TECHNICAL REPORT

ISO/TR 14999-2

Second edition
2019-07

Optics and photonics — Interferometric measurement of optical elements and optical systems —

Part 2: Measurement and evaluation techniques

Optique et photonique — Mesurage interférométrique de composants et systèmes optiques —

Partie 2: Mesurage et techniques d'évaluation

Reference number
ISO/TR 14999-2:2019(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword ..... **v**
Introduction ..... **vi**
1 Scope **1**
2 Normative references **1**
3 Terms and definitions **1**
4 Measurement objects **1**
4.1 Surfaces 1
4.1.1 Boundary surfaces of optical components 1
4.1.2 Reflection degree 2
4.1.3 Roughness 2
4.1.4 Topology of the regions (discontinuous regions) 2
4.1.5 Continuity of the surface; gradient of the surface 2
4.1.6 Stiffness of mirrors; finite-element-calculations 3
4.1.7 Temperature homogeneity of mirrors 3
4.1.8 Examples of measurement objects 3
4.2 Optical components in transmission 3
4.2.1 Single-pass versus double-pass testing 3
4.2.2 Windows (wavefront aberrations in transmission) 3
4.2.3 Prisms (wavefront aberrations and angle error) 3
4.2.4 Influence of temperature on the refractive index 3
4.3 Optical systems 4
4.3.1 Single-pass versus double-pass testing 4
4.3.2 Examination in the pupil 4
4.3.3 Chromatic aberrations 4
4.4 Indirect examination of the function of optical elements 4
4.4.1 Examination with different wavelengths 4
4.4.2 Examination with different beam paths 4
4.4.3 Tolerance range 4
5 Hardware aspects of an interferometer and test environment **4**
5.1 General 4
5.2 Construction principles and influences on the quality of measurements 6
5.2.1 General 6
5.2.2 Intrinsic instrument errors and the principle of common path 7
5.2.3 Optical compensation of errors 8
5.2.4 Mathematical compensation of errors 9
5.2.5 Contrast as a function of the irradiance in test and reference arm: methods to attain equilibrium in both arms 10
5.2.6 Contrast as a function of performance of the light source 14
5.3 Test environment 16
5.3.1 General 16
5.3.2 Influence of vibrations 17
5.3.3 Influence of gravity and support of the test piece 18
6 Methods for evaluating the optical path difference **20**
6.1 General 20
6.2 Visual inspection of interferograms 20
6.2.1 General 20
6.2.2 Example 1 — Fizeau interferometer 20
6.2.3 Example 2 — Twyman-Green interferometer 24
6.3 Manual evaluation of interferograms 26
6.4 Phase measurements with temporal carrier 28
6.4.1 General 28
6.4.2 Heterodyne interferometry 29

6.4.3 Phase lock interferometry (PLI) 30
6.4.4 Synchronous detection and phase-shifting interferometry 30
6.5 Phase measurements with spatial carrier 34
6.5.1 Fringe analysis by Fourier transform operations 34
6.5.2 Spatially synchronous fringe analysis 35
6.6 Removal of phase ambiguities (phase unwrapping) 37
6.7 Registration of wavefronts; coordinate systems, coordinate system definition 38
6.8 Polynomial and other representations of wavefronts 39
6.8.1 Representation of phase data 39
6.8.2 Zernike polynomials for a circular boundary 40
6.8.3 Use of Zernike polynomials for an elliptical boundary 40
6.8.4 Zernike-Tatian polynomials for a circular shape with a central hole (annular) 41
6.8.5 Legendre polynomials for a rectangle boundary 41
6.8.6 Orthogonal functions on “unusual areas” 41

7 Test reports and calibration certificates 41
7.1 General 41
7.2 Content of test reports and calibration certificates 42
7.3 Test reports 42
7.4 Calibration certificates 43
7.4.1 Basics 43
7.4.2 Specification 43
7.4.3 Adjustment or repair 43
7.5 Opinions and interpretations 43
7.6 Electronic transmission of results 43
7.7 Format of reports and certificates 44
7.8 Amendments to test reports and calibration certificates 44

8 Data format 44

Annex A (informative) **Orthogonal polynomials 45**

Bibliography 62

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 172, *Optics and Photonics*, Subcommittee SC 1, *Fundamental standards*.

