

Sphere diameter interferometry with nanometer uncertainty

Eric Stanfield, John Stoup, Michael Braine and Ted Doiron

National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899, United States of America

E-mail: eric.stanfield@nist.gov

Received 15 August 2019, revised 22 October 2019

Accepted for publication 25 October 2019

Published 12 February 2020



Abstract

In this paper we present both the design details and uncertainty budget for a National Institute of Standards and Technology (NIST)-developed high-accuracy Fizeau-type interferometer specifically adapted to measurement sphere diameter. To do this, we provide some historic details of the instrument's origins, an explanation of the clever accuracy-improving enhancements integrated into the design over the past several decades, and a detailed discussion of the role automation has played in achieving its current world-class accuracy. With utmost attention to the dominating components of a detailed uncertainty budget, we methodically improved a rather dated instrument design into a high-accuracy state-of-the-art system for sphere diameter measurement with an expanded uncertainty, $U(k = 2)$, of less than ± 10 nm for spheres with nominal diameters of between 0.4 mm and 26 mm. We conclude with a brief description of a novel application of this instrument for determining the phase-change upon reflection correction for both gauge blocks and gauge block platens. Applying corrections for phase-change upon reflection differences between different surfaces, gauge block and platen, is critically important for high-accuracy interferometric gauge block length measurements.

Keywords: sphere, diameter, interferometry, phase correction, gauge block, ball, Fizeau interferometer

(Some figures may appear in colour only in the online journal)

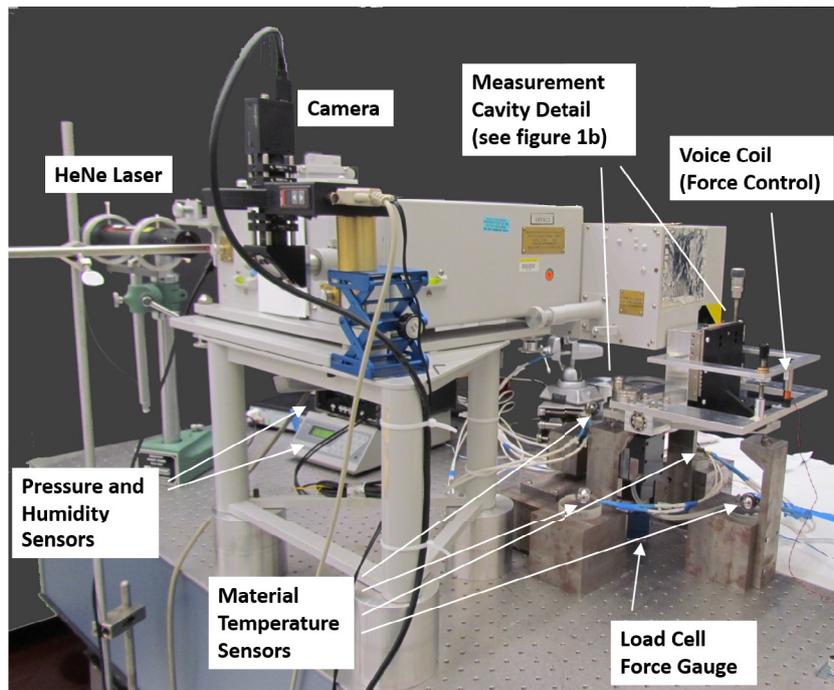
Background

Over the last 30 years the Dimensional Metrology Group at the National Institute of Standards and Technology (NIST) has used a specialized Fizeau-type interferometer for the diameter measurement of spheres. This instrument, known as the Strang Viewer, is used for measuring NIST's own master spheres and customer master spheres. The NIST masters are used in a mechanical comparison process for measuring customer spheres as a lower cost, less accurate, and quicker alternative to direct measurement by interferometry. NIST master spheres are also used on our Moore¹ M48 coordinate

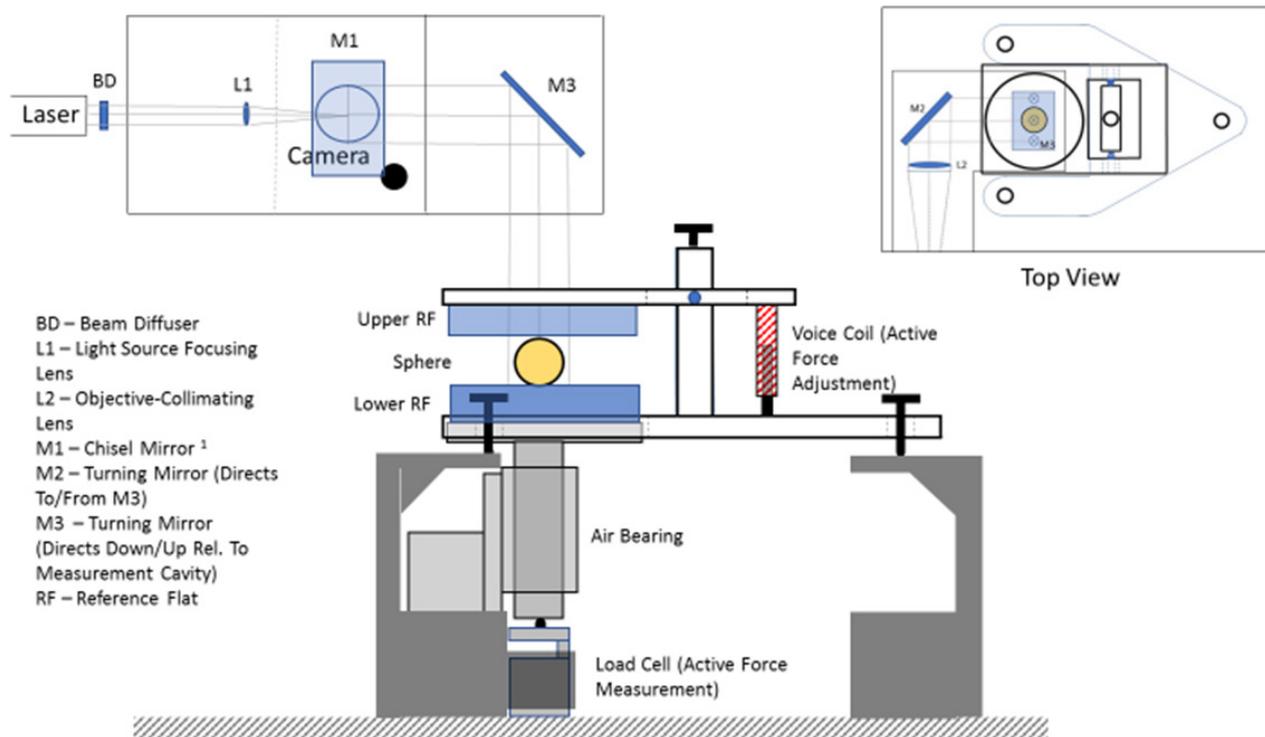
measuring machine (CMM) as qualification and reference spheres and are one of the key elements of the M48's world-class accuracy.

The Strang Viewer is named in honor of its designer and builder George Strang, who built several interferometric viewers in the 1960s to measure the diameters of spheres and cylinders. It is called a viewer rather than an interferometer, because the instrument was designed simply to produce a beam of parallel light and a means to view fringes produced by constructive and destructive interference of the light reflected from two precision flat surfaces. To obtain fringes, these reference surfaces must be oriented near, but not exactly, parallel to one another and orthogonal to the optical axis of the instrument. Furthermore, the upper reference flat must have a significant level of transparency to permit the transmittance of the reflected light from both surfaces, i.e. the bottom surface of the upper reference flat and the top surface of the lower reference flat. Fringe contrast is

¹ Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.



(a)



¹ This was designed originally to serve two purposes: 1) redirect the light source and 2) provide a two-position selectable slit width, one for the cadmium spectral lamp and the other for the mercury lamp.

(b)

Figure 1. (a) Strang viewer. (b) Strang viewer diagram.

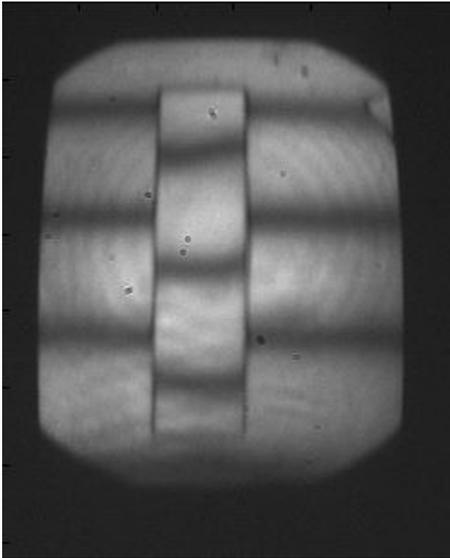


Figure 2. Gauge block interferometry fringes.

maximized when the reflectivity of both opposing surfaces of the reference flats are closely matched, therefore the upper surface is typically completely transparent when the lower surface is not highly reflective (i.e. uncoated lower reference flat) or coated to produce some small level of reflectivity, typically less than 10% when working with a highly reflective lower reference surfaces. The Strang Viewer becomes a Fizeau-type interferometer only when combined with the measurement cavity assembly, which is formed by the upper reference flat and lower reference flat. The Strang Viewer design is shown in figures 1(a) and (b). For sphere diameter measurements, both the upper and lower reference flat are uncoated fused silica to avoid introducing a phase change of the reflected light.

