram of wafer stage interferometer (IF) measurement disturbed by turbulence.

A. Alternate Methods: Grid Plates

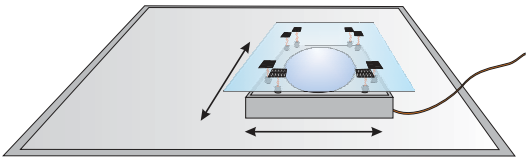

Fig. 30. Example of grating/scale method

Since turbulence happens along the beam path, the longer the beam path, the more susceptible the measurement is to turbulence. One way to minimize this problem is to use a different type of interferometer in combination with grating plates placed in proximity to the wafer scanner and in parallel with the main plane of motion as shown in Figure 30 [42],

[43]. The resolution of these systems is limited by the size of the features on the grating, rather than the optical wavelength, but the optical paths are quite short, minimizing but not eliminating, the effects of turbulence. The downside of this method is that it requires that either gratings be on the moving stage or lasers be on the moving stage. If lasers are placed on the stage, the cabling requirements to the stage go up dramatically, as does the drag and offset forces caused by the cables. If the gratings are on the stage, then these take up much of the area of the stage, which is supposed to mostly hold the wafer. Furthermore, large glass gratings with nanometer precision require incredible manufacturing precision, and at nanometer scales, glass distorts and flows, so such a system must be repeatedly recalibrated.

B. Closed-Loop Interferometer (solution to some problems with current interferometer systems)

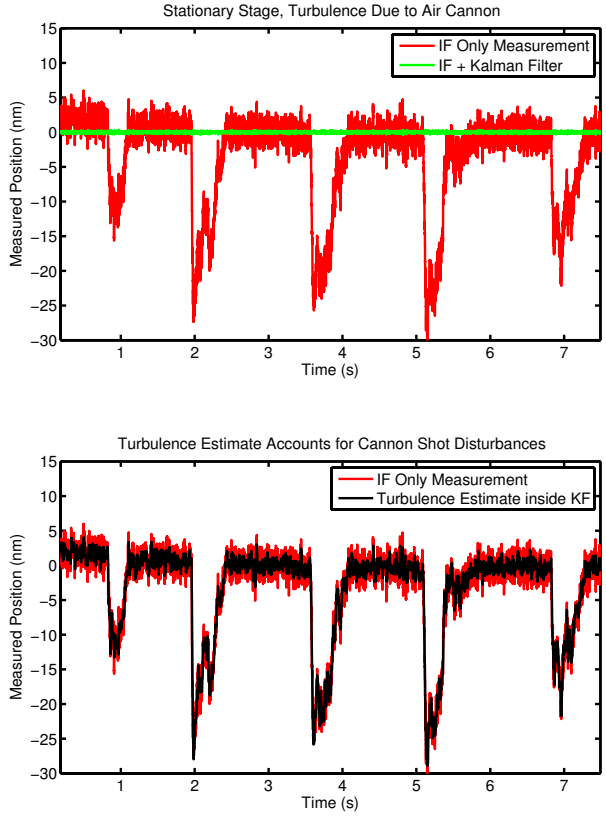

Fig. 31. Experimental results from QP algorithm [7] with a stationary stage and interferometer beam disturbed by air cannon shots.

A new interferometry method known as quintessential phase, QP, comes from the realization that turbulence bubbles such the one in Figure 28 are not noise and do not materialize instantly "out of thin air." They are in fact dynamic processes that can be detected (using a multi-segment detector), and if they can be detected, they can be tracked, using an Extended Kalman Filter (EKF) [44]. This work is described in detail in [7], but Figure 31 shows the method being used on a static stage test with turbulence pulses from an air cannon across the measurement beam. In short, interferometry using a single detector (red line) cannot detect the change in the wavefront as being anything different from a position change. The multi-segment detector makes the turbulence induced change in the wavefront observable, allowing the EKF to model it as turbulence (black line), and then remove this effect from the position estimate (green line).