This second edition cancels and replaces the first edition (ISO/TR 14999-2:2005) which has been technically revised. The main changes are:

a) Figure 1 has been updated.

b) A.1 has been aligned with the notation of ISO 14999-4.

c) Updated text referring to technologies that have evolved over the last 10+ years, such as lasers and detectors.

d) Improved clarity of the overall document (many minor edits made throughout the text).

A list of all parts in the ISO 14999 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

A series of International Standards on Indications in technical drawings for the representation of optical elements and optical systems has been prepared by ISO/TC 172/SC 1, and published as ISO 10110 under the title *Optics and photonics — Preparation of drawings for optical elements and systems*. When drafting this standards series and especially its Part 5, *Surface form tolerances* and Part 14, *Wavefront deformation tolerance*, it became evident to the experts involved that additional complementary documentation was required to describe how the necessary information on the conformance of the fabricated parts with the stated tolerances can be demonstrated. Therefore, the responsible ISO Committee ISO/TC 172/SC 1 decided to prepare an ISO Technical Report on Interferometric measurement of optical wavefronts and surface form of optical elements.

When discussing the topics which had to be included into or excluded from such a Technical Report, it was envisaged that it might be the first time, where an ISO Technical Report or Standard is prepared which deals with wave-optics, i.e., in which the ray approximation of geometrical optics is no longer valid. As a consequence, fewer references than usual were available, which made the task more difficult.

Envisaging the situation, that the topic of interferometry has so far been left blank in ISO, it was the natural wish to now be as comprehensive as possible. Therefore, the committee held discussions, whether important techniques such as interference microscopy (for characterizing the micro-roughness of optical parts), shearing interferometry (e.g. for characterizing corrected optical systems), multiple beam interferometry, coherence sensing techniques or phase conjugation techniques should be included or not. Other techniques, which are related to the classical two beam interferometry, like holographic interferometry, Moiré techniques and profilometry were also mentioned as well as Fourier transform spectroscopy or the polarization techniques, which are mainly for microscopic interferometry.

In the end, the committee adopted the guideline to include what presently are common techniques used for the purpose of characterizing the quality of optical parts as described in the ISO 10110 series. The decision was made to complete a first Technical Report, and to then update it by supplementing new parts, as required.

The committee intends that this document covers the need for qualifying optical parts and complete systems regarding the wavefront error produced by them. Such errors have a distribution over the spatial frequency scale; in this document only the low- and mid-frequency parts of this error-spectrum are covered, not the very high end of the spectrum. These high-frequency errors can be measured only by microscopy, measurement of the scattered light or by non-optical probing of the surface.

A similar statement can be made regarding the wavelength range of the radiation used for testing. ISO 14999 considers test methods with visible light as the typical case. In some cases, longer wavelength infrared radiation (e.g. 10,6 µm CO_2 lasers) is used for testing rough surfaces after grinding. A variety of laser wavelengths might be used for transmitted wavefront testing of optical systems at the application wavelength (e.g. near-infrared 1,55 µm or 1,06 µm, or ultraviolet 193 nm or 248 nm excimer lasers for microlithography optics). However, these are still rare cases, which are included in standards, that will not be dealt with in detail. The wavelength range outside these borders is not covered.

Optics and photonics — Interferometric measurement of optical elements and optical systems —

Part 2: Measurement and evaluation techniques

1 Scope

This document gives fundamental explanations to interferometric measurement objects, describes hardware aspects of interferometers and evaluation methods, and gives recommendations for test reports and calibration certificates.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

No terms and definitions are listed in this document.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

4 Measurement objects

4.1 Surfaces

4.1.1 Boundary surfaces of optical components

A common task in interferometry is measurement of the shape of a surface. This can be accomplished in two different ways. Either reflected light or the light transmitted through the surface could be used for the measurement.

Interferometric measurement is achieved by comparing the difference of two optical path lengths $\int nd$. Usually one path is called the reference path, the other the measurement path.

The resulting wave aberration, ΔW, for a displacement d of the surface, if measured in reflection, is $\Delta W = 2nd$. The same displacement measured in transmission results in the wave aberration $\Delta W = (n_2 - n_1)d$.