The design retains its historical capability of using Cd-114 spectral lamps while at the same time today's modern 633 nm Helium Neon stabilized laser. The Cd-114 spectral lamp produces multiple wavelengths corresponding to red, green, blue and violet wavelengths and is only relevant for determining proper fringe order, if the sphere diameter cannot be pre-estimated within $\lambda/4$ using our mechanical comparison process, where λ is the 633 nm wavelength of the laser. The Cd-114 wavelength values used, for this purpose, and their associated uncertainties are published by Bureau International des Poids et Mesures (BIPM) [1]. Currently, this capability is only needed for spheres less than 1 mm in diameter because they cannot easily be measured in the mechanical comparator. To enable automated multicolor interferometry using the Cd-114 spectral lamp, the instrument's video camera retrofit includes an optical filter wheel. This optical filter wheel is controlled using an electro-mechanical mechanism to view the desired wavelength.

Other similar interferometric based sphere diameter instruments have been developed over the years and the details of these designs can be found in [2–4]. This instrument's operating principles are different and independent of that developed by Saunders at NIST, for volume determination

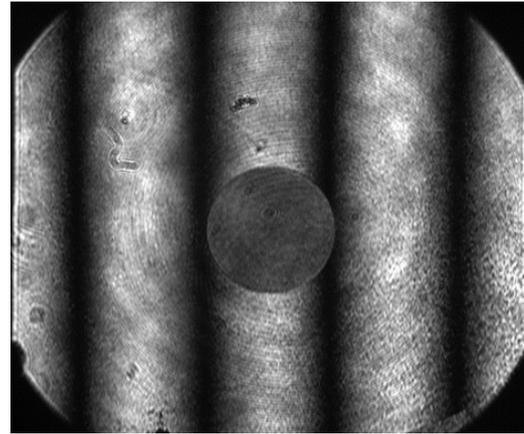


Figure 3. Sphere interferometry fringes.

from sphere and cylinder diameter measurement in the early 1970s [5].

Measurement principle

Sphere diameter interferometry, as performed at NIST, is similar to gauge block interferometry. When an artifact (sphere or a gauge block) is measured using the Strang viewer, static fringes are formed. These fringes are captured as digital images as shown in figures 2 and 3, and image processing techniques are used to calculate the fringe fraction. The fringe fraction, along with the knowledge of the fringe order is used to calculate the dimension of the artifact.

For gauge block interferometry, the upper reference flat is stationary. The platen fringes, represent the distance between the bottom surface of the upper reference flat and top surface of the platen (lower reference flat). The fringes on the gauge block, represent the distance between the top surface of the gauge block and the bottom surface of the upper reference flat. The gauge block does not contact the upper reference flat, so the length is derived as the difference between the distances. In figure 2, the fringe fraction, representing the difference, appears as the offset between the dark bands on the gauge block and the dark bands on the platen upon which the block is wrong [6].

The instrument arrangement is different when measuring a sphere than while measuring a gauge block. The sphere contacts both the upper reference flat and lower reference flat, thus the sphere's diameter is the actual separation between the two. Only one fringe pattern is formed, as shown in figure 3, and it represents the distance between the two reference flats.

Optimal adjustment of the fringes is achieved using the mounting assembly for the upper reference flat, as shown in figure 4, which is adjusted to accommodate a variety of sphere sizes. The empty cavity holding pin, when engaged, prevents the contact between the upper and lower reference flat when there is no sphere is in the cavity. This pin is disengaged when a sphere is placed in the cavity during a measurement. When the sphere is contacting both reference flats, the upper assembly vertical adjustment changes the wedge angle between the two

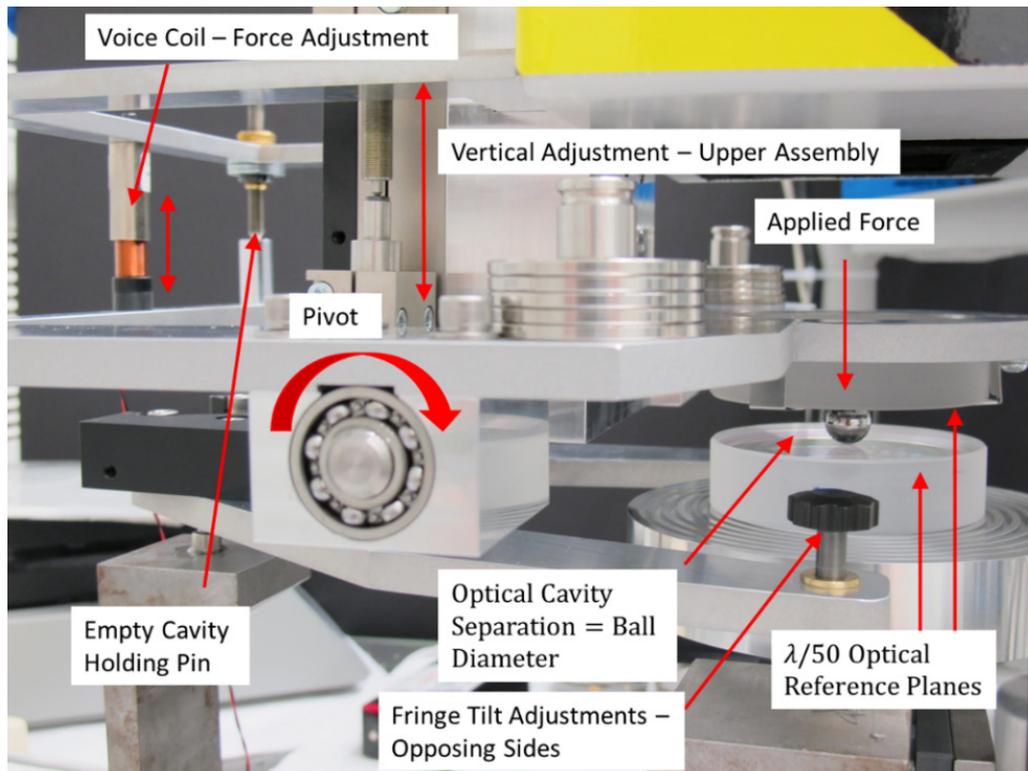


Figure 4. Measurement cavity with sphere to be measured.

reference flats (fringe spacing), since this causes the upper reference flat to pivot about the sphere. The angle of the fringes orthogonal to the wedge angle is adjusted using the two fringe tilt adjustment knobs on the arms of the bottom portion of the assembly. The upper reference flat pivot assembly also provides a means for counterweighting the upper reference flat to finely tune the applied force.

The field of view of the instrument optics limits the setup to the measurement of spheres ranging in size from, 0 mm (in theory), to approximately 25 mm. To set up the sphere in the field of view for measurement, it must be identifiable in the image. In the past, this has limited our measurements to spheres no smaller than 0.3 mm in diameter.

For analysis, rather than calculating the fringe fraction using the fringe of the gauge block at the gauge point relative to the banding platen fringes (figure 5), the center of the sphere is used, as determined from the image. The center is calculated from the least-squares circle fit to image data acquired by a simple edge detection routine (figure 6).

The fringe fraction, f , is determined in a similar manner to gauge block interferometry, except the center of the sphere is used to find ‘a’ rather than the center of the fringe on the gauge block (figure 6). The fringe fraction is calculated using NIST developed, fringe image analysis software, written in Python².

It is important to note; the thick wavy lines are actually very closely spaced data points representing the pixel location of the intensity minimum as fit to a line of pixels orthogonal to the dark band (fringe) in the image. A third order polynomial is

then fit using all the data points along the horizontal orientation in the figures. The a and b in the fringe fraction determination are determined where the vertical line, that intersects the center of the gauge block or sphere, intersects the fit line from the upper and lower fringe data. The assumption here is that the fit lines, from the fit of the pixel minimums, represents the surface that cannot be seen. For the gauge block, the platen or reference flat fringe representing the surface of interest is always hidden. However, for the sphere, depending on the location of the fringes, fringe spacing, and the size of the sphere, the fringe at the points of interest may or may not be hidden by the sphere shadow. The flatness of both, the reference flats and the gauge block, if gauge block interferometry is being done, is critical. In fact, in gauge block interferometry, the flatness of the gauge block surface is a dominant contribution to the overall uncertainty, given that the definition of length of a gauge block is the perpendicular distance between a point on the top of the gauge block at the center point to a plane to which the block is wrung.

The sphere diameter is calculated in three steps addressing various issues in each step. The first step addresses the correction of the wavelength of the light source for the index of refraction of air; the second step addresses the thermal expansion of the sphere based on its material, and the third step addresses the deformation of the measured sphere. In the first step, the fringe fraction, along with the environmental conditions, are used with the updated Edlén equation [7, 8] to calculate the deformed diameter of the sphere. In the second step, the diameter value is then corrected to exactly 20 °C using the measured material temperature and the coefficient of thermal expansion of the sphere material. The result now represents

² See footnote 1.

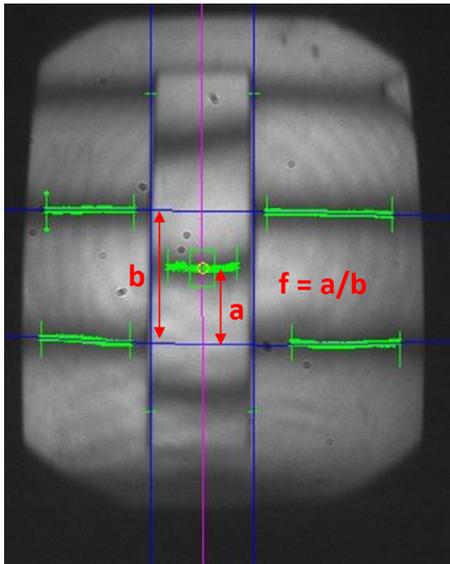


Figure 5. Gauge block fringe fraction.

