The new mise en pratique for the metre—a review of approaches for the practical realization of traceable length metrology from $10^{-11}$ m to $10^{13}$ m

To cite this article: René Schödel et al 2021 Metrologia 58 052002

View the article online for updates and enhancements.
Review

The new *mise en pratique* for the metre—a review of approaches for the practical realization of traceable length metrology from $10^{-11}$ m to $10^{13}$ m

René Schödel¹, Andrew Yacoot² and Andrew Lewis²

¹ Physikalisch-Technische Bundesanstalt (PTB), Bundesallee 100, 38116 Braunschweig, Germany
² National Physical Laboratory (NPL), Hampton Road, Teddington, Middlesex, TW11 0LW, United Kingdom

E-mail: rene.schoedel@ptb.de

Received 9 April 2021, revised 7 July 2021
Accepted for publication 14 July 2021
Published 5 August 2021

Abstract

The revised International System of Units (SI) came into force on May 20, 2019. Simultaneously, updated versions of supporting documents for the practical realization of the SI base units (*mise en pratique*) were published. This review paper provides an overview of the updated *mise en pratique* for the SI base unit of length, the metre, that now gives practical guidance on realisation of traceable length metrology spanning 24 orders of magnitude. The review begins by showing how the metre may be primarily realized through time of flight and interferometric techniques using a variety of types of interferometer. Examples of techniques for measuring the interferometric phase and coping when the integer interference order is unknown are then described, together with examples of typical uncertainty contributions that may be encountered. The requirements for traceable nanoscale metrology and the need for an alternative secondary metre as identified by the Consultative Committee for Length’s Working Group on Nanometrology are outlined. These led to the inclusion in the *mise en pratique* of secondary realisations of the length unit at the nanometre and sub nanometre scale, based on the lattice spacing of silicon. Three measurement techniques using this secondary realisation are then described in detail. The paper concludes by emphasising that measurements made today over 24 order of magnitude are still compatible with measurements made using the metre as adopted over 200 years ago.

Keywords: *mise en pratique*, SI, realization of the length, length measurement, silicon lattice, interferometry, speed of light

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

The International System of Units (SI) has been updated and refined over many decades since its foundations were laid at the end of the 18th and beginning of the 19th centuries, culminating in the first formalisation of the SI at the 11th General Conference on Weights and Measures (CGPM) in 1960. Since then, the SI has been updated and revised, adding a 7th base unit for use in chemistry, as well as refining definitions of derived and supplementary units, and providing rules for prefixes. Following recent advances in measurement of fundamental constants, the SI was revised in May 2019, leading to the current SI that is based on a set of seven defining constants, drawn from the fundamental constants of physics and other constants of nature, which is described in the 9th edition of the SI brochure [1]. Accordingly, the supporting documents for the constantsof nature, which is described in the 9th edition of the current SI that is based on a set of seven defining constants, drawn from the fundamental constants of physics and other constants of nature, which is described in the 9th edition of the SI brochure [1].

As with the other MeP, the fundamental revision of the MeP for the metre was led by the idea that the realization of an SI unit is related to primary measurement methods at the highest level. While, for good readability, some of the explanations contained in the new MeP for the metre [3] are repeated here, this paper is aimed at providing an extended background for different basic realization methods for the length unit.

The current SI definition of the metre is: ‘the metre, symbol m, is the SI unit of length. It is defined by taking the fixed numerical value of the speed of light in vacuum $c$ to be 299 792 458 when expressed in the unit m s$^{-1}$, where the second is defined in terms of the caesium frequency $\nu_{\text{Cs}}$’ [4]. It ensures continuity of the SI base unit of length and implies that ‘the metre is the length of the path travelled by light in vacuum during a time interval of $1/299 \, 792 \, 458$ of a second’, as stated in the previous definition of the metre from 1983 [5]. The revised MeP for the metre [3] does therefore provide the link to primary measurement methods resulting in a ‘path travelled by light’ from the light travel time. The fundamental equation underlying the above definition of the metre is a direct relationship between the speed of light $c$, a time interval $\Delta t$ and a length $l$, in vacuum:

$$l = c \cdot \Delta t. \quad (1)$$

As any previous length definition, including the earlier definition of the metre by the international prototype [6], the ‘path travelled by light’ represents a geometrical length rather than an optical path length, the latter being the product of the geometrical length and the refractive index of the medium through which the light propagates. Only in vacuum, to which, in a strict sense, the metre definition refers, does the discrepancy between the geometrical length and the optical length vanish. On the other hand, primary calibrations of lengths and distances most commonly need to be made in air which reduces the speed of light. Therefore, in the revised MeP for the metre [3] it is noted as a key point that the exact determination of the air refractive index, $n$, is of major importance. A distinction is necessary between $c$, the speed of light in vacuum and $c'$, the speed of light in general ($c' = c/n$):

$$l = c' \cdot \Delta t.$$ \hfill (2)

Under atmospheric pressure, the air refractive index reduces the speed of visible light with a relative effect of the order of $2.8 \times 10^{-5}$ ($n \approx 1.00028$), corresponding to about 280 $\mu$m per metre of measured length. In many cases, e.g., when using modulated light, it is important to consider the group refractive index of air, $n_g$, instead of the (phase-) refractive index, $n$:

$$l = c' \cdot n_g \cdot \Delta t.$$ \hfill (3)

This is particularly important for light pulses that can be viewed as a wave packet that travels as a unit. Generally, wave packets consist of an infinite set of sinusoidal waves interfering constructively only over a small region of space, and destructively elsewhere. The group refractive index can be simply derived from the dispersion relation resulting in the following expression:

$$n_g = n - \lambda_0 \cdot \frac{dn}{d\lambda_0} = n + f \cdot \frac{dn}{df}, \quad (4)$$

in which $f$ denotes the light frequency and $\lambda_0$ the so-called vacuum wavelength derived from it ($\lambda_0 = c/f$). Considering as an example green light ($\lambda \approx 520$ nm) in atmospheric air, $n_g - n$ is of the order of $10^{-5}$. This difference is comparable in size to the variation of the phase refractive index of air over the entire range of visible light ($n(380 \text{ nm}) - n(780 \text{ nm}) \approx 9 \times 10^{-6}$). Thus, using $n$ instead of $n_g$ can cause an error of about 10 $\mu$m per metre, which demonstrates the significance of $n_g$. Ultimately, primary methods for the practical realization of the metre must refer to the exact speed of light in air which is either $c' = c/n$ (equation (2)) or $c_g = c/n_g$ (equation (3)).

2. Primary methods for the practical realization of the definition of the metre

The practical approach to explore the travel time of light waves is significantly different from that of other waves. For example, while the propagation of sound waves can be tracked using various signal detectors along the wave propagation, this concept is inappropriate for light due to the speed of light itself, which is also the highest possible signal speed. Therefore, the approach to a setup for measuring the light travel time is generally based on dividing the light into two separate pathways, a reference path and a measuring path.

2.1. Direct measurement of the light’s travel time (time of light measurement)

The direct measurement of the travel time of light requires modulation of the light intensity in order to generate fiducial
features. The most vivid modulation is represented by a light pulse, e.g., a laser light pulse which is formed from a wave packet. The incident light pulse shown in figure 1 is split into two pulses, one of which travels a short reference path, the other one travels the measuring path. The mirrors in both pathways are arranged such that the light is retroreflected. After the second passage through the beam splitter the light pulse originating from the short reference path first hits a light detector and sets a first trigger at a defined threshold, defining a reference point in time. A second trigger is then generated by the delayed light pulse originating from the measurement path.

From the time delay $\Delta t$ between both triggers, the length difference between measurement and reference pathways is derived, which represents the length $\Delta z = l = \frac{1}{2}c_0 \cdot \Delta t$. For example, if the length, $l$, of 1 m, is traversed twice by the light pulse (outwards and return directions), then the time delay is approximately 6 ns. Instead of utilizing light pulses, representing a wave packet, amplitude-modulated light can be used to achieve higher resolutions. With this approach, the reference mirror (see figure 1) is omitted and the phase of the intensity modulation is determined electronically from the detector signal. The travel time for the light along the measuring path is then the difference $\Delta \phi_{\text{mod}} = \phi_{\text{mod}} - \phi_{\text{mod}}$, between the measured intensity modulation phases at the detector, $\phi_{\text{mod}} = 2\pi f_{\text{mod}} \cdot t$, and before travel, $\phi_{\text{mod}} = 2\pi f_{\text{mod}} \cdot t$, which leads to:

$$\Delta t = \frac{\Delta \phi_{\text{mod}}}{2\pi f_{\text{mod}}},$$

in which $f_{\text{mod}}$ is the modulation frequency. High modulation frequencies of the order of several gigahertz are used in sophisticated devices using external electro-optic [7] or electro-absorption modulators [8]. Using femtosecond lasers as a light source, multiple highly stable modulation frequencies from several hundred megahertz up to several tens of gigahertz can be generated in parallel [9].