4.1.2 Reflection degree

The Fresnel reflection from the boundary between two different media, R, can be calculated from the refractive index n_1 and n_2 at the boundary surface.

$$R = \left(\frac{n_2 - n_1}{n_2 + n_1} \right) \tag{1}$$

For most optical glasses this value is between 4 % and 6 %, so an average of 5 % is usually a good estimate.

This reflection causes a loss of light from the transmitted wavefront at every surface. On the other hand, this reflection is often used for the measurement itself. To obtain maximum fringe visibility, or contrast, the two interfering beams should have approximately the same intensity. Changing the reflectivity of the beam splitter within an interferometer only changes the amount of light in the interference pattern and does not change the beam intensity ratio of the two beams because the light in both arms is transmitted through and reflected by the beam splitter once. If the measurement path and reference path are separated, as in a Mach-Zehnder or Twyman-Green set-up, it is usually possible to adjust the intensities of the light in both arms.

A major problem arises in a Fizeau interferometer if the reference surface has high reflectance, the result will be multiple beam interference fringes resulting in narrow fringes as in a Fabry-Perot interferometer. If sinusoidal fringes are required as for the evaluation by phase shifting interferometry, the reference surface should have low reflection and an element needs to be introduced between the reference and the measurement surface that will absorb light without distorting the wave aberration.

The issue is solvable, when using a wavelength shifting interferometer or short coherence interferometer.

4.1.3 Roughness

For interferometric measurement the roughness of the measured surface should not exceed a certain limit that is a fraction of the wavelength and of the difference of indices of refraction, if used in transmission.

4.1.4 Topology of the regions (discontinuous regions)

Difficulties may arise with interferometer software when the wavefront area has breaks in it (e.g. because it is split into segments by the mechanical supports of the secondary mirror of a mirror telescope). Problems are most severe with static fringe analysis software that depends strongly on using neighbouring points to determine the position and continuity of fringes. Software analysis of phase differences is not affected to the same extent as it is a point-by-point evaluation of wave aberrations.

Similar problems may occur if the wavefront area has a complicated outline.

4.1.5 Continuity of the surface; gradient of the surface

Due to the inherent ambiguity of $\pm n2\pi$ phase difference, it is not possible to measure any arbitrary surface shape uniquely. The evaluation of a smooth surface is usually correct, if the wave aberration between two resolvable points is less than π. This effectively limits the largest slopes (or highest step discontinuities) that can unambiguously be measured with the interferometer.

The gradient of the surface under test relative to the reference surface results in a gradient of the measured wave aberration and in high-density or closely spaced fringes. Interferograms cannot be evaluated, if the fringe separation is less than twice the distance of two resolvable points. Thus the local gradient of the imaged wavefront needs to be less than (0,5 λ)/(detector element spacing) for unambiguous phase recovery. If this condition is not possible by adjustment, or by changing the measurement set-up, compensating optics may be required in some cases.

Some of the problems caused by phase difference ambiguity can be solved by multiple wavelength interferometry.

4.1.6 Stiffness of mirrors; finite-element-calculations

An optic under test should not be deformed in a manner other than it would be deformed under its intended application. It can be difficult to notice whether an optic is deformed during the measurement by the test fixture, which holds the optic in place during the measurement. As a first indication of the influence of the test fixture, the object can be measured by using two or more different test fixtures to hold the optic in different ways. In case of any doubt, a finite-element calculation is recommended to evaluate the effect of deformation on the optic.

4.1.7 Temperature homogeneity of mirrors

During measurement the object should have a homogeneous temperature. Inhomogeneous temperatures can cause deformations. The thermal expansion coefficient of optical materials is rather high and the thermal conductivity of optical materials is very low. Sufficient time should be allowed for optics under test to reach thermal equilibrium. In some cases, this can take minutes, but in others it might require several hours to reach thermal equilibrium.

4.1.8 Examples of measurement objects

Items that can be measured by interferometry include optical flats, windows, raw glass, convex and concave mirrors, lenses, prisms, and optical systems.

4.2 Optical components in transmission

4.2.1 Single-pass versus double-pass testing

Transmitting optical components can be measured in single-pass or double-pass, depending on the interferometric set-up. Double-pass measurement increases the sensitivity by a factor of two but may also include the effect of the reflecting surface. In double-pass measurements consideration should also be given to the possibility that the returning light passes back through the component at different locations.