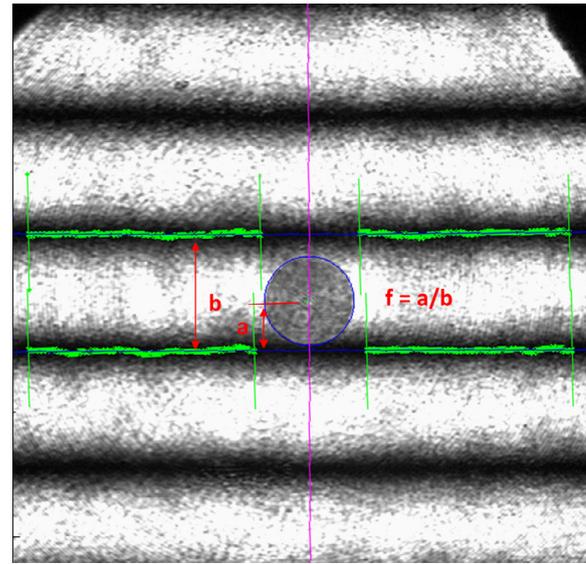


Figure 6. Sphere interferometry fringe fraction.

the deformed condition of the sphere and reference flats at the upper and lower contact points. The third step involves a methodology to account for elastic deformation, which is covered in more detail in the following section.

Compensation for elastic deformation

Though the sphere diameter measurement appears simple, there are several details that make this process more challenging. The optical length measured is the separation between the upper and lower reference flats and represents the ‘deformed’ condition of the sphere at the applied force. The physical contact between the sphere and each reference flat causes deformation of both the reference flats and the sphere. For example, when a 1 mm tungsten carbide sphere is placed between two fused silica reference flats of the instrument at 0.5N, the deformation of both the reference flats and the sphere is approximately 0.78 μm . About 90% of the deformation is on the surface of the reference flats and about 10% on the ball itself. The deformation is substantial and must be considered to obtain an accurate ‘undeformed’ diameter. If unaccounted for, it would represent an error more than 100 times larger than other uncertainty components in the measurement of sphere diameter. Using the specialized derivations of the Hertzian formulas from the Puttock and Thwaite publication [9], a fairly accurate estimate of the point contact deformation for both interfaces can be obtained. However, this correction has an uncertainty due to the uncertainty of the applied force and the elastic properties of the sphere and plane materials.

With the Strang Viewer design, the applied force of the upper reference flat is set by the user. This is done by counterweighing the upper reference flat mounting arm on the opposite side of the supporting pivot joint. Adjustment of the measurement force during the measurement process, from this initial setting, is done remotely using the voice coil.

Determining the applied force at each contact point is a challenge since the applied force varies based on the location

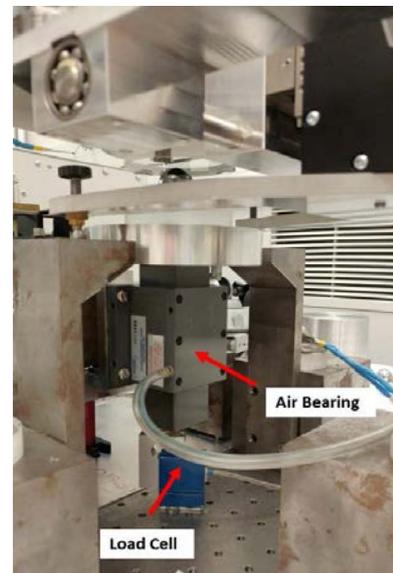


Figure 7. Load cell and air bearing.

of the sphere within the field of view, due to contact location between the upper and lower reference flats in relation to the pivot point of the pivot assembly. Complicating the conditions further, the applied force is different for the upper contact than it is for the bottom, due to the weight of the ball. The upper contact force is the weight applied by the upper contact; however, the bottom force is the weight of the upper reference flat plus the weight of the sphere. The weight of the sphere is easily measured independently on a calibrated scale. To accurately determine the applied force, an *in situ* force measurement solution is incorporated that uses a load cell and an air bearing mounted vertically as shown in figure 7. The load cell force sensitivity and accuracy are verified in place with the air bearing and lower reference flat using a calibrated set of weights. A direct *in situ* reading of the contact force is obtained through a series of steps. First, the empty cavity holding pin is engaged to ensure the two reference flats do not

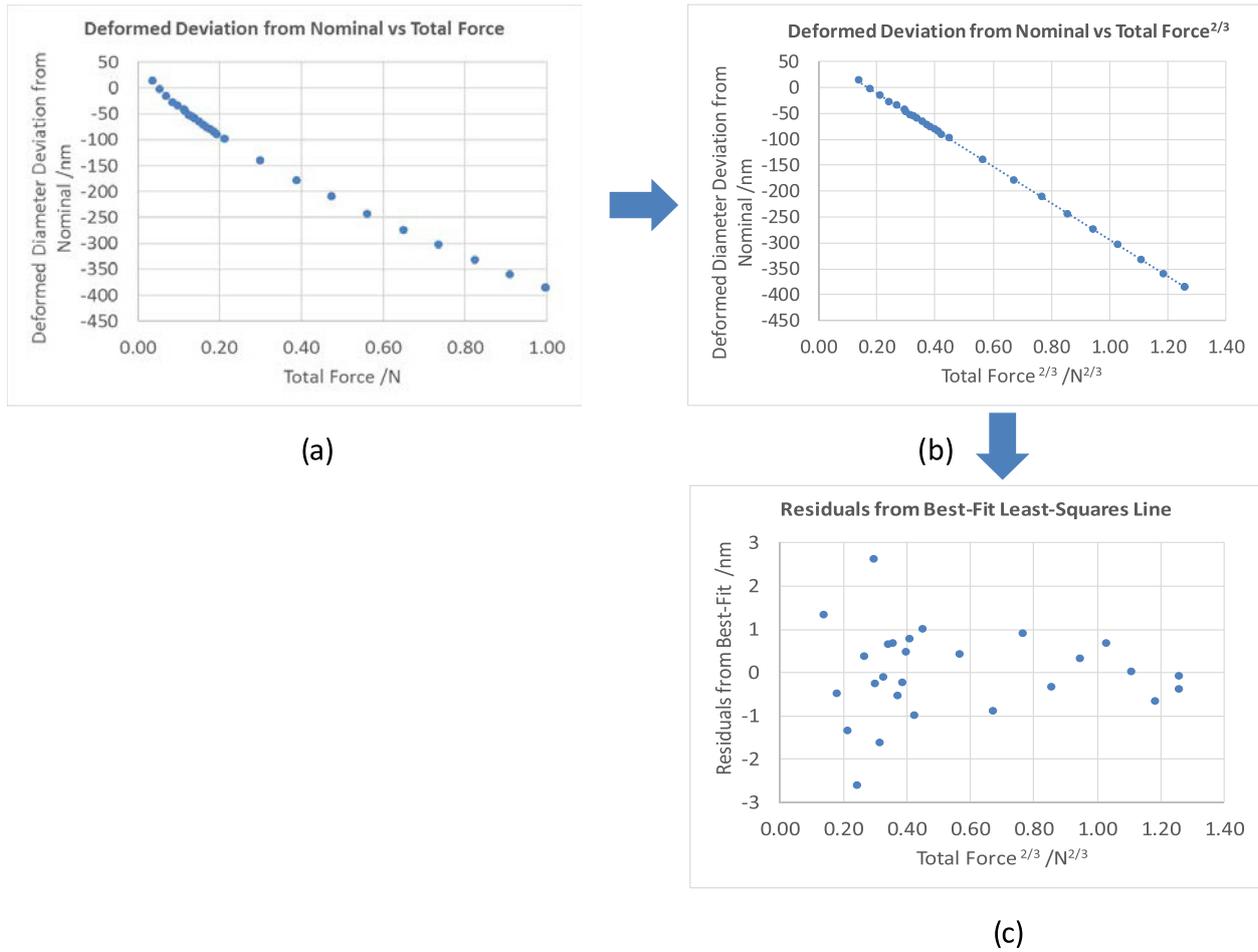


Figure 8. (a) (Top left) Deformed deviation from the nominal diameter versus force, (b) (top right) linearized-deformed deviation from the nominal diameter versus force^{2/3}, (c) (bottom center) residual from best-fit least squares line (after automation).

touch. Next, the weights of the air-bearing and lower reference flat are zeroed out from the load cell reading. Then, the empty cavity holding pin is disengaged, the sphere is placed in the measurement position and properly aligned. At this point, the load cell reading provides a real-time measurement of the instantaneous applied force between the sphere and lower reference flat due to the weight of the upper reference flat and the sphere. Using this force reading, the top force is then determined simply by subtracting the weight of the sphere. At present, it is estimated that both upper and bottom applied forces can be easily determined to better than 0.45 mN.

Sphere diameters are then measured at various forces, without moving the sphere. With the known contact forces and the fringe fractions representing the deformed diameters, the undeformed diameter is determined by a method of extrapolation. The relationship between the two variables is not linear as can be seen in figure 8(a). They are then linearized by plotting the sphere diameter (or deviation from the nominal diameter) and $f^{2/3}$, where f is the force (figure 8(b)). The $f^{2/3}$ component originates from the Puttock and Thwaite equation for point contact between a sphere and a plane for the contact force. The undeformed deviation from the nominal diameter is then determined from the Y-intercept of the linearized data.

This extrapolation method eliminates the uncertainty components associated with the mechanical properties of the

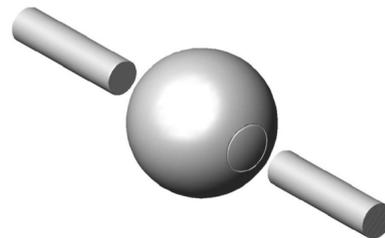


Figure 9. Annular marking with laser engraver to identify a specific two-point diameter, micrometer anvils added to show a two-point gauging situation.

sphere and reference flats. However, eliminating one source of error introduces another, thermal drift, due to the time-consuming nature of force adjustment, alignment, and fringe reading inherent in the original design. Prior to automation, measuring at eight different forces and repeating the measurement at the initial force could easily consume 45 min and resulted in noticeable thermal drift, of about 1 nm–2 nm per measurement. The drift was assumed to be linear and the results were drift corrected based on the difference between the first and last measurements at the initial force.