The direct measurement of the light’s travel time is largely used for measurements at geodetic scales. Here, the group refractive index of air, its homogeneity and invariance, are the limiting factors of the attainable measurement uncertainty. It is also used for even larger distances, e.g., from the earth to the Moon [10]. For the realization of short lengths interferometrical methods are more effective and generally more accurate.

### 2.2. The basic idea of indirect measurement of light travel time by utilizing interferometric techniques

For the realization of lengths below a few metres, as well as for the most accurate realization of length in general, interferometric techniques are far superior. Monochromatic light can be considered as an electromagnetic wave, the electric field of which is propagating along the measurement pathway (defined as z-direction):

$$E(z, t) = A \cdot \cos(\phi) = A \cdot \cos(k \cdot z - \omega \cdot t + \delta)$$

in which $A$ is the amplitude, $\phi$ the phase, $k$ the wave number, $\omega$ the angular frequency, $t$ the time, and $\delta$ the initial phase. The relationship between the parameters $k$ and $\omega$ with wavelength $\lambda$ and frequency $f$ is given by $k = 2\pi / \lambda$ and $\omega = 2\pi \cdot f$. Wavefronts travel the distance of a single wavelength during a single oscillation period $T$ ($T = 1/f$). The frequency of visible light waves lies in the range from 300 THz to approximately 600 THz. The respective oscillation period is therefore extremely short and, because the upper frequency limit of detectors is too low, the oscillation period could not be recorded directly by any real detector. The only measurable parameter of a single light wave is a mean intensity $\langle E^2 \rangle$. The infinite average $(t \rightarrow \infty)$ of the latter, $\lim_{t \rightarrow \infty} \frac{1}{T} \int_0^T (E(t, z))^2 \, dt$ results in the constant value of the light intensity: $I = \frac{1}{2}A^2$.

Two waves can be brought into interference with the simplest kind of interferometer shown in figure 2, left, which is basically the same arrangement as in figure 1. The amplitude ratio of the two waves, which interfere when leaving the

---

**Figure 1.** Scheme for the direct measurement of the light travel time which is determined from the time delay between light pulses travel pathways of different lengths before reaching a detector.
interferometer, depends on the reflectivity of the beam splitter and it can be assumed that both waves have the same initial phase and frequency. However, the phase of the light wave $E_2(z, t) = A_2 \cdot \cos [k \cdot (z + 2\Delta z) - \omega \cdot t + \delta]$ which double passes the measurement path is enlarged by $k \cdot 2\Delta z$ compared to the wave $E_1(z, t) = A_1 \cdot \cos [k \cdot z - \omega \cdot t + \delta]$ traversing the reference path. The sum of both waves represents the interference wave which can be written as:

$$E_{12} = E_1 + E_2 = A_{12} \cdot \cos (k \cdot z - \omega \cdot t + \delta_{12}).$$  

(7)

The amplitude of the interference wave is given by $A_{12} = \sqrt{A_1^2 + A_2^2 + 2A_1 \cdot A_2 \cdot \cos (2k \cdot \Delta z)}$ and the phase shift by $\delta_{12} = \arctan 2(A_1 \cdot \sin (2k \cdot \Delta z), A_1 \cdot \cos (2k \cdot \Delta z) + A_2)$. The infinite average of $\langle I_{12}^2 \rangle_{t \rightarrow \infty}$, which represents the intensity, is thus given by:

$$I_{12} = \frac{1}{2} A_{12}^2 = I_0 \left( 1 + \gamma \cos \left( 2\pi \cdot \frac{\Delta z}{\lambda/2} \right) \right),$$  

(8)

where $\gamma$ denotes the interference contrast $\gamma = 2\sqrt{I_1 I_2} / (I_1 + I_2) = (I_{\text{max}} - I_{\text{min}}) / (I_{\text{max}} + I_{\text{min}})$ and $I_0 = I_1 + I_2 = (A_1^2 + A_2^2)/2$.

As shown in figure 2, there is a relationship between the length of the measurement path and the number of oscillations, the so-called interference order, due to the phase change:

$$\frac{\Delta \varphi}{2\pi} = \frac{\Delta z}{\lambda/2}.$$  

(9)

Viewed from the other side, measurement of $\Delta \varphi / 2\pi$ allows determination of $\Delta z$ based on the cumulative phase change. In the simplest case the integer interference order, i.e. the integer part of $\Delta \varphi / 2\pi$, is obtained by counting the number of intensity periods while continuously shifting the measurement mirror.

$$a \cdot \cos (x + \alpha) + b \cdot \cos (x + \beta) = \sqrt{a^2 + b^2 + 2ab \cdot \cos (\alpha - \beta)} \cdot \cos (x + \delta),$$  

in which $\delta = a \tan 2(a \sin \alpha + b \sin \beta, a \cos \alpha + b \cos \beta)$ and $a \tan 2$ represents the extended arctangent.

In any case, the size of shift of the measurement mirror, i.e. a length, is measured as an arithmetic product of half of the light wavelength and the interference order:

$$l = \Delta z = \frac{1}{2} \frac{\Delta \varphi}{2\pi} = \frac{1}{2} \cdot \frac{\Delta \varphi}{\omega}.$$  

(10)

Accordingly, equation (10) reveals (see equation (2)) that the travel time of light is obtained from:

$$\Delta t = \frac{\Delta \varphi}{\omega} = \frac{\Delta \varphi}{2\pi f}.$$  

(11)

Again (for comparison see equation (5)), the travel time of light results from the quotient of a phase difference and a frequency.

Equation (11) clearly reveals that the indirect measurement of the travel time of light requires measurement of the following quantities:

(a) The frequency $f$ of the light;
(b) The phase difference $\Delta \varphi$ between the two interfering waves resulting from the observation of the intensity of interference using an interferometer.

2.3. Measurement of the frequency of the light

As equations (10) and (11) reveal, knowledge of the frequency of the light, $f$, is an essential requirement for the realization of the unit of length. It provides the scaling factor between a measured phase difference and the length that is realized by interferometry. For the highest demands on the accuracy of the light frequency, a light source could be synchronized to the primary frequency standards by an appropriate technique.

Generally, the amount of the frequency difference $|\Delta f|$ of a light source under test (wave $E_2(t) = A_2 \cdot \cos ((k + \Delta k) \cdot z - (\omega + \Delta \omega) \cdot t + \delta_2)$) with respect to the frequency of a reference source (wave $E_1(t) = A_1 \cdot \cos (k \cdot z - \omega \cdot t + \delta_1)$), is measured also by the interference detection technique. In this approach
the two light waves are brought into coincidence along the same path direction of propagation. In analogy to equation (7), the associated interference wave can be written as 

\[ E_{12}(t) = A_{12}(t) \cdot \cos (k \cdot z - \omega \cdot t + \delta(t)) \]

in which \( A_{12}(t) \) represents the oscillating amplitude 

\[ A_{12}(t) = \sqrt{A_1^2 + A_2^2 + 2A_1 \cdot A_2 \cos (\Delta k \cdot z - \Delta \omega \cdot t + \delta_2 - \delta_1)} \]

and \( z \) the position of the detector. Again, the high frequency term of \( E_{12}(t) \), which includes the light frequency itself, cannot be acquired with any detector. Therefore, the measurable interference intensity is given by \( I_{12}(t) = \frac{1}{2}A_{12}^2(t)^2 = I_0(1 + \gamma \cos (\Delta k \cdot z - \Delta \omega \cdot t + \delta_2 - \delta_1)) \). For the purpose of further considerations, made in the following paragraphs, \( \Delta k = k_1 - k_2 \) is expressed as \( \frac{1}{2}(n_1 \cdot \omega_1 - n_2 \cdot \omega_2) \) and \( n_2 = n_1 + (\omega_2 - \omega_1) \cdot \frac{ df}{dz} \). This results in:

\[
I_{12}(t) = \frac{1}{2}A_{12}(t)^2 = I_0 \left( 1 + \gamma \cos \left( -2\pi \cdot \Delta f \cdot \left( \frac{1}{c}n_g \cdot z \right) + \delta_2 - \delta_1 \right) \right),
\]

in which \( n_g \) is the group refractive index. Equation (12) represents the so-called beat signal from which \( |\Delta f| \) can be derived electronically. The frequency of a light source (\( E \)) under test (\( f_{\text{EUT}} \)) is therefore related to the frequency of a standard: 

\[ f_{\text{EUT}} = f_{\text{standard}} \pm f_{\text{beat}}. \]

To achieve unambiguity, it is necessary to have more than one standard frequency light source available, or to be able to vary the standard frequency in a known direction, e.g. by reducing the frequency by a small amount and observing the change in \( f_{\text{beat}} \).

As an alternative to direct measurement of the light frequency, one may use one of the values from the list of ‘recommended values of standard frequencies for applications including the practical realization of the metre and secondary representations of the second’ [11]. This list is updated periodically by recommendation of new candidate standard frequencies. Candidate frequencies are examined according to a published set of guidelines and procedures [12] and only those that pass the necessary checks are recommended to the CIPM for entry into the list.