XI. Conclusions

This tutorial has hopefully given the reader an introduction to laser interferometers as a measurement device with strong applications to feedback loops. The ability to make non-contact, highly precise, multi-dimensional, high speed measurements of moving physical objects should be quite attractive to most control engineers. The bulk of the document has been spent on giving the reader an understanding of how these position transducers work, what can limit their accuracy, and the incredible design innovations that have gone into improving the accuracy, robustness, and speed of these devices.

References

[1] Wikipedia, "Visible spectrum." en.wikipedia.org/wiki/Visible_spectrum, 2013.

[2] Wikipedia, "Interferometry." en.wikipedia.org/wiki/Interferometry, 2013.

[3] A. A. Michelson and E. W. Morley, "On the relative motion of the earth and the luminiferous ether," *American Journal of Science*, vol. 34, no. 203, pp. 333–345, 1887.

[4] G. R. Fowles, *Introduction to Modern Optics*. New York, NY: Dover, second ed., 1975.

[5] E. Hecht and A. Zajac, *Optics*. Addison-Wesley Series in Physics, Reading, MA: Addison-Wesley, 1979.

[6] S. Ramo, J. R. Whinnery, and T. V. Duzer, *Fields and Waves in Communication Electronics*. New York, NY: John Wiley & Sons, second ed., 1984.

[7] E. Johnstone and D. Y. Abramovitch, "Quintessential phase: A method of mitigating turbulence effects in interferometer measurements of precision motion," in *Proceedings of the 2013 American Control Conference*, (Washington, DC), AACC, IEEE, June 17–19 2013.

[8] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference, and Diffraction of Light*. Oxford, England: Pergamon Press, sixth ed., 1980.

[9] J. William H. Hayt, *Engineering Electromagnetics*. Electrical & Electronic Engineering, McGraw-Hill Inc., fourth ed., 1981.

[10] C. Nave, "Quarter wave plate." hyperphysics.phy-astr.gsu.edu/hbase/phyopt/quarwv.html, 2013.

[11] Wikipedia, "Waveplate." en.wikipedia.org/wiki/Waveplate, 2013.

[12] J. N. Dukes and G. B. Gordon, "A two-hundred-foot yardstick with graduations every microinch," *Hewlett-Packard Journal*, pp. 2–9, August 1970.

[13] Hewlett Packard, *10764B Fast Pulse Converter*, part number 10764-90012 ed., April 1982.

[14] Hewlett Packard, *5527A/B Laser Position Transducer Designers Guide*, part number 05527-90015 ed., August 1992.

[15] D. C. Chu, "Phase digitizer," United States Patent 6,480,126, Agilent Technologies, Palo Alto, CA USA, November 12 2002.

[16] D. C. Chu, C. J. Courvile, and L. C. Kalem, "Phase digitizer for signals in imperfect quadrature," United States Patent 6,952,175, Agilent Technologies, Palo Alto, CA USA, October 5 2005.

[17] D. C. Chu, L. C. Kalem, and C. Schluchter, "System and method for interferometry non-linearity compensation," United States Patent 7,436,519, Agilent Technologies, Palo Alto, CA USA, October 14 2008.

[18] D. Y. Abramovitch and G. F. Franklin, "A brief history of disk drive control," *IEEE Control Systems Magazine*, vol. 22, pp. 28–42, June 2002.

[19] J. Nie and R. Horowitz, "Control design of concentric self-servo track writing systems for hard disk drives," in *Proceedings of the 2010 American Control Conference*, (Baltimore, MD), pp. 2631–2640, AACC, IEEE, June 2010.

[20] H. Butler, "Position control in lithographic equipment: An enabler for current-day chip manufacturing," *IEEE Control Systems Magazine*, pp. 28–47, October 2011.

[21] R. M. Schmidt, G. Schitter, and J. V. Eijk, *The Design of High Performance Mechatronics: High-Tech Functionality by Multidisciplinary System Integration*. Delft, NL: IOS Press (Delft University Press), September 15 2011.

[22] R. Kneppers, "HP laser interferometers," *Vaisala News*, pp. 34–37, 1999.

[23] T. Baer, F. V. Kowalski, and J. L. Hall, "Frequency stabilization of a 0.633-μm He-Ne longitudinal Zeeman laser," *Applied Optics*, vol. 19, pp. 3173–3177, September 15 1980.