Significantly reducing or eliminating the drift correction and uncertainty associated with it became the primary objective for achieving further improvement in accuracy. This is

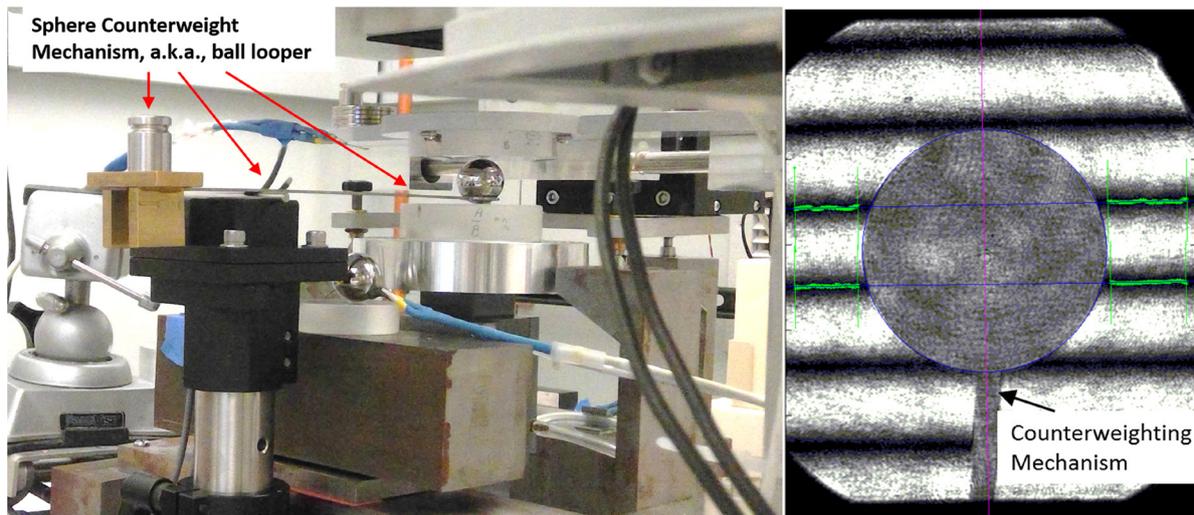


Figure 10. Sphere counterweighting mechanism.

where automation came into play. With the integration of a 1.4 megapixel CCD camera for remote fringe image acquisition, the voice coil for automated force adjustment, and the NIST developed image analysis software for fringe fraction determination; the entire measurement process could be done quickly and remotely, removing the operator from the measurement area completely. These modifications provided a level of automation that reduced the multi-force measurement process time by a factor of roughly 30 while more than doubling the number of measurement forces. The current multi-force measured process was reduced from 45 min to approximately 90 s. With this, drift during measurement was reduced to a negligible 0.1 nm per measurement or less. Figure 8(c) shows the residuals from the best-fit least-squares line, after automation, corrected for any remaining linear thermal drift.

Sphere geometry

The standard measurement process for sphere diameter involves repeating the multi-force extrapolation process for multiple individual diameter measurements to produce a representative average diameter of the sphere. Between individual diameter measurements, the sphere is removed, cleaned, then repositioned in the measurement cavity. The process is typically repeated 10 times and the standard deviation of the ten measurements is divided by the square-root of the number of measurements to produce the standard deviation of the average, which goes into the uncertainty budget and represents a combined term for repeatability and artifact geometry. While the repeatability is excellent, including artifact geometry often makes this term the dominating component of the uncertainty budget. To produce a more accurate sphere diameter measurement, in the absence of a perfectly round sphere, the measurand can be redefined to limit the contribution of out-of-roundness if the sphere can be used in the same manner and satisfy the needs of the customer as a two-point diameter standard. To do so, the sphere can be marked with an annular ring as shown in figure 9 and measured only within the area

enclosed by the circle to limit the measurement to one specific two-point diameter. This is done by physically marking the sphere with a laser engraver. With a specific diameter marked, multiple measurements can then be performed at roughly the same two-point location after removing, cleaning, and repositioning the sphere. This repositioning is done to sample and randomize any diameter variation due to localized geometric form error, of both the sphere (inside the annular ring) and reference flats.

Counterweighting and reference flat geometry

A method for counterweighting the sphere was developed that allows for measurement of larger spheres near zero total force. Doing so, reduces the extrapolation distance from the measurement data to the y-intercept. In addition, reference flats with best available geometry have been introduced. Both modifications are embodied in current design of the instrument.

The counterweighting mechanism can be seen in figures 10 and 15. This counterweighting mechanism provides a means to minimize the contribution of the weight of the sphere to the total applied force between the sphere and lower reference flat. This enabled measurements on large diameter spheres at closer to 0 N force, which effectively reduces the uncertainty associated with the force extrapolation method.

To further minimize uncertainty contributions due to form of the reference flats, both reference flats were replaced with best-available $\lambda/50$ flats. The fringe fitting software does a very good job predicting the form of the cavity close to contact point or center of the sphere, even when one or both of those fringes are hidden by the shadow of the sphere. However, any departures from the fit estimation due to local unseen geometry perturbations in the hidden region closer to where the sphere makes contact, will introduce error. As can be seen in figure 10, the fringes in the field of view, have almost no detectable curve. The assumption is still made that there are no discontinuities at the contact point, however the multiple setup and reading process would reveal such as situation by

making individual outlier identification easy. For example, an outlier related to contact geometry could be the result of a scratch, pit, or dirt.

Uncertainty analysis

The previous sections provide some insight into the measurement method and the improvements that have been made. However, to specify the uncertainty in sphere diameter measurement with highest confidence, an uncertainty analysis was performed by studying the sources of error in the instrument and the measurement process. The sources of error in measuring the sphere diameter are listed below:

- Scale
 - Laser wavelength
 - Index of refraction of air
 - Environmental measurements for wavelength compensation
 - CO₂ content in air assumption
 - Air temperature
 - Atmospheric pressure
 - Partial pressure of water vapor
- Repeatability (includes geometry of sphere and reference flats)
- Temperature related errors
 - Thermometer
 - Coefficient of thermal expansion
 - Temperature gradients
- Instrument geometry
 - Obliquity correction
- Elastic deformation
 - Force extrapolation intercept method

Scale

The current design of the Strang Viewer uses a stabilized HeNe laser with a calibrated vacuum wavelength that has a relative standard uncertainty of 0.6×10^{-8} from the calibration against the NIST maintained iodine stabilized reference. While this component alone is negligible in this application, there is more uncertainty associated with the wavelength to be considered. These include, the intrinsic uncertainty of the Edlén Equation itself, the potential error associated with the assumed CO₂ concentration, and the atmospheric inputs to the equation for pressure, temperature, and partial pressure of water vapor. The Edlén Equation has an intrinsic relative standard uncertainty of 1.0×10^{-8} [8, 18]. For our application, the CO₂ concentration in air is not continuously measured rather assumed to be 450 μmol mol⁻¹. This assumption has an uncertainty and based on experimental data the uncertainty budget includes a potential variability of $\pm 150 \mu\text{mol mol}^{-1}$. The standard uncertainties associated with atmospheric pressure, air temperature, and relative humidity (near standard conditions and converted to the partial pressure of water vapor) are 2.0 Pa, 0.012 °C, and 0.7% respectively. These correspond to relative

Table 1. Scale related relative standard uncertainties.

| Uncertainty source | Length dependent relative standard uncertainty(10 ⁻⁸) |
|---|---|
| Laser vacuum wavelength | 0.6 |
| Index of refraction (modified Edlén equation) | 1.0 |
| CO ₂ content in air assumption | 2.3 |
| Atmospheric pressure | 0.5 |
| Air temperature | 1.1 |
| Relative humidity | 0.7 |

standard uncertainties in length of 5.4×10^{-9} , 1.1×10^{-8} and 7×10^{-9} , respectively, based on known sensitivities [10].

The uncertainties of the pressure and relative humidity readings (inputs) used in the Edlén Equation for the wavelength correction, are dominated by the calibration uncertainty of the sensor reported by the respective department at NIST that performs the calibration. However, the uncertainty of the air temperature reading is dominated by the magnitude of the temperature gradients in the beam path of the optical cavity that exist during measurement, more so than the uncertainty attributed to the calibration of the air sensor versus our primary thermometers. Therefore, the uncertainty of the air temperature sensor reading during measurements includes both the calibration uncertainty of the sensor and the potential gradient between the sensing location during measurement and the sphere being measured. The uncertainty associated with this gradient is estimated from a rectangular distribution with a half-width of 0.02 °C, which yields a 0.012 °C air temperature reading standard uncertainty. The 0.0014 °C uncertainty of the temperature sensor calibration is ignored for this term since the uncertainty due to the gradients is ten times larger. The standard uncertainties associated with the scale are summarized in table 1.