From the calibrated frequency of the light source, the wavelength \( \lambda_0 \) can be calculated which in vacuum is generally assumed to be identical to the path length corresponding to a single order of interference (\( l = \frac{1}{2}\lambda \)). While this assumption is correct for an ideal interference of two plane waves, it does not apply if the light is not perfectly monochromatic, or for non-planar waves (see section 4.5 for the corresponding effect to the measurement uncertainty).

3. Measurement of the phase difference \( \Delta \varphi \) in various interferometric approaches

As described above, the basic principle of distance measuring interferometers is the observation of the periodically changing detector signal of the interference intensity during a change of the distance of a measuring mirror with respect to a stable reference mirror, (see figure 2). Each period corresponds to a unitary change in the interference order, i.e. to a change in distance by half the wavelength of the light used (equation (10)). In such interferometers, the moving mirrors are usually positioned on carriages. To cope with angular changes during the movement, retroreflectors are typically used instead of flat mirrors which makes the observed interference signal insensitive to small tilts. Over the distance to be measured a huge change in the interference order is counted, e.g. approximately 3 million oscillations per metre when using a typical red laser (\( \lambda \approx 633 \text{ nm} \)). Due to the large coherence length of today’s laser light sources, it would be theoretically possible to measure distances in the kilometre range. However, air turbulences make counting of integer interference order changes practically impossible at larger distances, especially in uncontrolled environments such as outdoors. A further limit is the mechanical stability of the interferometer. Generally, an interferometer using a single wavelength is called a homodyne interferometer.

3.1. Homodyne interferometers with quadrature detection scheme for measuring the phase difference

With the arrangement shown in figure 2, the direction in which the measuring mirror is moved cannot be detected (see equation (8)). In extreme cases, moving the measurement mirror forward and backward can even simulate uniform motion along the \( z \)-axis. For this reason, most homodyne interferometers are equipped with additional components, using the ‘quadrature procedure’ in order to establish a unique relationship to the position of the measuring mirror. As shown in figure 3, a linearly polarized laser beam whose polarization axis is tilted by 45° hits a polarizing beam splitter which splits the incident light into the two polarization components that are perpendicular to each other. While the component which is polarized parallel to the incident plane (p-beam) entirely passes through the beam splitter, the vertically polarized component (s-beam) is fully reflected. These two beams are reflected and then brought together at the polarizing beam splitter. For the generation of interference between the two mutually perpendicular components s and p polarizers are required, the axis of which is tilted by 45°.

Additionally, the beam which is transmitted at the beam splitter traverses a \( \lambda/4 \) delay plate before the polarizer which causes a mutual phase shift of \( \pi/2 \). Thus, the signal at this detector is shifted by \( \pi/2 \) compared to the signal of the other detector. The quadrature detection thus provides two interference signals that are shifted by \( \pi/2 \) with respect to each other and are thus called sine and cosine signals, respectively. The signals \( I_{\text{sin}} \) and \( I_{\text{cos}} \) can be represented as shown in figure 3 on the right, e.g. by means of an oscilloscope in the \( xy \) mode. A shift of the measuring mirror then leads to a circle being displayed. Depending on the direction of displacement, the corresponding vector rotates in one or the other direction. Hereby, a full revolution corresponds to one integer interference order (\( \Delta \varphi = 2\pi \)) and thus to a shifting of the measuring mirror by \( \Delta z = \lambda/2 \), which can be clearly detected with this procedure. While the number of revolutions can be counted electronically, in this approach the fractional part of the phase difference \( \Delta \varphi \) is represented by the angle \( \varphi \) = arctan 2 \((I_{\text{sin}}/I_{\text{cos}})\) (see figure 3, right). In reality unavoidable imperfections, e.g. unequal
detector gains and offsets can cause distortions of the circle resulting in an ellipse that can be ‘corrected’ back to a circle in the post processing by a Heydemann or similar correction (see [13–15] and references therein). Incomplete separation of the differently polarized partial beams in a polarizing interferometer as well as multiple back reflections, so-called optical mixing errors also produce periodic non-linearities (see [16–19] and references therein).

3.2. Heterodyne interferometers

Laser sources can be designed to emit two or more coherent light waves of slightly different frequency that are polarized perpendicular to each other. For He–Ne lasers, this is often achieved by external magnetic fields inducing the Zeeman effect [16], leading to frequency differences of typically 1 MHz to 3 MHz. Larger frequency differences can be generated by suitable laser resonator designs [20]. Furthermore, it is possible to shift the frequency of a light wave by a defined value, e.g. by means of acousto-optic modulators. Figure 4 shows a scheme of such an interferometer known as a heterodyne interferometer. The laser light source used here generates two waves with different frequencies: \( E_1(t) = A_1 \cdot \cos(k_1 \cdot z - \omega_1 \cdot t + \delta_1) \) and \( E_2(t) = A_2 \cdot \cos(k_2 \cdot z - \omega_2 \cdot t + \delta_2) \) that are polarized perpendicular to each other (‘s’ indicates vertical polarization with respect to the image plane and ‘p’ parallel polarization). First, both light waves are split using a non-polarizing beam splitter. The two reflected waves then hit a polarizer, the axis of which is inclined by 45° with respect to the two polarization directions. As a result, the now equally polarized components of the interference waves \( E_1 \) and \( E_2 \) lead to a beat signal at the reference detector (see derivation of equation (12) above):

\[
I_{\text{ref}} (t) \propto 1 + \gamma \cos \left( - (\omega_2 - \omega_1) \cdot \left( t - \frac{1}{c} n_g \cdot z_{\text{ref}} \right) + \delta_2 - \delta_1 \right),
\]

in which \( z_{\text{ref}} \) is the length of pathway travelled by \( E_1 \) and \( E_2 \) to the reference detector. The two light waves that pass through the non-polarizing beam splitter hit a polarizing beam splitter where \( E_1 \) is reflected and \( E_2 \) is transmitted and pass the measurement path. After both waves have been reflected by the retroreflectors, they are combined after the second pass of the polarization beam splitter. After passing through a polarizer, the axis of which is inclined by 45° in relation to the two polarization directions, a second beat signal is generated at the measuring detector:

\[
I_{\text{meas}} (t) \propto 1 + \gamma \cos \left( k_2 \cdot 2 \Delta z \cdot (\omega_2 - \omega_1) \cdot \left( t - \frac{1}{c} n_g \cdot z_{\text{meas}} \right) + \delta_2 - \delta_1 \right),
\]

in which \( z_{\text{meas}} \) is the length of pathway of \( E_1 \) via the reference mirror to the measuring detector and \( \Delta z \) is again the length difference between the distances from measuring mirror and reference mirror, respectively, to the beam splitter \( (\Delta z = z_{\text{meas}} - z_{\text{ref}}) \). Considering equations (13) and (14), the difference between the phases of \( I_{\text{meas}} \) with respect to \( I_{\text{ref}} \), which can be measured electronically, is represented by:

\[
\varphi_{\text{meas}} - \varphi_{\text{ref}} = k_2 \cdot 2 \Delta z + \varphi_0,
\]

in which \( k_2 = \frac{2}{c} \omega_2 \cdot n_g \) represents the wave number of \( E_2 \) and \( \varphi_0 = (\omega_2 - \omega_1) \cdot \frac{1}{c} n_g \cdot (z_{\text{meas}} - z_{\text{ref}}) \) can be considered as a constant value, provided that positions of detectors and the reference mirror, but also \( n_g \) are unchanged. Therefore, when considering a movement of the measuring mirror in a distance measurement, the length of the distance is again obtained from equation (10). The effect of non-linearities in heterodyne displacement interferometry and how to reduce them is widely discussed in the literature (e.g. see [16, 21] and references therein).

3.3. Frequency sweeping interferometers (FSI)

The larger the length to be measured, the more likely that beam interruptions can occur, the more impractical the movement of a measuring mirror over a distance becomes. The relationship
between a path length and the phase difference via equation (10) can be considered inversely, i.e. $\Delta \varphi = 2\pi \cdot n \cdot l = \frac{4\pi}{c} \cdot n \cdot l$. When the frequency of the light source can be altered, e.g. when using a suitable tunable laser, that allows a time dependent frequency $f(t)$, phase differences are observed as a function of the light frequency. In this approach, from the measured slope of $\frac{\Delta \varphi}{\Delta f} = \frac{2\pi}{c} \cdot l \cdot \left( n + f \cdot \frac{\partial f}{\partial t} \right) = \frac{4\pi}{c} \cdot l \cdot n_g$ the length is determined:

$$l = \frac{1}{4\pi} \cdot \frac{c}{n_g} \cdot \frac{\Delta \varphi}{\Delta f},$$

in which $n_g$ is the group refractive index of air (see equation (4)). While various concepts for frequency sweeping interferometer (FSI) [22–26] form the basis of versatile length measurements, the uncertainty in the resulting length is typically much larger than one-quarter of a wavelength. Critical to the traceability of FSIs is the determination of the rate of frequency change with a link to the SI second. This can be performed by simultaneously observing the laser output through an absorption medium which offers quantum transitions at well-known frequencies, such as the spectra of acetylene and HCN in the infra-red part of the spectrum. Currently the CIPM has approved a set of 111 recommended frequencies for transitions in various species of the acetylene molecule [11] at around 19 THz (1.5 μm wavelength) with uncertainties at the level of a few parts in $10^{11}$, and work is underway internationally to secure agreement on similar transitions in HCN [27].