[24] M. A. Zumberge, "Frequency stability of a Zeeman-stabilized laser," *Applied Optics*, vol. 24, pp. 1902–1904, July 1 1985.

[25] G. S. Sasagawa and M. A. Zumberge, "Five-year frequency stability of a Zeeman stabilized laser," *Applied Optics*, vol. 28, pp. 824–825, March 1989.

[26] Agilent Technologies, *Agilent Laser and Optics User's Manual, Volume I*, fifth, part number 05517-90086 ed., September 2007.

[27] G. Fedotova, "Analysis of the measurement error of the parameters of mechanical vibrations," *Measurement Techniques*, vol. 23, no. 7, pp. 577–580, 1980.

[28] R. Quenelle, "Nonlinearity in interferometric measurements," *Hewlett-Packard Journal*, vol. 34, no. 4, p. 10, 1983.

[29] C. Sutton, "Nonlinearity in length measurements using heterodyne laser Michelson interferometry," *Journal of Physics E: Scientific Instrumentation*, vol. 20, no. 1290-1292, p. 20, 1987.

[30] T. L. Schmitz, V. Ganguly, J. Yun, and R. Loughridge, "Periodic error correction in heterodyne interferometry," in *Proceedings of the 2013 American Control Conference*, (Washington, DC), AACC, IEEE, June 17–19 2013.

[31] V. Ganguly, D. White, T. L. Schmitz, and J. Yun, "Experimental evaluation of phase digitizing nonlinearity correction in heterodyne interferometry," in *Proceedings of the 2012 Precision Engineering Annual Meeting*, (San Diego, CA), ASPE, ASPE, October 21–26 2012.

[32] E. S. Johnstone, J. J. Bockman, A. B. Ray, and K. Bagwell, "Interferometer using beam retracing to eliminate beam walk-off," United States Patent 6,897,962, Agilent Technologies, Palo Alto, CA USA, May 24 2005.

[33] Agilent Technologies, *Agilent Laser and Optics User's Manual, Volume II*, fifth, part number 05517-90086 ed., July 2007.

[34] B. Edlén, "The refractive index of air," *Metrologia*, vol. 2, pp. 71–80, April 1966.

[35] K. Birch and M. Downs, "An updated Edlén equation for the refractive index of air," *Metrologia*, vol. 30, no. 3, pp. 155–162, 1993.

[36] P. E. Ciddor, "Refractive index of air: New equations for the visible and near infrared," *Applied Optics*, vol. 35, pp. 1566–1573, March 20 1996.

[37] P. E. Ciddor and R. J. Hill, "Refractive index of air: 2. group index," *Applied Optics*, vol. 38, pp. 1663–1667, March 20 1999.

[38] P. E. Ciddor, "Refractive index of air: 3. the roles of CO_2, H_2O, and refractivity virials," *Applied Optics*, vol. 41, pp. 2292–2298, April 20 2002.

[39] D. Musinski, "Displacement-measuring interferometers provide precise metrology," *Laser Focus World*, December 2003.

[40] Agilent Technologies, *Agilent N1231B PCI Three-Axis Board with External Sampling: User Guide*, first, part number n1231- 90006 ed., October 2004.

[41] X. Chen and M. Tomizuka, "Control methodologies for precision positioning systems," in *Proceedings of the 2013 American Control Conference*, (Washington, DC), AACC, IEEE, June 17–19 2013.

[42] D. C. Chu and E. R. Whitney, "System for sensing an absolute position in two dimensions using a target pattern," United States Patent 7,489,409, Agilent Technologies, Santa Clara, CA USA, February 10 2009.

[43] J. Yun, D. C. Chu, M. D. TEnuta, R. Yeung, and N. T. Nguyen, "System architecture for sensing an absolute position using a target pattern," United States Patent 7,869,662, Agilent Technologies, Santa Clara, CA USA, January 11 2011.

[44] A. E. Bryson and Y. C. Ho, *Applied Optimal Control*. 1010 Vermont Ave., N. W., Washington, D.C. 20005: Hemisphere Publishing Co., 1975.