In summary, below are the sources of uncertainty for scale:

Repeatability

Repeatability, in its most basic form, is defined as the level of agreement between multiple measurements without changing anything. Given this, we have quantified the basic repeatability for this instrument and force extrapolation method to be 0.8 nm. This represents multiple measurements of the same diameter of the same sphere without moving the sphere. This repeatability figure does sample variability in the fringe fraction algorithm, the effect of phase noise, and fringe separation. This is especially true since the changes in the atmospheric pressure over the course of the multiple diameter measurements would introduce changes in the fringe pattern and fringe fraction. It could also be considered to sample any variability due to vibration and random errors associated with the wavelength in air and random fluctuations in sphere diameter due to temperature. The 0.8 nm represents the best repeatability achievable. In practice, however, this term is calculated with every sphere measured and since multiple different diameters are sampled to obtain an average, it will include a significant source of uncertainty due to form of both

the sphere and reference flats. For each sphere measured, multiple measurements are made, repositioning the sphere each time and allowing the system to thermally equilibrate. No matter how well a specific two-point diameter measurement location is defined using fiducials on the sphere, there is still variability due to a localized sampling of sphere geometry and the geometry of the reference flats. Instead of devising a separate approach to quantify this contribution independently for each sphere measured, it is sampled as part of the repeatability. Thus, for each sphere measured, this expanded term becomes a dynamic part of the uncertainty budget calculation.

It is important to note that two-point diameter variation is due to the out-of-roundness of the sphere; however, out-of-roundness may not cause two-point diameter variation (e.g. odd lobing in 2D). If a customer requests the average diameter of the sphere, with no orientation reference, the diameter variation due to out-of-roundness will most likely be the dominant source in the reported average diameter uncertainty. In addition to diameter variation due to sphere out-of-roundness, each repositioning of the sphere results in a measurement sampling of different local flatness of the upper and lower reference flat. Incorporating $\lambda/50$ reference flats has made this error source almost negligible. For example, when using a $\lambda/50$ reference flat of 76 mm diameter, along with a HeNe light source of 632 nm, the flatness over the entire surface could be 12.6 nm; however, if only the central 10% of the flat is sampled, this value decreases significantly, but not necessarily in a proportional manner because of the typical profile shapes of reference flats. Including the actual contribution due to the geometry of the reference flats and sphere, the repeatability, emulates the actual measurement-to-measurement process, thus more meaningful than the pure static repeatability. Typically, the measurement process involves measuring the diameter 10 times, setting up the sphere each time between measurements. The repeatability of the of the measurements is simply the standard deviation of the measured diameters divided by the square root of the number of measurements or 10, for this process.

The diameter variability, due to out-of-roundness, is the most dominant source of uncertainty in the uncertainty budget and is highly dependent of the geometry of the sphere being measured and whether or not the sphere is marked to specify a specific diameter

$$u_{\text{Repeatability}} \approx 1.3 \text{ nm (best capability)}.$$

This represents the best repeatability achievable without any contribution due to the form of the sphere but it does add variability due to the geometry of the reference flats to the ‘static’ 0.8 nm repeatability as defined previously.

Temperature related errors

Determining the temperature of the sphere being measured involves the accuracy of the thermometer, the coefficient of thermal expansion of the sphere, and the possible gradient

between the temperature sensing location and the actual sphere temperature. Due to the interferometric nature of the measurement and the measurement process sensitivities (i.e. applied forces), direct measurement of the temperature of the sphere being measured is not possible.

Thermometer uncertainty

The thermometer system uses extremely stable thermistor probes. The uncertainty of the thermometer comes from the calibration of the unit and the additional uncertainties attributed to instability of the unit’s calibration over the calibration interval and any non-linearities, not compensated through the linearization of the probe from the calibration, over the temperature range of use. Our range of use is $20 \text{ }^\circ\text{C} \pm 0.1 \text{ }^\circ\text{C}$. The standard uncertainty of the thermometer reading is $0.0014 \text{ }^\circ\text{C}$. In terms of length, this is then multiplied by nominal coefficient of the thermal expansion (CTE) of the sphere which is being measured. For a steel sphere, the CTE or α is approximately $11.5 \times 10^{-6} \text{ K}^{-1}$, thus the relative standard uncertainty from the thermometer is given by:

$$u_{\text{Thermometer}} = (u_{\text{Thermometer Calibration}}) \alpha = (0.0014 \text{ }^\circ\text{C}) \\ (11.5 \times 10^{-6} \text{ K}^{-1}) = 1.6 \times 10^{-8}.$$

For clarity, it is important to differentiate between the uncertainty of the air temperature and the material temperature. These are two distinctly different measurements and thus are subject to different influences. The material temperature uncertainty is much better than the air temperature measurement uncertainty because the material thermal mass, which the probe is imbedded, is much greater than the thermal mass of the probe alone in air. This results in the material mass with imbedded sensor acting as a long pass filter thus it does not exhibit the high frequency variation seen in the air. This variability contributes significantly to the overall air temperature measurement uncertainty.

Coefficient of thermal expansion uncertainty

The coefficient of thermal expansion of the sphere is not directly measured at NIST, rather the CTE provided by the manufacturer is used. For simplicity, we assume this value to be good to $\pm 10\%$ which is typically the uncertainty reported in material property handbooks. As an example, a steel artifact with a $\alpha = 11.5 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$ would have a potential error in the CTE of $\pm 1.2 \times 10^{-6} \text{ K}^{-1}$ or 10%. Since the exact probability distribution this represents is not known, we assume a rectangular distribution. The maximum difference between the actual temperature of the sphere and $20 \text{ }^\circ\text{C}$ is $\pm 0.1 \text{ }^\circ\text{C}$ which is based on laboratory performance and spatial gradients from one location in the room to another. Given this, we calculate the relative standard uncertainty by:

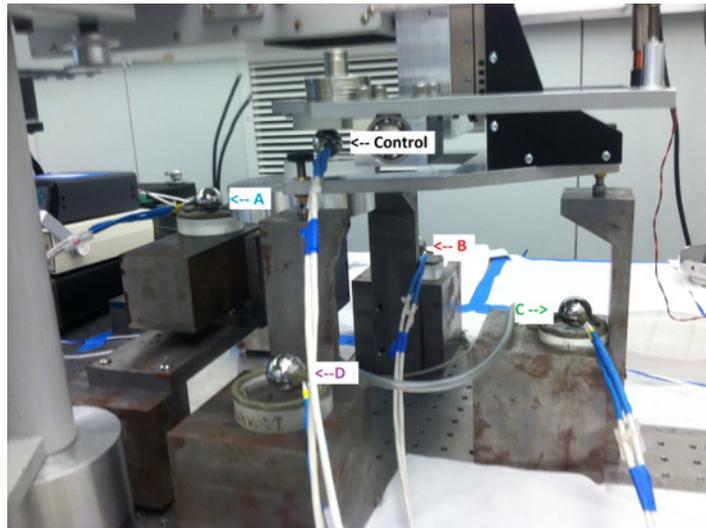


Figure 11. Temperature monitoring system.

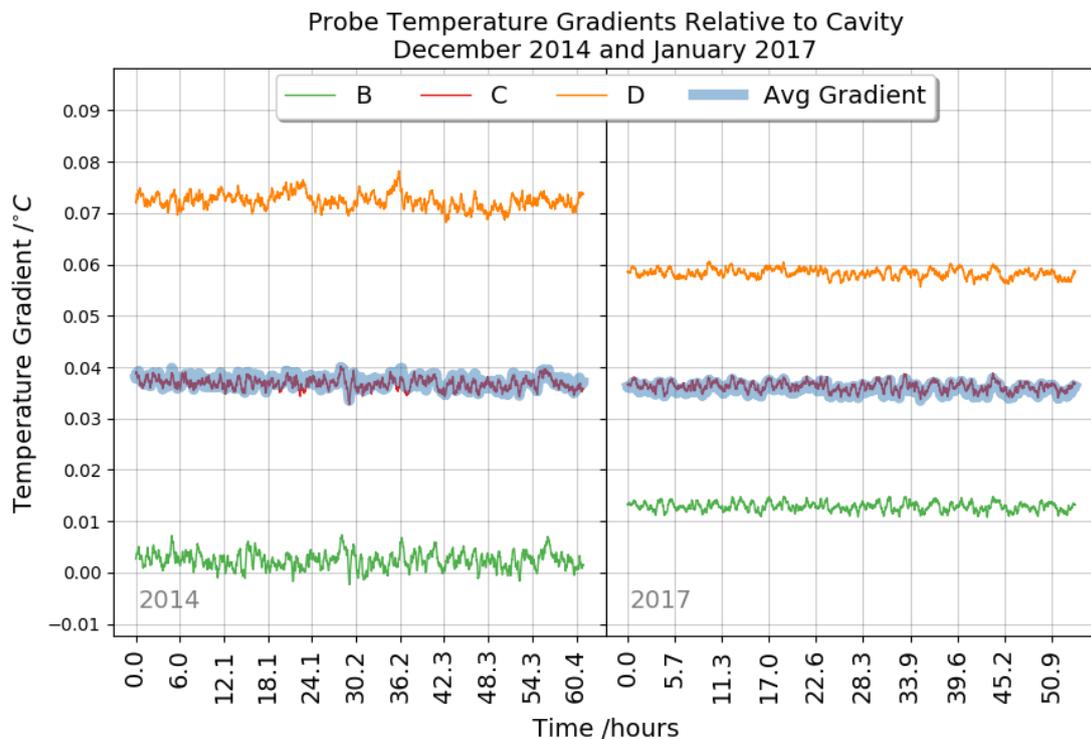


Figure 12. Offset of each temperature sphere/location and average temperature of all ‘non-control’ locations relative to the ‘control’.

$$u_{CTE} = \Delta T \frac{0.1\alpha}{\sqrt{3}} = 0.1 \text{ } ^\circ\text{C} \left[\frac{1.2 \times 10^{-6} \text{ K}^{-1}}{\sqrt{3}} \right] = 6.9 \times 10^{-8}.$$