4.2.2 Windows (wavefront aberrations in transmission)

For windows the shape error of the surfaces is usually not important. Also, the measured transmitted wavefront will include the homogeneity of the material. Depending on the application, a certain amount of power may be tolerated separate from the other wave aberrations. Also, a tolerated wedge can be measured by interferometry. However, it can be more convenient to measure angular errors by different equipment.

4.2.3 Prisms (wavefront aberrations and angle error)

As in the case for windows, the wavefront aberrations and angular errors of prisms can be measured by different equipment. However, if the angular tolerances are in the interferometric region, and many parts are to be measured, it can be more convenient to measure both features by interferometry. In this case a fixed set-up, or a master specimen, is used as a reference.

4.2.4 Influence of temperature on the refractive index

For measurement of an optical component in transmission, it should be noted that not only the objects might be deformed by the thermal expansion but, also, that the refractive index of the material changes with temperature. Therefore, thermal settling of the test piece before testing is even more important.

4.3 Optical systems

4.3.1 Single-pass versus double-pass testing

Complete optical systems can be measured by interferometry in a manner similar to the testing of single components. It is, however, important that systems be measured in the same geometry as they were designed to be used. This can lead to a complicated set-up in single or double pass. For long systems tested in double pass and in the presence of severe aberrations, it is necessary to consider that the return light path can be considerably different from the incident light path.

4.3.2 Examination in the pupil

Interferometric measurements should be made in the exit pupil of the optical system.

4.3.3 Chromatic aberrations

If systems are measured at wavelengths different than those they are designed for, the effects caused by chromatic aberrations should be computed. There will be some systems, where the wave aberrations can be simply scaled by the ratio of the test and design wavelengths, whereas other systems are so different that a measurement is not possible.

For transmitted wavefront measurements with a wavelength-shifting interferometer, consider whether the classical refractive index or the group refractive index should be used to describe the medium.

4.4 Indirect examination of the function of optical elements

4.4.1 Examination with different wavelengths

In some cases, examination of flat optical elements is possible at wavelengths other than the application wavelength. In these cases, corrections can be scaled to the application wavelength. It should be noted, however, that inhomogeneities of optical materials may to some degree depend on the wavelength range. Because of the presence of chromatic aberrations no universal recommendation is possible.

4.4.2 Examination with different beam paths

It is preferred that the measurement set-up should be as similar as possible to the application. However, in some cases it may be more convenient to measure optical elements in a way that is different from their use. In this case, it may be difficult to find a correlation between the measured wave aberration and how the application is affected by these aberrations.

4.4.3 Tolerance range

Sometimes the relationship between the interferometric measurement and the tolerances of the measured objects is not clear. Usually the complete test set-up should be considered and, if possible, the sensitivity to measurement tolerances analyzed with finite-element calculations. Optical design calculations may also be used as an evaluation method.

5 Hardware aspects of an interferometer and test environment

5.1 General

The purpose of this clause is to acquaint the user with common issues associated with interferometric measurements that can affect the accuracy of measurements. It is a matter of fact that two different persons using the same hardware and doing their measurements in the same laboratory, will not necessarily achieve identical results with their measurements. The skilled user might achieve a highly accurate result, whereas the unskilled user might have severe errors in his result that he might not be

aware of. It is important to keep in mind that good reproducibility of measurement is no guarantee for a correct result, because systematic sources of errors might have influenced the measured results. Knowledge about such possible influences, and how to avoid them, is what experimental skill is about.

Sources of errors in interferometric measurements include, for example:

— improper use of the measuring instrument, because the optical principles are not well understood, e.g. failure to image the surface under test onto the CCD camera of the interferometer;

— use of unsuitable fixtures to hold the test piece, inducing stress which causes bending;

— influence of gravity on the test piece;

— vibrations of the test set-up, which might induce phase-measuring errors;

— unsuitable use of polynomial fits with respect to the given shape of the aperture (for example due to some obscured parts of the circular shape) and adjacent subtraction of error terms like tilt and focus terms, due to a violation of the orthogonality assumption;

— presence of stable layers of air with different temperatures in the interferometer cavity, causing systematic low-order aberration;

— flipping (mirroring), or some other mismatch, of a calibration error map with respect to the actual orientation, shape or magnification of the measured field;

— influence of different temperature or different focus settings between calibration and measurement;

— use of test pieces that are not homogeneous in temperature and have a considerable coefficient of temperature expansion;

— adjustments with tilt or focus subtraction can lead to unnoticed misalignments.