### 3.4. Multiple-beam interferometers

Figure 5 shows a diagram for a multi-beam interferometer, in which the interference of the light is observed in reflection. After passing through the beam splitter, the light passes through a semi-transparent optical plate. Some of the light reflected by the mirror is reflected by the semi-transparent surface and goes back to the mirror. The higher the reflectivity $R$ of the semi-transparent surface, the more often this sequence is repeated.

After passing through the beam splitter, the interference wave resulting from the multi-beam interference hits a detector. The amplitude of the interference wave can be derived from the infinite sum of partial light waves (see, for example, [28]). As in the case of the two-beam interference, the associated interference intensity is a periodic function of $\Delta \varphi = \Delta c / (\frac{1}{2} \lambda)$, i.e. the periodicity is given by half the wavelength of the light used. The interference intensity, the formula of which is shown in figure 5, leads to an increasingly sharper structure with increasing reflectivity $R$ of the semi-transparent plate (see insert in figure 5).

A multiple-beam interferometer is called a ‘Fizeau interferometer’ if the distance between the two flat surfaces on which the reflection takes place is large compared to the wavelength of the light. An advantage of such interferometers is the spatial narrowness of the fringe minima which can be accurately localised using e.g. CCD detectors. On the other hand, the dependency of the interference signal on the path difference is more complicated for Fizeau interferometers. With an absorbing film used as a beam-splitter, the interference profile of the reflected light will even be asymmetric.

A Fabry–Pérot interferometer (FPI) is a Fizeau interferometer operated with two semi-transparent mirrors whose reflectivity $R$ is typically large. The so-called ‘finesse’ serves to characterize the resonator. It is defined as the ratio of the so-called free spectral range $\Delta \lambda$ to the full width at half maximum $\delta \lambda$ of an individual maximum of the interference intensity: $F = \Delta \lambda / \delta \lambda = \pi \sqrt{R} / (1 - R)$. Extremely stable resonators with high finesse can be used for short-term frequency stabilization of laser sources. FPI resonators made of monocrystalline silicon, for example, enable frequency stabilization of commercial laser systems of better than $10^{-16}$ [29]. In a limited range of the length FPIs can also function as displacement sensors [30]. Suitable FPIs can also be used to accurately determine the refractive index of air [31].

### 3.5. Large-field imaging interferometers

The interferometers described above have beams with small diameter, typically a few millimetres, however, there is another class of interferometers, namely large field interferometers which are equipped with optical components to generate beam diameters of about 50 mm or larger. Besides collimating lenses, with a focal length of typically several hundred millimetres, comparatively large beam splitter and mirrors...
are necessary. Such interferometers are often combined with imaging optics, which make it possible to locate objects (e.g. the front of a gauge block) in relation to the pixel coordinates of a camera sensor. These large-field imaging interferometers can be designed as two-beam interferometer or multi-beam interferometer (Fizeau interferometer). These are suitable for the measurement of the length of bar shaped material measures, e.g. gauge blocks [32] wrung to a flat platen forming a step height. Although the basic physical principles of modern large field interferometers are unchanged for more than a century [33–36], today’s interferometers benefit from highly monochromatic stabilized lasers, high dynamic range cameras, computer-controlled equipment including phase stepping and sophisticated software analysis (e.g. see [37]). Further, modern light sources can be stabilized to frequency standards by using optical frequency comb techniques (e.g. see [38]). The basic scheme of a two-beam large field interferometers is shown in figure 6. A point light source located in the focus of a collimator creates a large bundle of parallel rays that covers the prismatic body, such as a gauges block, as shown in figure 6 along with its end platen. In this way, the interference intensities for each partial beam of the bundle are laterally resolved:

\[ I = I_0 \left\{ 1 - \gamma \cos \left( \frac{2\pi}{\lambda/2} (z_2(x,y) - z_1(x,y)) \right) \right\}, \quad (17) \]

where \( z_{1/2}(x,y) \) represents the distribution of the geometric paths (1: reference, 2: measurement) perpendicular to the optical axis. Recall \( \gamma \) from equation (8). If the reflecting surfaces were completely flat, the path difference \( z_2(x,y) - z_1(x,y) \) would describe a plane whose tilt depends on the alignment of the surfaces with one another.

The cosine function in the equation is the main term responsible for the appearance of the fringes with maxima for path differences of \( m\lambda/2 \) and minima for \( (m + \frac{1}{2}) \lambda/2 \), where \( m \) is an integer. The prismatic body in the measuring arm of the interferometer represented in figure 6 causes the occurrence of two offset fringe systems. The value of this offset in relation to the fringe spacing corresponds to the length, \( l \), of the artefact, but only the fractional interference order, \( q \), is visible. Together with the integer interference order, \( i \), which must be determined separately, the length of a prismatic body can then be obtained:

\[ l = \frac{\lambda}{2} \cdot (i + q). \quad (18) \]

The sharpness of the interferograms is affected by the finite size of the aperture, which is positioned in the focal plane of the focusing lens at the output of the interferometer. This aperture serves to suppress disturbing secondary reflections. The application of digital processing techniques [39, 40] to phase shift interferometry has significantly improved measurement efficiency and accuracy. The basic concept of phase shift interferometry is that one can determine the phase modulo \( 2\pi \) over a beam diameter by acquiring multiple interferograms, each phase shifted by a certain amount, wherein the phase shift might be generated by variation of the interferometer’s reference pathway. Considerable efforts have been made to develop effective and error-compensating phase extraction algorithms [41–47]. The algorithm proposed by Tang [47] is based, for example, on five intensities \( I_k = I_0 \{ 1 - \gamma \cos (k - 3) \alpha \} \), \( k = 1 \ldots 5 \), which are recorded successively with equally spaced phase steps \( \alpha \), which leads to the phase difference:

\[ \Delta \varphi = \arctan \left( \frac{\sqrt{2(I_4 - I_5) + (I_1 - I_5)[2(I_4 - I_5) - (I_5 - I_0)]}}{I_1 + I_5 - 2I_3} \right), \quad (19) \]

as illustrated in figure 7.

For Fizeau interferences there exists a dedicated phase determination algorithm [48] which takes into account the non-sinusoidal intensity of fringes in Fizeau interferometers.

The installing of a permanently evacuated cell in the measuring arm of a large-field interferometer allows accurate determination of the refractive index of air for the specific
wavelength of the light used. Here, the refractive index of air is determined from the difference of the optical path lengths in vacuum and air, along a common cell length [49]. If the collimated beam of a large-field interferometer is large enough to contain both a gauge block and a vacuum cell next to it, the effective refractive index of the air can be determined from a common phase topography. A rough estimate of the air refractive index will suffice [50].

3.6. Extension of the unambiguity range

When the integer interference order cannot be determined, e.g. in the measurement of long distances, the ambiguity range is limited to about a quarter of the wavelength. The range of unambiguity can be enlarged by applying more than a single light source, with a different wavelength in a length measurement by interferometry.

Considering the simplest case of two different light sources, for a given length \( l \) the respective phase differences can each be written as \( \Delta \varphi_i = 2\pi \cdot l / (\lambda_i) \) (see equation (10)). Thus, the difference \( \Delta \varphi_2 - \Delta \varphi_1 \), which is called the synthetic phase \( \Phi_{\text{synth}} \), corresponds to

\[
\Phi_{\text{synth}} = \Delta \varphi_2 - \Delta \varphi_1 = 2\pi \cdot l / \Lambda_{\text{synth}},
\]

in which \( \Lambda_{\text{synth}} \) denotes the so-called synthetic wavelength (\( \Lambda_{\text{synth}} = \lambda_1 \cdot \lambda_2 / (\lambda_1 - \lambda_2) \)).

Consequently, the length can be rewritten as

\[
l = \frac{\Lambda_{\text{synth}} \cdot \Phi_{\text{synth}}}{2\pi}. \tag{21}
\]

Compared to equation (10), the unambiguity range in a length measurement that is based on equation (21) is greatly enlarged due to the fact that the synthetic wavelength is much larger than a single wavelength. On the other hand, the accuracy of \( \Phi_{\text{synth}} \) is limited by the accuracies of the individual phase differences \( \Delta \varphi \). Therefore, the uncertainty of the length is generally up-scaled by a factor of approximately \( \lambda / \Delta \lambda \).