Temperature gradient uncertainty

The temperature gradient represents the potential error between the temperature at the location of the sensors relative to the temperature of the sphere during measurement in the optical cavity, since the temperature of the sphere under test cannot be measured directly. Normally, to determine this, multiple temperature sensors are set up to sample this error over a

short period of time, then it is assumed that the gradient does not change because of the exceptional temperature control in the lab. The specification for temperature stability in the NIST Advanced Measurement Laboratory (AML) room housing the Strang Viewer is $\pm 0.01 \text{ } ^\circ\text{C}$, but performance indicates 12h stability of $\pm 0.002 \text{ } ^\circ\text{C}$ or better. Since we are striving to maintain an extraordinarily low uncertainty, a system was devised which would remain in place and could be monitored to ensure that temperature gradient conditions do not change. This system is comprised of five 25 mm diameter spheres with 2 mm holes drilled in them to embed the 5 k Ω 4-wire thermistors. Figure 11 shows the location of the five temperature spheres. Each sphere has two sensors, in case one fails. The

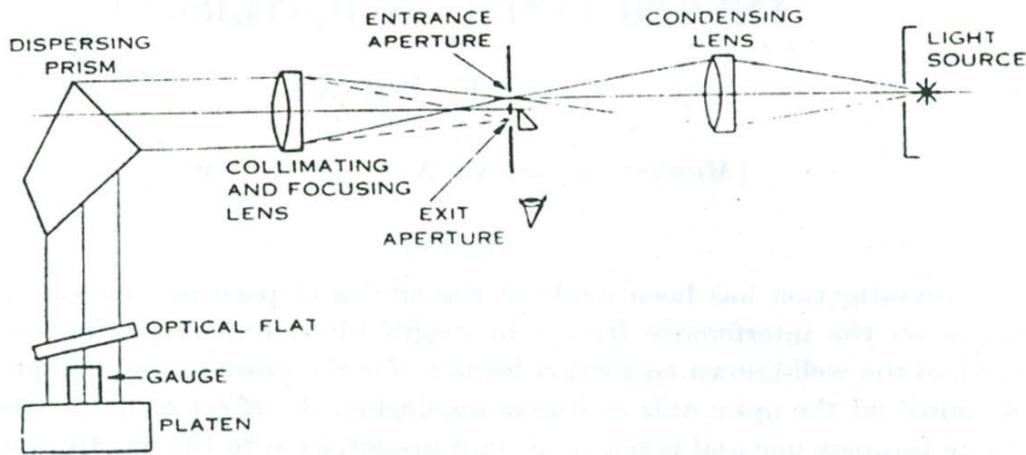


Figure 13. Example of the optical configuration of a Fizeau type interferometer (N P L gauge interferometer) that results in an obliquity error and subsequent correction. (Reproduced from Bruce (1955) with permission from CSIRO Publishing [11].)

fifth sphere is mounted where the sphere under test is located during measurement. Data is collected for many days to monitor the differences between each temperature sphere relative to each other and to the fifth temperature sphere inside the cavity to estimate the real sphere temperature.

Periodically, all ten temperature probes are mounted together in a thermal copper block to determine the offsets between each one so that the individual readings can then be corrected during use. This procedure ensures that they read the same relative to one another. The fifth temperature probe, labeled as the ‘control’ in figure 11, is placed in the measurement cavity to periodically repeat the offset experiment and track gradients. In practice, the temperature of the sphere whose diameter is being measured, is being measured indirectly from the average of the four (A, B, C and D) permanently mounted temperature spheres and corrected based on the data from the periodic offset and gradient experiment. The ‘control’ temperature sphere is set aside and not used during normal measurements.

Figure 12 shows the temperature of each sphere relative to the ‘control’ sphere and the average of the three spheres. We show the results for three temperature spheres instead for 4, as shown in figure 11, because one of the temperature spheres had to be taken offline due to a sensor being broken around the time of the 2017 measurements. Probe D tracks with the average. Probes A and C vary significantly because they are sheltered from the vertical airflow of the room by the instrument and rear of the upper reference flat mounting assembly respectively. At 300 air changes h^{-1} , this can be significant. Comparing 2014 and 2017, the average difference between the average of the surrounding spheres relative to the control sphere is the same to $<0.002\text{ }^{\circ}\text{C}$, however the individual differences between the control location and A and C has become noticeably smaller. This was due to reconfiguration of the probes and ancillary items with respect to the instrument. The graph shows remarkable local stability over 50–60 h. Absolute temperature does move $\pm 0.01\text{ }^{\circ}\text{C}$ day-to-day, but the gradients are quite constant.

During normal operation without the control sphere, the temperature relationship between average of temperature

spheres A and C versus D is closely monitored. If there is a temperature difference of more than $0.002\text{ }^{\circ}\text{C}$, measurements are suspended, and a potential cause is investigated. If the condition does not self-correct or is not confirmed against some laboratory control system change or problem that cannot be fixed, then the experiment with the control temperature sphere is repeated and the new offset relationships are determined.

The relative standard uncertainty of this process is derived as follows using the standard deviation of the gradients (σ_G) from the experiment and the typical worst case CTE (α) for a steel sphere.

$$u_{\text{Gradient}} = \sigma_G \alpha = (0.0025\text{ }^{\circ}\text{C}) (11.5 \times 10^{-6}\text{ K}^{-1}) = 2.9 \times 10^{-8}.$$

Instrument geometry

The Fizeau design of this interferometer results in an obliquity correction [11]. This obliquity correction, which is an optical path angular error resulting from the offset of the entrance (source) and exit (image) apertures relative to the optical axis, can be calculated but is most easily obtained by direct measurement if internal dimensions are not available (refer to figure 13).

To determine the correction, a length standard comprised of a 101.6 mm round DoAll³ Cer-Vit gauge block with a 2.54 mm steel gauge block wrung to the top is wrung to small steel platen as shown in figure 14.

Due to the transparent nature of the Cer-Vit gauge block, a 2.54 mm steel gauge block is wrung to the top to provide a surface that produces fringes with good contrast. This seemingly complicated gauge block combination is used instead of a single steel gage block because ultra-low CTE of the Cer-Vit minimizes the temperature related uncertainty contributions to the measurement. The block and platen assembly were measured in an Automated Gauge Block Interferometer (AGI)^{4,5} of the

³ See footnote 1.

⁴ See footnote 1.

⁵ This instrument was designed by NPL (National Physical Laboratory, UK) and manufactured by Hexagon (formerly TESA).

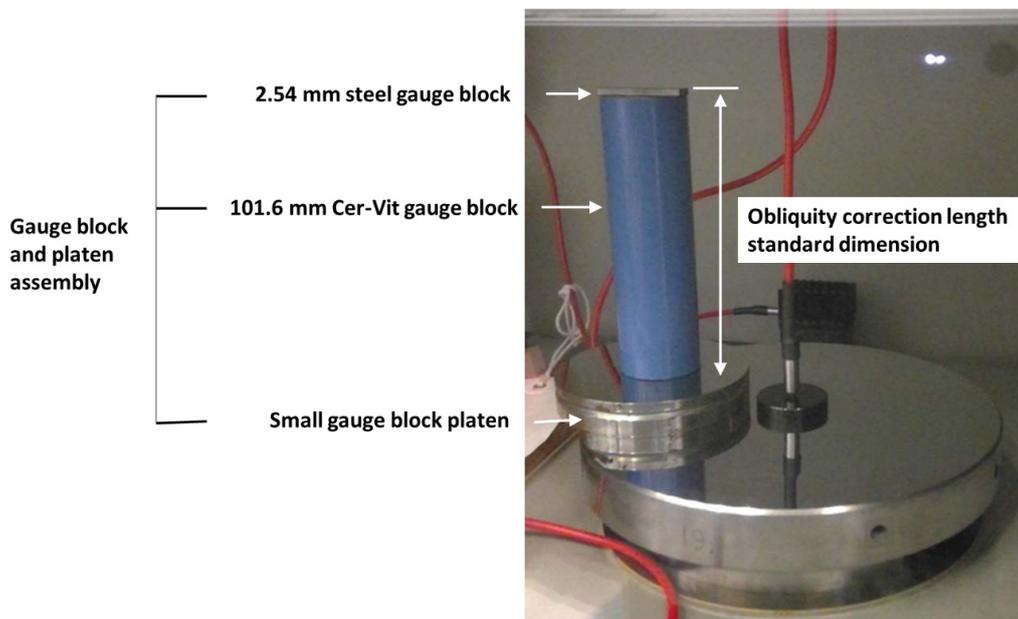


Figure 14. Cer-Vit block in Twyman–Green Interferometer for reference measurement to determine obliquity correction of the Strang viewer.

Twyman–Green design, that does not have an obliquity correction. Then, the block and platen assembly are transferred to the Strang Viewer for measurement. The measurement is performed using the Strang Viewer by removing the bottom assembly (load cell and air bearing) and inserting the block and platen assembly in its place. The Strang Viewer is now configured as a gauge block interferometer with the upper reference flat fixed in place. The measurements in both instruments are repeated over a period of several weeks to sample the range of fringe fraction readings that result with the change in atmospheric pressure. This was done to randomize any subtle, non-systematic, fringe reading errors. By using the same block, platen, and wring, many of the sources of uncertainty associated with a typical gauge block interferometric measurement become common mode (i.e. phase correction and wringing variability), thus make no contribution to the uncertainty of the obliquity correction determination. Thermal errors become negligible through the choice of the Cer-Vit material which has a CTE of less than $1 \times 10^{-6} \text{ K}^{-1}$.

From this obliquity correction determination procedure, the standard uncertainty of the difference between the measurements by both instruments of the combined length of the 101.6 mm Cer-Vit gauge block and 2.54 mm steel gauge block as wrung to the gauge block platen was determined to be 5.2 nm. This results in a relative standard uncertainty of the obliquity correction of

$$u_{\text{Instrument}} = 5.1 \times 10^{-8}.$$

This uncertainty is dominated by the repeatability of both instruments and the uncertainties associated with the corrected laser wavelengths for each instrument.