These are only examples; although there are a much greater number of “typical” sources of error. The only way to overcome such types of error, which depend very much on the actual test situation and the demands for the final accuracy, is that the operator planning and assembling the test should be aware of possible influences on the accuracy of the measurement, which might be of an optical or mechanical nature.

Conceptually, it is very important not to believe blindly the results which the instrument shows. At the same time, it is equally important not to blame the instrument, or the principle of the interferometric measurement, if there are inexplicable results. Note that in the majority of cases the instrument shows the “correct” readings from what is presented to it, even if that is not the measurement task in question. If, for example, the measured error map does not rotate by 72° when the test piece is rotated physically by 72°, this might indicate that the reference surface may contribute a considerable amount to the total error. The support of the test piece can also influence the measurement, etc.

Another test might be to repeat the measurement after 1 h without touching anything in the meantime. If the results deviate from each other, the temperature of the surface under test or its support structure may have had an uneven temperature distribution in the first test. The time needed to achieve thermal stability adequate for the measurement can vary significantly, based on the part and fixture dimensions, coefficient of thermal expansion, dn/dT, thermal conductivity, heat capacity, initial temperature distribution, and the desired measurement accuracy. Also, the temperature in the laboratory might have changed, the instrument might have warmed up, etc.

Such tests are imperative in order to exclude at least the most common sources of error. It is strongly recommended to repeat a measurement at least three times and compare the results; this repetition should include the demounting and remounting of the part in the test set-up, as well as all the adjustments of the set-up and the settings of the interferometer. It is even better to repeat the whole test procedure on another day, and even by another operator.

All measurement conditions and settings should be documented and the final data sets should be stored in the computer in an organized way. Ideally, the documentation should be stored together with

the measured data sets. Any further treatment like subtraction of tilt or even higher order (Zernike) functions, number of averages, any filtering like smoothing with a spatial low pass or median filter to remove "spikes", should be documented and stored together with the data set. Such information is part of the result and when not given together with the measured surface map, the result is useless and cannot be used for proof of quality for the part under test.

5.2 Construction principles and influences on the quality of measurements

5.2.1 General

When the wavefront deviation of a test piece is measured by an interferometer, the test piece becomes part of the optics of the instrument. The auto-collimation condition should be met, as well as the condition to image the surface under test onto the detector. In order to achieve high flexibility of possible locations for the surface under test and for different test configurations, there will be stringent requirements on the spatial and temporal coherence of the light source that need to be fulfilled. These can easily be attained by use of a laser and, together with a very high intensity compared to other light sources, are the reason that the laser is the typical light source for interferometers.

One of the consequences of the very high coherence of lasers is that all kinds of imperfections, such as impurities of substrates, optical cements and coatings, tiny scratches, bubbles, holes, dust particles, micro-roughness of surfaces, which can occur at any part of the light path through the interferometer, are "collected" and are superimposed as an uncleanliness, i.e., unwanted amplitude and phase modulations of the wavefronts. The further away the disturbing defects are from an image plane of the detector, the more the defects are altered in their phase distributions due to Fresnel diffraction and in spatial frequency. A very narrow defect located on a surface near an image of the light source might spread out to a big size in the detector plane. The specification of optical parts used in an interferometer set-up therefore have to be much more stringent than in conventional optical instruments and depend on the position of the part in the ray path. For surfaces near the image of the light source (where the diameters of the ray bundles are small) ultra-high surface quality requirements should be maintained. Generally speaking, the higher the test accuracy needs to be, the more severe are the demands for the quality of all interferometer parts.