Application of such ‘synthetic wavelength interferometry’ in air benefits from a separate consideration of the air refractive index. Continuing the above example of two wavelengths, each written as \( \lambda_k = \lambda_0 \cdot n_k \), in which \( \lambda_0 \) denote so called vacuum wavelengths (\( \lambda_0 = c / f_k \)), the synthetic phase corresponds to

\[
\Phi_{\text{synth}} = 4\pi \cdot l \cdot \frac{\lambda_0^2 - \lambda_1^0}{\lambda_1^0 \cdot \lambda_2^0} \left( n_1 - \lambda_0^0 \cdot n_2 / \lambda_2^0 - \lambda_1^0 \right). \tag{22}
\]

For small differences \( \Delta \lambda^0 = \lambda_2^0 - \lambda_1^0 \) it can be easily demonstrated that the term \( n_1 - \lambda_0^0 \cdot (n_2 - n_1) / (\lambda_2^0 - \lambda_1^0) \) is close\(^6\) to the group refractive index \( n_g \) (see equation (4)) considered at the average vacuum wavelength \( \lambda_0^0 + \frac{1}{2} \Delta \lambda^0 \).

\(^6\) Agreement is within \( 10^{-8} \) for wavelengths \( > 450 \) nm and \( \Delta \lambda^0 < 10 \) nm.
Accordingly, the length can then be rewritten as

\[ l = \frac{1}{2} \frac{\Lambda^0_{\text{synth}}}{\delta g} \frac{\Phi_{\text{synth}}}{2\pi}, \tag{23} \]

in which \( \Lambda^0_{\text{synth}} \) denotes the synthetic vacuum wavelength \( (\Lambda^0_{\text{synth}} = \lambda_1^0 + \lambda_2^0 / (\lambda_2^0 - \lambda_1^0)) \). Such two-colour synthetic wavelength interferometry is suitable for absolute measurement of even very large distances (e.g. see [51, 52] and references therein).

While measurement of the synthetic phase and consideration of the synthetic wavelength can provide appropriate estimates in a length measurement [53], the accuracy of the length according to equation (23) is typically larger than a quarter of a wavelength.

The exact total number of integer interference orders can be determined by the so-called ‘method of exact fractions’ which was first suggested by Benoit [54]. Here, the phase difference, \( \frac{1}{\lambda} \Delta \varphi \), representing the total interference order, is split into an integer part, \( i \), and a fractional part \( q \), i.e.

\[ l = \frac{1}{\lambda} \cdot (i + q). \]

When different wavelengths, \( \lambda_k \), are available for each wavelength \( l \) is expressed as:

\[ l_k = (i_k + \delta_k + q_k) \lambda_k / 2, \tag{24} \]

in which \( i_k \) represent integer orders, calculated from an estimate of the length, \( \text{est} \), as \( i_k = \text{round} (N \frac{l}{\lambda_k} - \frac{1}{4} \lambda_k) \), where ‘round’ represents the operator for rounding to the nearest integer) and \( \delta_k \) are variation integers. For a set of \( N \) different wavelengths the mean length, \( \bar{T} \), and the average deviation from the mean length, \( \Delta \), can be considered as described in more detail in [50]:

\[ \bar{T} = \frac{1}{N} \sum_{k=1}^{N} l_k, \quad \Delta = \frac{1}{N} \sum_{k=1}^{N} |T - l_k|. \tag{25} \]

The resulting set \( \{\bar{T}, \Delta\} \) can be displayed as scatterplot which represents a coincidence pattern. Considering the case of two wavelengths, the value of \( \Delta \) is exactly given by \( \frac{1}{2} |l_1 - l_2| \). Assuming that \( \delta_1 \) and \( \delta_2 \) correspond to the correct integer orders \( (\delta_1 = 0, \delta_2 = 0) \), the two lengths, written as \( l_1 = (i_1 + q_1) \lambda_1 / 2 \) and \( l_2 = (i_2 + q_2) \lambda_2 / 2 \), coincide \( \Delta = 0 \) when \( q_1 \) and \( q_2 \) are given ‘exactly’. A relation for further variation integers \( (\delta_1 \neq 0, \delta_2 \neq 0) \) for which coincidence is obtained can be extracted by setting \( \Delta = 4 |\bar{T}_1 - \bar{T}_2| = \frac{1}{\lambda} \frac{|(i_1 + q_1) \lambda_1 / 2 - (i_2 + q_2) \lambda_2 / 2|}{2} = 0 \), in which \( \bar{T}_1 \) and \( \bar{T}_2 \) correspond to blunder lengths. This situation is expressed with equation (26):

\[ \Delta = 0 = |\lambda_1 \delta_1 - \lambda_2 \delta_2| \Rightarrow \Delta = 0 \text{ for } \frac{\delta_2}{\delta_1} = \pm \frac{\lambda_1}{\lambda_2}. \tag{26} \]

Therefore, integer interference orders could be calculated incorrectly when their ratio is equal to the ratio of the wavelengths used in the measurements, i.e., a rational number. In particular, the issue occurs when the first wavelength is given by an integer multiple of the second wavelength, i.e. \( \lambda_1 = a \lambda_2 \) \((a \text{ is an integer})\), coincidence \( \Delta = 0 \) is obtained when \( \delta_2 = a \delta_1 \).

Assuming the two fictive wavelengths 532.3 nm and 548 nm, the example coincidence pattern shown in figure 8 demonstrates that at the length deviation \( \bar{T} - l_{\text{est}} = \delta l \) there exists a minimum of \( \Delta \) that is exactly zero (marked by the dotted circle). Further local minima exist which are separated by the amount of half the synthetic wavelength, \( \lambda_{\text{synth}} = \lambda_1 \lambda_2 / (\lambda_2 - \lambda_1) \), i.e. at positions \( \bar{T} - l_{\text{est}} = \delta l \pm \lambda_{\text{synth}} / 2 \) in which \( m \) is an integer. However, it should be noted that the coincidence pattern (grouped values of \( \Delta \) belonging to a certain number of \( m \)) is always different. The next neighbouring minima for which \( \Delta \) is nearly zero can be found at specific multiples of \( \lambda_{\text{synth}} / 2 \). In the example of figure 8 these minima are localized at \( m = \pm 3 \). This number is strongly dependent on the wavelengths \( \lambda_1 \) and \( \lambda_2 \). The amount of \( m \) could be calculated in the same way as the ‘range multiplier’ suggested by de Groot (see equation (14) of [53]).

The lengths which are closest to the correct minimum \( \bar{T} - l_{\text{est}} = \delta l \) are separate by a single integer interference order, i.e. at \( \bar{T} - l_{\text{est}} = \delta l \pm (\lambda_1 + \lambda_2) / 2 \). The corresponding values of \( \Delta \) amount to \( |\lambda_1 - \lambda_2| / 4 \). This reveals a very important requirement for the applicability of the method of exact fractions: the wavelengths must be clearly separate. Otherwise, the value of \( \Delta \) for the neighbouring lengths becomes too small. When, e.g., \( \lambda_1 \) and \( \lambda_2 \) would be just 2 nm apart from each other, for \( \{\delta_1 \pm 1, \delta_2 \pm 1\} \) values of \( \Delta \) would result in 0.5 nm which is very close to zero (\( \Delta = 0 \) is obtained for the correct variation numbers \( \{\delta_1, \delta_2\} \)). It is therefore important to note that in such case, despite the small values of \( \Delta \), the corresponding lengths are apart by relatively large amounts, namely \( \pm (\lambda_1 + \lambda_2) / 2 \). In other words, the ambiguity range critically depends on the uncertainty in the measurement of the fractional interference orders \( q_i \), with the remark that uncertainty contributions to the measured length, which are the same for all wavelengths used, have no influence on this ‘method of exact fraction’.

4. Typical uncertainty contributions in the practical realization of the length unit by interferometry

When considering the model equation(s) representing a length measurement, as required in order to derive the sensitivity coefficients for a GUM-compliant [55] uncertainty budget, a selection of contributions to the measurement uncertainty are
observed which are common across many length measurement systems based on optical interferometry. Typically, the largest source of uncertainty, when measuring the length of an artefact comes from correction of the measured length to the standard temperature of 20 °C, as prescribed in ISO 1 [56], due to inaccuracies in the measurement of the artefact temperature and lack of precise knowledge of the coefficient of thermal expansion (CTE). These issues can be obviated when the artefact is made of low CTE material, and then the remaining major uncertainty contributions come from the application of the optical interferometry to the length measurement. The uncertainty of the optical frequency is then usually the smallest source of uncertainty, when measuring the length of an artefact, leading to overall uncertainties of the order of several parts in 10^{-6}.