Elastic deformation

There are two approaches to correcting for deformation that include: (1) applying a calculated correction based on applied

force, contact materials and their associated elastic properties, and the contact geometry/dimensions using specialized Hertzian derivations published by Puttock and Twaite [9], and (2) calculating the undeformed diameter from the Y -intercept using data obtained from measuring the deformed diameter at numerous forces. Plotting the deformed diameter versus the applied force to the two-thirds power linearizes the relationship and linear regression analysis provides for the easy determination of the intercept. This paper focuses on the second approach described next, in that it is the most accurate because it does not require knowledge of the elastic properties of the materials involved, including contributions of their relative uncertainties.

The entire measurement sequence for a sphere is automated. Control and data acquisition is performed remotely from an external control room. The automation includes changing the force by means of a voice coil, acquiring fringe/sphere images with a video camera, data acquisition for the measurement of force, temperature, pressure and humidity, fringe fraction analysis and Edlén Equation calculations to obtain the deformed diameters at each force. Data acquisition and analysis is performed using the Python programming language. Each measurement sequence is easily customizable with regards to the number for force settings, and the maximum and minimum applied force. Typical sequences involve measurement using 20 or more force values with a randomized application order. For thermal drift correction purposes, the first and last readings are at the same force. After completion of the measurement, the operator performs a zero-force check of the force measuring device. At each force setting, multiple images from the camera were acquired, analyzed, and the results are averaged. With the automation, the magnitude of the thermal drift and correspondingly the uncertainty associated with the drift correction, that dominated the uncertainty budget when the system was manually operated,

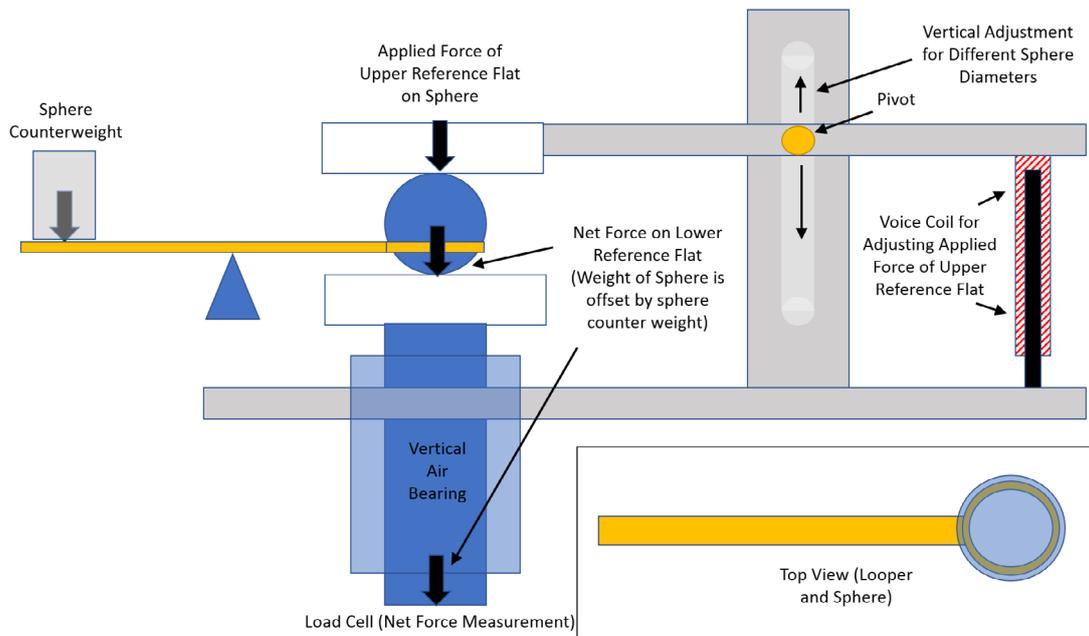


Figure 15. Ball weight counterweighting mechanism-sphere looper.

has been reduced to just a negligible level of a few tenths of a nanometer. A sequence of 15 to 30 measurements at different force can now be completed in 90 s to 120 s.

The uncertainty of the linear regression approach includes the uncertainty of the load cell readings, fringe reading repeatability, non-linear thermal drift, and error in the measured weight of the sphere under test.

The absolute accuracy of the sphere weight is critical as any errors here will systematically affect every force between the upper reference flat and the sphere, since the sphere weight is subtracted from the measured force between the sphere and lower reference flat. The weight of the vertically mounted air bearing and lower reference flat are electronically zeroed from the load cell reading before loading the sphere into the measurement cavity and bringing the upper reference flat in contact. By zeroing these out, the combination of the sphere weight and the upper contact (varied by the voice coil) is measured *in situ*.

The effects of the nonlinearities over the course of the measurement and the fringe reading repeatability are significantly reduced through the linear regression and any remaining variability due to these is captured in the standard error of the intercept from the regression analysis. This sequence is repeated for ten different setups of the sphere being measured and the standard error of the intercept for each run is then pooled and included in the uncertainty budget for as the standard uncertainty due to the elastic deformation correction process. This uncertainty budget entry is updated for each sphere measured. Although small, this number can increase with sphere diameter due to thermal influences and due to variability associated with surface finish.

$$u_{\text{Elastic Deformation}} \approx 1.0 \text{ nm (typical)}.$$

To achieve a very low standard error on the calculated Y -intercept, it was determined that it was important to include

measurements as close as possible to zero force to limit the distance over which to extrapolate the best-fit least-squares line from the data set to the Y -axis. This is simple for small spheres that do not weigh much but for spheres over roughly 12 mm, depending on material density, this can be significant. Because of this, a means had to be found to compensate for the weight of the sphere. The clever approach devised is called the 'sphere looper' and is depicted in figure 15 and it appears in use in figure 10. The sphere looper is made from metal for stiffness and the dimensions of the ring are such that it straddles the sphere about 1/2 the radius of the sphere and so the outer edge of the loop is smaller than the diameter of the sphere. The latter is necessary to ensure the loop does not interfere with image of the sphere, which is used to determine the center. If the design is done carefully, one looper can accommodate a range of sphere sizes to limit the number of loopers to 2 or 3 to cover the range of sphere sizes from 12 mm to 25 mm in diameter. The sphere looper is balanced over an externally mounted pivot point and loaded with weights until the weight of the sphere, as read by the force device, is just a gram or two more than the force exerted by the upper reference flat. Although an initial concern, this setup does not degrade the repeatability of the results.

Uncertainty summary

In table 2, the coefficient of thermal expansion, α , is assumed to be that of typical dimensional artifact steel or $11.5 \times 10^{-6} \text{ K}^{-1}$. The relative standard uncertainty due to the uncertainty in the coefficient of thermal expansion, t_g , represents the maximum temperature from 20 °C the gauge is measured. For this measurement process, $20 - t_g$ is 0.1 °C.

The best case expanded uncertainty, U , with a coverage factor of $k = 2$ would be for a 1 mm diameter steel sphere at $\pm 3.4 \text{ nm}$. For a larger sphere, such as, a 25 mm diameter

Table 2. Uncertainty budget summary.

| Uncertainty source | Standard uncertainty $u(x_i)$ | Sensitivity coefficient $c_{x_i} = \frac{\partial f}{\partial x_i}$ | Length uncertainty $u_i(d) = c_{x_i} \times u(x_i)$ | |
|---|--------------------------------------|---|---|------------------------|
| | | | Fixed | Length dependent |
| Process, ref. flats, and sphere geometry | | | | |
| u repeatability (best) | 1.3 nm | 1 | 1.3 nm | |
| Scale | | | | |
| u laser vacuum wavelength | 6.0×10^{-9} | L | 0.6×10^{-8} 1 | |
| u Edlén | 1.0×10^{-8} | L | 1.0×10^{-8} 1 | |
| u index of refraction input—CO ₂ content assumption | 150×10^{-6} | 1.5×10^{-4} 1 | 2.3×10^{-8} 1 | |
| u index of refraction input—air temperature | 0.012 °C | 9.3×10^{-7} 1 | 1.1×10^{-8} 1 | |
| u index of refraction input—pressure | 2.0 Pa | 2.7×10^{-9} 1 | 0.5×10^{-8} 1 | |
| u index of refraction input—partial pressure of water vapor from RH | 0.7% | 1.0×10^{-8} 1 | 0.7×10^{-8} 1 | |
| Sphere temperature | | | | |
| u thermometer | 0.0014 °C | $\alpha \times L$ | 1.6×10^{-8} 1 | |
| u CTE | 6.9×10^{-7} K ⁻¹ | $ (20 - t_g) \times L$ | 6.9×10^{-8} 1 | |
| u gradients | 0.0025 °C | $\alpha \times L$ | 2.9×10^{-8} 1 | |
| Instrument geometry | | | | |
| u obliquity | 5.1×10^{-8} | L | 5.1×10^{-8} 1 | |
| Deformation | | | | |
| u deformation | 0.9 nm | 1 | 0.9 nm | |
| Combined standard uncertainty, u_C | | | 1.6 nm | 9.6×10^{-8} 1 |