As discussed in ISO/TR 14999-1, it is very important to image the wavefront under test onto the detector plane. If the location of this wavefront relative to the instrument changes from one test set-up to another, a possibility of refocusing the detector to this new location should be provided. In some instruments, provision is made to alter the magnification with which the wavefront under test is imaged onto the detector. In some cases, this is done in fixed steps, in other cases this is done continuously over a certain range. On the other hand, it is necessary to attain a good optical wavefront-correction when "tailoring" the wavefronts in the instrument to the desired shape and at the same time realizing a good optical transfer function for the amplitude and phase when imaging the wavefront under test at the detector plane. An interferometer's ability to transfer different spatial frequency features (with amplitudes a small fraction of a wavelength) to the detector is quantified with an instrument transfer function (ITF). This is also sometimes termed system transfer function (STF) or height transfer function (HTF). Definitions and methods for quantifying the ITF can be found in References [5], [6] and [7] and elsewhere in the literature.

All such possibilities and demands cause a certain amount of complexity of the optical layout of such an instrument, leading to optical systems with multiple surfaces. It is obvious that it is more difficult to keep the unwanted additional disturbances by the optical parts small when more optical parts are necessary to achieve the desired functionality. The skill of the designer of an instrument lies in finding the best compromise between the degree of aberration correction (keeping the wavefront errors with low spatial frequency small) and the degree of noise (i.e., high spatial frequency errors). The noise increases with every additional surface which might be necessary for aberration correction. Since the complexity of the optical layout grows with the universality of the use of the instrument, it is much easier to construct a high-quality single-purpose instrument.

Due to higher cost of production, and the deterioration in the appearance of the interferograms obvious to any customer, companies tend to minimize the number of optical elements and seek to achieve the best correction for wavefront aberrations. This might have consequences for the handling of the

instrument. If, for example, the transmission sphere, used for spherical testing, is not aligned properly when attached to the instrument, coma and astigmatism might be introduced into the measured wavefront. If deviations in the alignment of the focus setting between calibration measurements and final measurements for the parts under test exist, this again might cause wavefront errors in the final results. While an instrument that incorporates more wavefront corrective components may be "robust" against higher order aberration errors, it may lead to measurement results with a higher amount of coherent noise due to the increased number of surfaces.

The opposing criteria for the way to design a laser interferometer require a compromise between wavefront quality, field correction, versatility on the one hand and number and location of surfaces on the other hand.

5.2.2 Intrinsic instrument errors and the principle of common path

The task gets more and more difficult when the errors, which have to be measured accurately, become smaller and smaller. It might be concluded from this that it would be nearly impossible to get reliable measuring results. Needless to say, it is necessary that the "intrinsic errors" caused by the instrument itself should be at least not higher than the errors caused by the test piece. Example: suppose the test piece is a well-polished spherical surface of a lens. The interferometer itself might include in total 12 lens surfaces and another 10 surfaces of plane plates. Therefore, it would be necessary to fabricate the 22 surfaces within the instrument to a degree of perfection that is at least 22 times better than that of the test piece in order to attain the same disturbance from the instrument (i.e. "intrinsic errors") and from the test piece. Or, it might be concluded that a factor 4 to 5 would be sufficient for randomly distributed errors. Even in the latter case, it would be nearly impossible to fabricate and maintain an instrument with such a degree of perfection.

This argument is both right and wrong. It is the principle of interference that errors common to both waves, i.e., the test wave and the reference wave, cancel out. The ultimate use of this principle is apparent in the Fizeau-type interferometer, where all but the last surface before the surface to be tested and the air between these two surfaces are common to both waves. The quantity that is measured by a Fizeau interferometer is the optical path difference between the two surfaces facing each other (the "optical thickness distribution" of the air gap; this includes the distribution of the refractive index of the air). So far the argument is wrong, but it is right for very small-scale errors. It is never possible for the rays to travel exactly the same path, so the principle of "common path" with cancellation of common errors is always violated, if high spatial frequency noise is in demand. So, even if Fizeau interferometers are more robust for errors with low spatial frequencies, this is not the case for coherent noise.

In order to check the sensitivity of an instrument to alignment errors as well as for intrinsic high frequency noise, the following two procedures can be useful.

a) The following procedure should be repeated with different orientations of the fringes and also with the highest number of fringes the instrument is capable of measuring. This test is one measure for the robustness of the instrument against misalignments of all kinds. Proceed as follows.

   1) Place a reference flat in front of the transmission flat of the instrument and adjust for about 25 fringes of tilt.

   2) Perform a measurement and store the result.

   3) Adjust for zero-fringes and perform another measurement.