4.1. Determination of the air refractive index

As mentioned in section 1, for measurement in air, the refractive index is significant and must be measured directly (e.g. through comparison of equal optical path lengthsin air and in vacuum in a refractometer) or determined from measurement systems based on optical interferometry. Typically, the observed which are common across many length measurement systems based on optical interferometry. Typically, the largest source of uncertainty, when measuring the length of an artefact comes from correction of the measured length to the standard temperature of 20 °C, as prescribed in ISO 1 [56], due to inaccuracies in the measurement of the artefact temperature and lack of precise knowledge of the coefficient of thermal expansion (CTE). These issues can be obviated when the artefact is made of low CTE material, and then the remaining major uncertainty contributions come from the application of the optical interferometry to the length measurement. The uncertainty of the optical frequency is then usually the smallest source of uncertainty, when measuring the length of an artefact, leading to overall uncertainties of the order of several parts in 10^{-6}.

4.2. Cosine error

An angular misalignment between the measurand length and the light path length in the measuring arm of an interferometer contributes an error directly proportional to the cosine of the angle, θ, between the two lengths. For small misalignments, using the small angle approximation, the fractional error is approximately θ^2/2, e.g. for θ = 1 second of arc misalignment, the fractional error, given by δl/l = d · cos θ, is 1.2 × 10^{-11}. Minimisation of cosine error requires careful alignment of the interferometer with respect to the measurand.

4.3. Abbe error

When the measured length and the light path length are parallel, but laterally offset from one another, angular error, α, in the motion of the reflecting mirror in the measurement arm causes a length measurement error proportional to both the tangent of the angle, and the lateral offset, d, between the two lengths, e.g. for α = 5 s of arc and d = 100 mm, the error, given by δl/l = d · tan α, is 2.4 μm. Note that this error is length independent—it depends on the offset, d, and the motion error, α, not on l. Abbe error can be minimised by either reducing rotational errors caused by imperfect motion axes, or by minimising the lateral offset between the measurand length and the light path.

4.4. Light source aperture correction

Interferometers with collimated beams are subject to obliquity errors due to the finite (non-zero) size of the source. Some forms of interferometer require the source aperture to be offset from the focal point of the collimator lens (i.e. not on the principal axis). In such cases, the off-axis source causes a small angular deviation of the measurement beam with respect to the measurand axis, similar to a cosine error; for a negligible sized source, positioned a distance s off-axis with a collimator lens of focal length f, the error scales as δl/l = s^2/2f^2 [61]. For an interferometer with the source perfectly on axis, if the source size is finite, then elements of the source will still be off-axis and the resulting effect may be calculated by integrating the effect of all such infinitesimally small elements each contributing a cosine error. The result shows that the

Table 1. Main influence parameters affecting the air refractive index, their values at standard conditions, and their sensitivity coefficients evaluated at standard conditions.

<table>
<thead>
<tr>
<th>Influence parameter</th>
<th>Value at standard conditions</th>
<th>Refractive index sensitivity coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>20 °C</td>
<td>-9.2 × 10^{-7} K^{-1}</td>
</tr>
<tr>
<td>Pressure</td>
<td>101 325 Pa</td>
<td>+2.7 × 10^{-9} Pa^{-1}</td>
</tr>
<tr>
<td>CO₂ content</td>
<td>400 ppm</td>
<td>+1.4 × 10^{-10} (ppm)^{-1}</td>
</tr>
<tr>
<td>Humidity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>As relative humidity</td>
<td>50% RH</td>
<td>-8.7 × 10^{-9} (% RH)^{-1}</td>
</tr>
<tr>
<td>As dew point</td>
<td>9.27 °C</td>
<td>-1.5 × 10^{-8} K^{-1}</td>
</tr>
<tr>
<td>As water vapour pressure</td>
<td>1168 Pa</td>
<td>-1.8 × 10^{-10} Pa^{-1}</td>
</tr>
</tbody>
</table>
length error, for a source of diameter $d$ and a collimator lens of focal length $f$, is given by $\delta l = (d/4f)^2$ [61], e.g. for $d = 1 \text{ mm}, f = 1000 \text{ mm}$, the relative error is $6.25 \times 10^{-8}$.

### 4.5. Wavefront aberrations and non-planar wavefronts

For wide field interferometers optical aberrations, especially those unique to either the reference or measurement arm, lead to topographic errors in measured surfaces, with a one-to-one correspondence, as a minimum (some errors will accumulate during propagation). Typical commercial optics can achieve around $\lambda/20$ to $\lambda/40$ wavefront quality, leading to 15 nm to 30 nm measurement error across the aperture-limited surface. Most of this so-called optical error can be corrected by using an almost ideally flat surface (e.g. a flatness standard). The remaining optical error after correction is then of the order of magnitude of the flatness deviation of the standard.

Even for ‘perfect’ optical components, the physical wavefronts are not perfectly planar due to diffraction and the distance travelled by a wavefront during one period of the oscillation of the electromagnetic wave differs from that of an ideal plane wave and will also depend on location. A typical example of this is the well-known TEM$_{00}$ mode emitted by He–Ne lasers, which is Gaussian. For well-designed interferometers operating near the diffraction limit, the effect can be shown to be of the order of 1 or 2 nm [62, 63].

### 4.6. Polarisation effects (transport and crosstalk)

Both homodyne and heterodyne interferometers often use polarisation to separate the light for the two arms of the interferometer. The separated polarisation states are often achieved using polarisation-dependent optics. Polarisation jitter and imperfect extinction during splitting and recombination appear as phase errors, as do any extraneous rotations of the polarisation state during path length traversal, leading to length measurement errors [16–19], typically of the order of a few nanometres.

### 4.7. Parasitic reflections and non-linearity of fringes

All interferometers, and especially those utilising polarization optics for beam steering and phase measurement are prone to phase errors caused by parasitic beams which cause optical fringe interpolation to become non-linear, leading to errors of a few nanometres. For some effects, compensation is possible through sampling across several fringe periods [13] or improving the design of interferometer, but for the most demanding applications, where multiple parasitic beams produce a spectrum of phase non-linearities, detailed modelling is required to achieve the necessary compensation [64].

### 4.8. Reference path instability

Apart from Fizeau designs of interferometer, in which the reference path is of zero length, all interferometers have a finite length of the reference path (usually the separation between the beam splitter and the reference mirror). Erroneous changes in the reference path length directly contribute to length measurement error. Variation of the reference path length can be caused by e.g. vibrations or thermal changes in the mounting mechanism, e.g. consider a 1 m mechanical arm made of steel (CTE $10.7 \times 10^{-6} \text{K}^{-1}$), a 0.1 °C change in temperature would change the arm length by 1.07 μm, leading to a length error of the same value if the change occurred during the time the measurement arm was being monitored.

### 4.9. AC detection issues

When using AC detection, the detector itself can influence the phase measurement. e.g. typical detectors contain a small active area with some focussing optics in which any local inhomogeneity will couple with beam wandering (due to turbulence or surface form errors in optics) to cause phase errors. Other effects such as amplitude to phase coupling may be present depending on the type of sensor. It is possible that errors up to 1/2 fringe may be encountered.

### 4.10. Optical component errors and parasitic motions

For length measuring interferometers where a physical motion is monitored using a laser beam (fringe counting) parasitic motions of the moving part such as roll, pitch, and even yaw may cause errors in the reflected beam which returns to the detector. If plane mirrors are used to reflect the beam, roll and pitch cause angular deflections in the returned beam, which can cause sine errors if large aperture detectors are used. Even with cube corner retro reflectors, roll, pitch, and yaw will contribute errors which can be as large as 0.5 μm for parasitic rotations of 0.36° [65].

### 4.11. Light optical frequency (vacuum wavelength)

An uncalibrated, unstabilized 633 nm He–Ne laser can be assumed to have a vacuum wavelength $\lambda = 632.9908 \text{ nm}$ with a relative standard uncertainty of $1.5 \times 10^{-6}$, according to [66]. If this is insufficient, stabilized lasers are available commercially which can achieve frequency stability of around $10^{-9}$, frequency stabilized lasers can be calibrated by beat frequency against a reference laser or optical frequency comb with an uncertainty of a few parts in $10^{11}$, thus for the majority of the length scale, the instability or calibration error of the light source is usually the least significant error source.

It is possible for a 633 nm He–Ne laser to have a small secondary mode lasing on the He–Ne $3s^2 \rightarrow 2p^2$ transition at 640 nm, which might cause a non-linearity in the fringe interpolation, so it may be desirable to check for the presence of such modes [66]. About 50 years ago, an NMI reported that a laser purporting to operate at 633 nm was found to actually operate almost exclusively on the 640 nm transition. Rather than a non-linearity, this extreme case causes a systematic length-proportional error of 1.1% due to the incorrect value of the wavelength. We are not aware of more recent occurrences of this extreme problem but there is still a small possibility of a small amount of 640 nm light contaminating the 633 nm output.