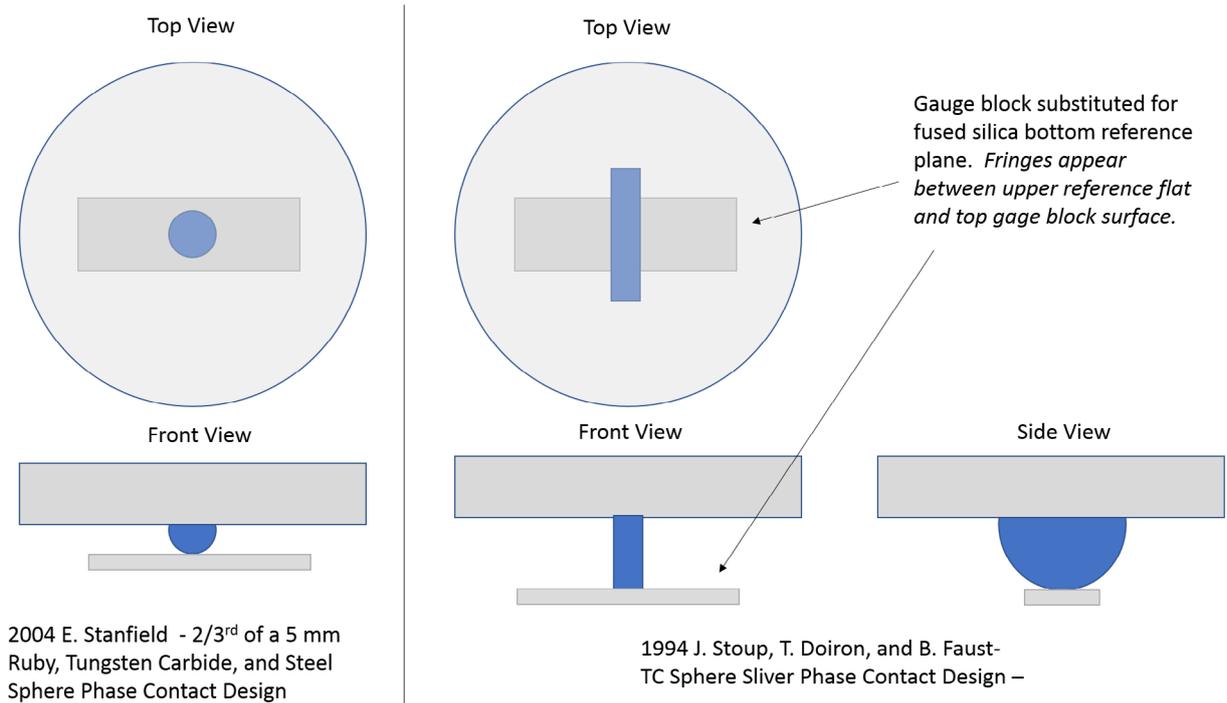


Figure 16. Specially designed spherical contacts used for gauge block and gauge block phase correction measurement.

steel sphere, the expanded uncertainty, U , with a coverage factor of $k = 2$ would be ± 8.9 nm. The uncertainty for the larger spheres is significantly influenced by the sphere material, specifically the three components associated with uncertainty related to the sphere temperature ($u_{\text{Thermometer}}$, u_{CTE} , $u_{\text{Gradients}}$). For materials with lower CTE values, like tungsten

carbide, ruby, silicon nitride, and especially Zerodur⁶ these contributions become much less significant resulting in an expanded uncertainty, U , with a coverage factor of $k = 2$, potentially as low as ± 6.1 nm.

⁶ See footnote 1.

Master spheres measured by the Strang Viewer support ultra-high accuracy characterization of the M48 coordinate measuring machine (CMM) probe tip so that the uncertainty in the diameter of the probe trip is insignificant when volumetric measurements are subsequently performed with the CMM. Spheres measured using the Strang Viewer also serve more directly as reference masters for special artifact calibrations using the M48 CMM, as documented in numerous publications [12–16].

Unique application

In 1998, Stoup, Doiron, and Faust used the Strang Viewer to measure the phase correction for gauge blocks [17]. Their work, used a slice of sphere with wringable flat surface that was mounted/wrung to the underside of the upper reference flat, as shown in figure 16. The flat was created by truncating the sphere, leaving the equator intact to ensure that the equator shadow could be visualized for center determination. Doing this fixed the contact point for the sphere which guarantees the exact same contact point, from setup to setup, unlike the variability of the contact point on a sphere even when marked. With the fixed sphere slice in place, the force extrapolation approach was used to measure the dimension (size or length) of the slice with different lower reference flats. When the dimension of the sphere slice is measured on the Strang Viewer using Fused Silica or Quartz as the lower reference flat, that dimension is known to have 180° phase change upon reflection exactly, with no uncertainty. This measured dimension becomes the reference for our comparison. Next, a gauge block or gauge block platen is substituted as the lower reference flat, in place of the Fused Silica reference flat, and the measurement process repeated. The resulting sphere slice dimension will be different from the reference measurement, assuming the gauge block or gauge block platen is not Fused Silica or Quartz. This difference is the optical length correction resulting from difference in the actual phase correction from that of the Fused Silica reference, at the frequency of the light source used during the measurements. In gauge block length interferometry, we refer to this as the ‘phase correction’.

Since this time, the instrument has experienced significant upgrades, that include the stabilized laser as light source, full automation, and even a redesign of the spherical contact. The configuration of the spherical contact is shown in figure 16. Today, NIST uses two-thirds of a 5 mm diameter sphere and we have made these spheres out of different materials to prove that the same results can be obtained even though the slope of the linearized deformation is different for each material. The phase correction measurement process results associated with these improvements are not quite complete but expect to be published at a later date.

Conclusion

The Strang Viewer is an instrument designed to measure sphere diameter and is based on a Fizeau interferometer. The optical cavity design, the application, and decades of refinement have

resulted in a very high-accuracy instrument for sphere diameter measurement. Its adaptation to the measurement of gauge block and gauge block platen optical phase correction by physical contact is quite unique, in that it is simple and completely independent of the wringing layer limitations of the traditional stack method and pack experiments. The current state of the instrument is a great example of the advantages of automation, which clearly helped us overcome the dominant uncertainty contributions associated with manual operation like thermal drift. Validation and confidence in the method has been verified through the measurement of three spheres using both the Strang Viewer and another absolute measurement system that uses a laser displacement interferometer based NIST developed micrometer for the measurement of cylindrical diameter. The measurement of sphere diameter, although cumbersome, provided results with an estimated expanded uncertainty, U , of ± 20 nm, with a coverage factor of $k = 2$. This comparison was done for several different sized spheres and the worst agreement was 15 nm, ensuring the utmost confidence in both systems. Further validation and confidence of the method, the instrument, and the analyses is expected upon publication of the results of the Consultative Committee for Length (CCL) K-4 intercomparison of diameter standards.

Acknowledgment

We would like to acknowledge Mr. Ralph Veale (NIST retired) for supporting and encouraging efforts to develop and modernize the instrument during his tenure as Group Leader. We would also like to acknowledge the technical contributions of Dr. Lowell Howard (Staff Fellow at the U.S. Food and Drug Administration, formerly at NIST) for developing the automatic force adjustment and image acquisition systems.

References

- [1] Quinn T 2003 Practical realization of the definition of the metre, including recommended radiations of other optical frequency standards (2001) *Metrologia* **40** 103–33
- [2] Haitjema H and Kotte G J 1998 A ball-diameter-measuring instrument in a gauge block interferometer *SPIE Conf. Proc.* **3477** 101–8
- [3] Desiraju K and Hocken R 2005 Interferometric measurement of sphere diameter *ASPE Annual Conf. Poster* (UNCC)
- [4] Jennekens R J 2006 Modification of a ball diameter-measuring instrument for NMI *DCT rapporten; Vol. 2006.059* (Eindhoven: Technische Universiteit Eindhoven)
- [5] Saunders J 1972 Ball and cylinder interferometer *J. Res. Natl Bur. Stand.* **76C** 11–20
- [6] Rolt F H 1929 *Gauges and Fine Measurements* vol 1 (London: Macmillan) ch 11
- [7] Birch K and Downs M 1993 An updated equation for the refractive index of air *Metrologia* **30** 155–62
- [8] Bonsch G and Potulski E 1998 Measurement of the refractive index of air and comparison with modified Edlén’s formulae *Metrologia* **35** 133–9
- [9] Puttock M and Thwaite E 1969 Elastic compression of spheres and cylinders at point and line contact *National Standards Laboratory Technical Paper* No. 25 Division of Applied Physics, National Standards Laboratory, Commonwealth Scientific and Industrial Research Organization (CSIRO)

- [10] Estler W T 1985 High-accuracy displacement interferometry in air *Appl. Opt.* **24** 808–15
- [11] Bruce C F 1955 The effects of collimation and oblique incidence in length interferometers *Aust. J. Phys.* **8** 224–40
- [12] Muralikrishnan B, Stone J, Stoup J and Sahay C 2010 Micro-feature dimensional and form measurements with the NIST fiber probe on a CMM *Cal Lab: Int. J. Metrol.* **17** 25–30
- [13] Stoup J and Faust B 2011 Measuring step gauges using the NIST M48 CMM *J. Meas. Sci.* **6** 66–73
- [14] Stoup J and Doiron T 2003 Measurements of large silicon spheres using the NIST M48 coordinate measuring machine *SPIE Conf. Proc.* **5190** 277–88
- [15] Stoup J and Doiron T 2001 The accuracy and versatility of the NIST M48 coordinate measuring machine *SPIE Conf. Proc.* **4401** 136–46
- [16] Stoup J 2009 A few case studies in uncertainty using the NIST M48 CMM *Conf. Proc. of the ASPE Spring Topical Meeting: Mechanical Measurements and Measurement Uncertainty*
- [17] Stoup J, Faust B and Doiron T 1998 Minimizing errors in phase change correction measurements for gage blocks using a spherical contact technique *SPIE Conf. Proc.* **3477** 161–72
- [18] Stone J A and Zimmerman J H 2011 Engineering metrology toolbox (<https://emtoolbox.nist.gov/Wavelength/Documentation.asp#AppendixAV>)