   4) Subtract the step 3 measurement data from the step 2 data and compute the Zernike terms for the resulting difference-data set; besides other errors, the induced wavefront tilt with respect

to the optical axis make visible the optical wavefront aberrations which go along with the small angle of the test wave.

b) The second procedure checks the intrinsic high frequency noise and therefore the ability to detect small-scale errors which normally go along with very small amplitudes. Proceed as follows.

   1) Take the data of the difference from the measurement on-axis and the measurement with the 25 fringes of tilt.

   2) Subtract the first 36 Zernike terms. The remaining surface map shows mainly the intrinsic coherent noise of the instrument.

   3) When "spikes", which occur at the boundary of the measured part, are removed, the rms value is a quality number for the intrinsic noise. This noise should be uncorrelated when the experiment is repeated with different orientations of the fringes and therefore reduce with the square root law, if measurements with different fringe orientations are averaged and the difference of those averages are calculated instead of the difference of only two measurements.

Together with these tests, it is recommended that two other simple checks be performed to assess the proper alignment of the instrument:

The collimation of the plane-wave leaving the system can be checked with the help of a thick (>30 mm) plane-parallel plate of known good optical quality by inserting the plane-parallel plate with an incident angle of 45° into the beam and projecting the lateral shearing interferogram onto a screen. If the plate has no wedge-angle, no fringes should be visible, but they will be present if the wavefront from the interferometer converges or diverges.

The adjustment for the alignment reticle, or other means for the alignment of the beam, can also be checked using a corner-cube mirror or prism of known good quality. The returning beam should be incident precisely at the centre of the alignment device. This is also a method for adjusting the reference surface perpendicular to the beam. By tilting the reference surface, the interference fringes formed by the reference wave and a wave reflected by the triple-mirror (having three surfaces with angles of 90° between them) should be made as broad as possible.

5.2.3 Optical compensation of errors

It is a very useful property of two-beam interference that the physical principle can help to suppress the errors caused by component parts of the interferometer. If the two interfering beams experience the same disturbances when passing through the optical parts, the wavefront errors impressed on them are identical and cancel out in the final wavefront difference. Therefore, the two wavefronts should travel almost the same path through the instrument so ensuring that "optical compensation" of errors takes place. This is achieved best by the arrangement of a Fizeau test set-up, where the test surface and the reference surface face each other without any component in between. This cancellation is not necessarily perfect if there are deviations in the optical path when the beams are tilted with respect to each other.

Another deviation from perfect symmetry of the interfering waves is due to the imaging conditions of the two surfaces in question onto the detector. The two surfaces in a Fizeau configuration, the test surface and the reference surface, cannot both be imaged exactly onto the detector surface at the same time. Normally, the reference surface is larger than necessary, so that the surface under test defines the final aperture for the size of the interferogram. In this case, no errors are introduced to a first order by Fresnel diffraction at the boundary of the reference surface, when this boundary is larger than the diameter used. Nevertheless, there is another higher-order error, which will be explained in conjunction with the testing of spherical surfaces with a so-called "Fizeau transmission sphere".

Figure 1 shows the optical conditions when testing a spherical surface B with respect to a reference surface A. In this example, it is assumed that the reference surface A (this is the last surface of the transmission sphere) as well as the test-surface B are both concave. The apexes of the surfaces are called S_A and S_B. The Fizeau cavity is set up correctly when both surfaces are arranged so that they

have a common centre of curvature. If the Fizeau lens is calculated and manufactured properly, this centre point coincides with the focal point F of the spherical wave, illuminating the reference surface A.