A certain amount of 'parasitic light', which comes from an undesired mode of the resonator of a laser, can even lead to a relative error of a few parts in $10^7$ when measuring a length, although the light frequency of the main mode was calibrated
to a relative uncertainty of $10^{-12}$. As shown in [67] for a single parasitic light mode, the maximum effect on a length measurement is given approximately by $0.16/\eta$ fractional interference orders, where $\eta$ is the ratio between the intensity of the main light mode and that of the parasitic light mode. It therefore appears desirable to have information about the amount of parasitic light. An additional error source in stabilized lasers comes from parasitic reflections inside the laser system (often from the optics used in the stabilization scheme) [68]. The phase of such reflections varies with the optical path length and wavelength of the laser; for lasers employing modulation-based stabilization schemes, such as the less-often used Lamb dip (e.g. 32 kHz modulation to 16 MHz optical frequency depth), the phase difference modulates, giving rise to an intensity modulation at the same frequency as the modulation signal. This can bias the stabilization circuit, especially during warm-up. In two-mode stabilized and transverse Zeeman stabilized lasers which operate with two orthogonally polarized modes, optical feedback can cause changes in the quality factor of the optical cavity, causing reduction in the frequency stability performance [68].

4.12. Phase change on reflection

Reflection and transmission of light can be described by the well-known Fresnel equations. If a light wave, written in complex notation as $E = \text{Re} \{ A \cdot e^{ikz - \omega t + \phi} \}$, is reflected at perpendicular incidence at the surface of a material, the ratio between reflected ($E_r$) and incident light wave ($E$) is generally given by:

$$E_r/E = (\hat{n}_1 - \hat{n}_2) / (\hat{n}_1 + \hat{n}_2),$$

where $\hat{n}_1$ is the complex refractive index of the medium in which the wave propagates and $\hat{n}_2$ is the complex refractive index of the material from which the light is reflected. Assuming that $\hat{n}_1$ is a real number $n_1$ (such as for air, glass or vacuum and other dielectric materials), this ratio results to: $E_r/E = (n_1 - (n_2 + i \cdot \kappa_2)) / (n_1 + (n_2 + i \cdot \kappa_2))$ and can be written as:

$$E_r/E = \left[ E_r/E \right] \cdot e^{i \phi_{\text{RP}}}$$

where $\phi_{\text{RP}}$ describes the phase change on reflection:

$$\phi_{\text{RP}} = \arg \left[ E_r/E \right] = \tan^{-1} \left[ -2 \cdot n_1 \cdot n_2, n_1^2 - n_2^2 - \kappa_2^2 \right].$$

While for $\kappa_2 = 0$ and $n_2 > n_1$ equation (29) returns exactly $\pi$ radians, which is the known phase change on reflection for dielectric materials, increased values of $\kappa_2$ lead to an increased phase change for non-dielectric materials, e.g. metals (see [69] and references therein). This effect can be interpreted as the virtual penetration of the light into the material and must be generally taken into consideration in the measurement of the (geometrical) length.

When measuring step heights in a wide field interferometer, e.g. the measurement of a gauge block central length when wrung to a platen, any difference in the complex refractive index between the upper and lower surfaces will cause an erroneous difference in the phase of the reflected wavefronts leading to a fringe shift and a length measurement error. Measurements of $n_2$ and $\kappa_2$ for a number of steel gauge blocks using ellipsometry [70] indicated a 20% variation in $\kappa_2$, resulting in a $5^\circ$ variation of $\phi_{\text{RP}}$, corresponding to a variation in measured length of $\sim 5$ nm, for $\lambda \approx 633$ nm [71]. Typical values include $\sim 20$ nm difference between e.g. steel and glass.

Semi-transparent surface layers on the material cause multiple reflections that lead to an additional phase change. This must be taken into account in particular if the desired measurement uncertainty is in the sub-nm range. A prominent example is the volume determination of highly enriched $^{28}$Si spheres [72], where it is important to collect information about the composition and thickness of several surface layers and to calculate the influence of these layers on the measurement. Another example is the correction for the phase change occurring on dielectric mirrors of a Fabry–Perot cavity refractometer [73].

5. Secondary methods of realising the metre for dimensional nanometrology

A major addition to the mise en pratique for the metre with the 2019 revision of the SI was the introduction of secondary realisations of the metre based on the silicon lattice parameter. Details of this can be found in the mise en pratique [3], associated guidance documents [74–76] and a recently published overview [77]. A summary is presented here.

The growth in nanotechnology was originally predicted by Feynman [78]. Some years later Taniguchi [79] plotted the machining accuracy for different levels of machining as a function of time over the twentieth century extrapolating into the mid twenty-first century. His plot (a modified version of his graph has been reproduced in figure 9) showed that in the first quarter of the current century nanoscale metrology would be routinely required. Also shown in figure 9 are the primary scale attributes of three ways of realising the metre: the historically used 1 m platinum–iridium International Prototype metre (which was made obsolete by the revised SI definition of the metre in 1960), the vacuum wavelength of the popular 633 nm He–Ne laser stabilised to saturated absorption in iodine, and the 192 pm $d_{220}$ silicon lattice spacing. The relative uncertainties are also given for the latter two. The error sources described earlier in the paper (section 4) show that the frequency uncertainty of suitable lasers cannot be equated with the measurement uncertainty in the primary realisation of the metre. In fact, these sources of error dominate the overall measurement uncertainty, i.e. the laser frequency uncertainty is typically only a small contributor when using optical interferometry. For example, the relative frequency uncertainty of the iodine stabilised He–Ne laser is $2.1 \times 10^{-11}$. On the other hand, the measurement uncertainty for the primary realisation of lengths in air is a few parts in $10^8$. Achieving sub-nanometre accuracy with optical interferometers requires considerable care, often supported by independent verification that the non-linearity terms have been reduced to a sufficiently low level [64, 80]. Taniguchi’s modified graph in figure 9 shows that
to achieve sub-nanometre uncertainties will become increasingly challenging for optical interferometry, especially when working in air. To meet this demand for nanoscale dimensional metrology, the Consultative Committee for Length’s Working Group for Nanometrology considered alternative metre realisations at the nanoscale. A bottom-up approach in line with Feynmann’s prediction seemed logical; the Working Group recommended the lattice parameter of silicon. The silicon $d_{220}$ lattice spacing has been measured extensively as part of the Avogadro project [81–86]. It is listed in the Committee on Data for Science and Technology (CODATA) list of constants [87] with a value of $d_{220} = 192.0155714 \times 10^{-12}$ m, with a standard uncertainty of $0.000032 \times 10^{-12}$ m, i.e. relative uncertainty of $\Delta d/d = 1.67 \times 10^{-8}$ at a temperature of 22.5 °C in vacuum. Recent lattice parameter measurements have been performed on an isotopically enriched silicon 28 [88]. For the purposes of length metrology at the nanometre and picometre scale the lattice spacing uncertainty is sufficient for dimensional metrology. A 1 mm displacement would be equivalent to $5.2083 \times 10^6$ lattice spacings. With the quoted uncertainty of the lattice spacing, the corresponding relative uncertainty on the measurement of 1 mm based purely on the lattice spacing would be $\sqrt{n} \cdot U_C(d_{220}) \sim 10^{-14}$ m, where $n$ is the number of atoms per unit cell.

Obviously other factors need to be taken into account to calculate the overall uncertainty, depending on the technique used to realise the use of the lattice spacing, but nevertheless the potential of the lattice spacing is clear.

X-ray interferometry was originally developed by Bonse and Hart [89] with Hart suggesting the potential of the x-ray interferometer for dimensional metrology in his paper, ‘the angstrom ruler’ [90]. The theory of the operating principle of an x-ray interferometer for displacement measurement was described by Bonse and Hart [91]. Figure 10 shows a schematic plan view of a monolithic x-ray interferometer. The monolithic interferometer has been machined from a single crystal of defect free silicon that has been orientated such that planes from which x-rays have been diffracted (220) are perpendicular to the faces of three vertical lamellae (B, M and A) shown in figure 10. Incoming x-rays are incident at the Bragg angle and diffracted by the first lamella, B, that acts as a beam splitter. The two inward diffracted beams are diffracted again from the second lamella, M, and recombine at the third lamella, A, to form an interference pattern that is given by the lattice spacing rather than the wavelength of the x-rays being diffracted. This can be compared to a grating interferometer [92].

The third lamella has a flexure mechanism machined around it so that it can be moved parallel to the lamella face. When it is translated, a moiré fringe pattern is formed between the interfering x-ray beams and the planes in the third lamella.
from which x-rays have been diffracted. With careful alignment the fringes can be ‘fluffed out’ across the full width of the x-ray beam. This monolithic construction is favoured from the point of ease of construction, alignment, translation of the third lamella, and stability, however, its scanning range is typically limited to a few micrometres which is nevertheless suitable for many dimensional metrology applications. For lattice parameter measurements, a longer range of up to 1 mm is required and the interferometers used for such experiments comprise two separate components made from the same original piece of silicon. Since x-ray interferometry is the route for traceable measurement of the silicon lattice spacing, it is appropriate that it is one of the three methods chosen for realisation of the lattice parameter of silicon as a route to traceability for dimensional metrology.

To overcome the limited range of a monolithic x-ray interferometer, it has been combined with an optical interferometer (COXI) to provide a long range (1 mm) facility for the calibration of displacement measuring transducers [93]. Subsequently, x-ray interferometers have been used both as actuators to generate traceable sub nanometre displacements for positioning [94] and to measure sub nanometre errors in optical interferometers and encoders [95, 96]. Bergamin et al first proposed the idea of servo control of x-ray interferometers for quantised fringe positioning [97] in the combined optical and x-ray interferometer project [93]. Using this technique, the scanning stage of the x-ray interferometer can either be held fixed on the linear part of an x-ray fringe or moved in discrete steps of x-ray fringes. The development of a digital control system for an x-ray interferometer made during the Nanotrace project [80], where NMI-developed interferometers were assessed for non-linearity, enabled sub-fringe positioning and quadrature detection of x-ray signals [98] making x-ray interferometry a picometre ruler.

The second method chosen by the Working Group is based on transmission electron microscopy (TEM) to support linewidth metrology for the semiconductor industry. Linewidth metrology is typically performed using a critical dimension atomic force microscope (CD-AFM) that measures the sidewalls of linewidth structures. Unlike a conventional AFM, a CD-AFM [99, 100] servo controls the position of the AFM tip in a lateral direction rather than the traditional vertical direction and it uses a double anvil shaped AFM tip with the tips being horizontal. The measurement of linewidth is subtly different from measurement of pitch. When measuring pitch, the separation between identical structures in a row or column is being measured and the effects of dilation of the size of a structure due to probe sample interaction effects is cancelled out as common features are being measured. On the other hand, when measuring the width of a structure, the effects of probe sample interaction dilate the width of the structure; the probe can be either a tactile probe where the tip shape has an effect or electrons from an SEM where scattering can broaden the shape of the structure [101]. For several techniques, models have been developed to estimate the uncertainty contribution of the probe sample interaction. This limited the typical measurement uncertainty of linewidth structures to 2 nm for sub-micrometre wide structures. Knowledge of the lattice spacing of silicon combined with lattice resolving techniques for TEM offers the potential for a reduced measurement uncertainty for linewidth metrology. However, the destructive nature of sample preparation in the TEM means that TEM must be used in conjunction with other microscopy techniques such as scanning electron, optical or atomic force. Both NIST and PTB have been extremely active in this area developing measurement techniques and samples so that the effects of the tip width on CD-AFM measurements could be calculated and taken into account in the final measurement uncertainty [102–107]. Samples were developed that comprised a series of linewidth structures that could be measured using CD-AFM. One of the structures was then cross-sectioned so that it could be examined using TEM and the number of silicon atoms on the structure counted. This allowed the true width of the structure to be determined and then the CD-AFM measurements of all
Figure 11. (a) An overview image of the alignment marks and five line features measured by STEM (smallest line feature L1 marked on the right); (b) the line L1 with a nominal feature width of 50 nm measured by STEM; (c) the averaged intensity profile of the STEM image at the marked area in (b) with an inset figure showing the periodic crystal lattice planes of the silicon. Reproduced from [103]. © IOP Publishing Ltd. All rights reserved.

Figure 12. Optical images of (a) step free region (b) staircase and (c) amphitheatre structures. In sub-figures (b) and (c) the height between successive terraces is 0.314 nm. Reproduced with permission from [76].

the structures could be compensated based on the difference between the CD-AFM and TEM measurements of the same structure. Using this approach, the measurement uncertainty has been reduced to below 2 nm \((k = 2)\) with the uncertainty associated with the TEM-CD-AFM experiment 0.6 nm \((k = 2)\) [102–106].

Figure 11 shows an example of SEM and AFM analysis of a linewidth structure by PTB.

The third method by which the silicon lattice spacing can be realised is via monoatomic silicon steps for the calibration of scanning probe microscopes. The calibration of the topography axis of a scanning probe microscope is usually realised through height standards that have been calibrated by a national measurement institute using a metrological atomic force microscope [108]. These standards comprise a series of pillars or troughs of known height or depth that are positioned such that the substrate either side of the step can also be measured. This enables compensation for the effects of tilt of the sample in accordance with measurement methods described in ISO 5436 [109] and ISO 11952.

16
to be used. There are currently only a few suitable standards with heights below 10 nm and nothing available in the sub-nanometre range. To address this, partners in the European Metrology Research Programme project CRYSTAL [111] developed standards based on monoatomic terraces of silicon. This follows on from initial work undertaken at NIST [112, 113]. Although their steps demonstrated the principle of using the silicon lattice spacing, the steps were unidirectional and too narrow. Within the CRYSTAL project three types of sample have been produced and examples are shown in figure 12.

Figure 12(a) shows an example of a large area atomically flat region that can be used for identifying instrumentation artefacts since there should be no topography on the region scanned other than an apparent slope caused by tilting of the sample with respect to the lateral scan axes. Any topography present is likely to be an artefact of the instrument caused, for example, by stray light from the scanning AFM cantilever detection system.

Figure 12(b) shows the staircase of single (111) silicon steps for which the lattice spacing is 0.314 nm and figure 12(c) shows an example of the amphitheatre like structure. These structures are much more appropriate for the calibration of the vertical axes as not only do they have a series of steps, but they allow measurements to be made both descending the amphitheatre and then ascending so that the effect of sample tilt can be taken into account. Further work by Garnæs et al. [114] developed algorithms specifically for the areal processing of AFM measurements of atomic steps to aid the calibration of SPMS. These algorithms are available in the open source AFM image processing software Gwyddion [115].

6. Conclusions

In the new mise en pratique for the metre, published in 2019, the relationship to the primary parameters was clearly presented and the achievable measurement uncertainty as well as the influencing variables and corrections to be considered for the representation of the metre are specifically identified. Examination of the influencing variables clearly shows that the limiting factor for achieving traceability in length measurement is not the realisation of the metre in terms of a length derived from an optical frequency, but relating such an SI-traceable length to a physical length to be measured. At the nanometre scale, the achievable resolution from optical wavelength sub-division becomes the limiting factor and alternative methods for the secondary realisation of the metre in the field of nanometrology were also proposed and described in section 3 guidelines of the advisory committee for length (CCL), which are based, among other things, on the work within the framework of the international Avogadro cooperation.

The motto of the International System of Units is ‘for all people, for all time’ expressing the desire that the system is both accessible (‘for all people’) and eternal (‘for all time’). Since the first definition of the metre, in 1793, as a fraction of the Earth polar meridian, each successive re-definition has maintained consistency with its antecedents while adapting to the latest technology and scientific knowledge. The 1983 redefinition based on the speed of light was the first metre definition to deliberately move away from specifying any particular material or object (metal bars, particular wavelengths), choosing instead the more fundamental and abstract basis of the speed of light itself, taking advantage of recent work at the time on linking the krypton wavelength to the caesium clock frequency. The MeP for the metre was created to take care of the practical details and to collect together the details of the known optical frequency standards for reference, in anticipation that optical frequency sources would become cheaper and readily available for use in length measuring systems (such as laser interferometers).

This anticipation of emerging uses for high-precision length standards continues with the 2019 redefinition, where a further fundamental change to the MeP is to include a secondary realisation of the metre, based on the silicon lattice, opening the way for access to sub-nanometre standards for nanotechnology. But we should not lose sight that although the SI length scale is now extended to the world of the very small, the very same scale still covers the original range of metres to planetary dimensions, and much further. At the present, the Voyager spacecraft represent the longest range traceable dimensional measurement to a human-manufactured object,—through time of flight [116]. The SI traceable, physical, length scale now runs from $10^{-13}$ m to over $10^{10}$ m and yet is still compatible with measurements made using the metre as first adopted over 200 years ago—a remarkable achievement.

Acknowledgments

The authors wish to thank the CIPM Consultative Committee for Length and associated working groups, in particular, the Working Group for Nanometrology, for their support. The writing of this paper was partially funded by the Engineering Measurement Programme from the National Measurement System, UK.

ORCID iDs

Réne Schödel https://orcid.org/0000-0002-7597-9036
Andrew Yacoot https://orcid.org/0000-0001-6740-821X
Andrew Lewis https://orcid.org/0000-0001-9556-3890

References

[25] EMPIR 2021 Large volume metrology applications http://empir.npl.co.uk/lava/
[28] Haitjema H, Schellekens P H J and Wetzel's S F C L 2000 Calibration of displacement sensors up to 300 m with nanometre accuracy and direct traceability to a primary standard of length Metrologia 37 25
[34] Kösters W 1926 Ein neuer interferenzkomparator für unmittel- baren wellenlängenanschluß Z. Feinmechanik Przízíon 34 55–9
[46] Bönsch G and Böhme H 1989 Phase-determination of Fizeau interferences by phase-shifting interferometry Optik 82 161–4
[47] Bönsch G and Potulski E 1998 Measurement of the refractive index of air and comparison with modified Edlen’s formulae Metrologia 35 133
plane mirror optical interferometer Meas. Sci. Technol. 11 1126–30


[111] EMRP 2016 Crystalline and self-assembled structures as length standards https://ptb.de/emrp/sib61-home.html