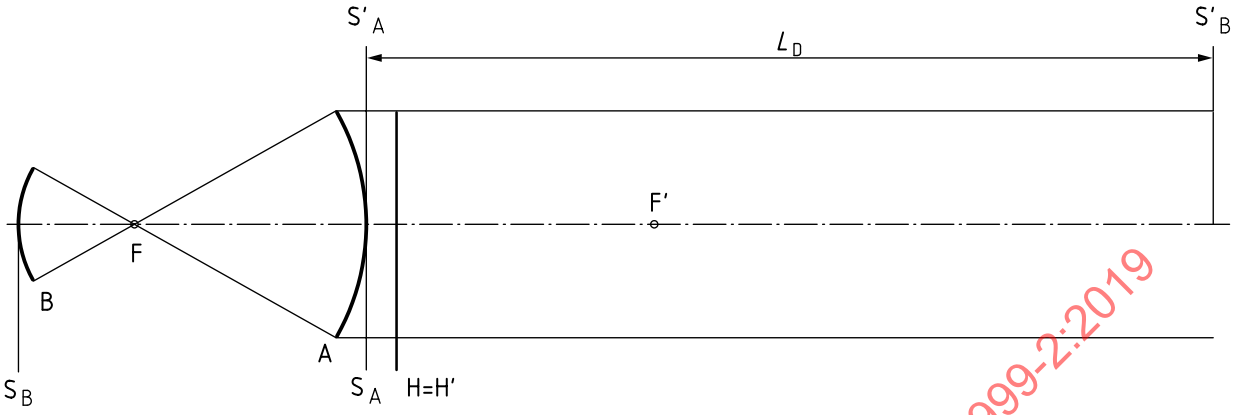


Figure 1 — The images S_A' and S_B' of the apexes of the reference surface S_A and the test surface S_B are defocused with respect to each other

As can be seen from Figure 1, both surfaces in question are imaged by the transmission sphere, which is optically represented by the principal plane H = H′ and the focal points F and F′, into different locations at the optical axis. The sizes of the images match, but the axial distance between the images, L_D (the "defocus") is shown in Figure 1. Now, it is supposed that there might be aberrations already present in the wavefront illuminating the transmission sphere as well as additional aberrations added by the transmission sphere itself. The transmission sphere should image an infinite object into the focus point F without introducing additional spherical aberration. Furthermore, the test surface apex S_B should be imaged to S'_B and the reference surface apex S_A into S_A' without adding different phase-terms into these images.

As was explained in ISO/TR 14999-1:2005, 2.11, only ideal plane waves do not alter their shape when they spread out. So, even if we suppose that the two wavefronts at the locations S'_A and S'_B might still have the same shape, i.e., the same aberrations, the fact that one of them has to travel the distance "defocus" will cause them to deviate from each other when they meet to interfere on the detector. This introduces a systematic wavefront error, which depends on the wavefront quality of the interferometer "common path" and is the more pronounced the larger the distance "defocus" is compared to the diameter of the wavefront. As a rule of thumb, the radius of the test surface should not be smaller than 10 % of the focal length of the transmission sphere. Also, the transmission sphere should not introduce an error greater than $\lambda/2$ in double pass, though higher local slopes are more problematic. For high-precision measurements, these tolerances should even be more stringent.

It is important to keep in mind that any wavefront error, which is already present in the plane wavefront entering the transmission sphere, will be made visible also by this effect of defocused images. There is no way to overcome this problem with Fizeau interferometers other than to keep the defocus as small as possible by using a transmission sphere with the smallest possible air gap. The influence of this error can be minimized by appropriate calibration measurements.

5.2.4 Mathematical compensation of errors

The great advantage of "optical compensation" of errors due to the principle of "common path through the optics" is that this compensation takes place in "real time", i.e., continuously during the measurement.

In contrast to that approach, a better outcome can be attained when two measurements are performed, including all the errors of a test set-up "left over" by the optical compensation scheme (Fizeau or Twyman-Green or others). The first measurement is with a "calibration master" and the second with the test piece. The resulting error maps are stored in computer memory. Suppose that nothing but the master and test pieces have changed between the measurements, the difference of the two measurements should show only the difference in the shape of the master piece and the test piece. All other errors should cancel out by this "mathematical compensation for the intrinsic errors" of the test

set-up. Therefore, the master plays a similar role as the reference wavefront, however, in this case some higher order errors, like "defocus" discussed in the context with Figure 1, also cancel out due to the higher degree of "symmetry" compared to the normal Fizeau test.

The drawback with this method is that, in principle, two measurements are necessary to get the result which increases the statistical errors by a factor of $\sqrt{2}$, and that the two measurements cannot be performed both at the same time. In order that this mathematical compensation of errors be effective, it should be assured that the calibration measurement is still valid for the test piece. It is therefore highly recommended to take the calibration measurement at a time immediately before or after (or both) the measurement of the test surface.

This "mathematical compensation" can have some important advantages compared to the "optical compensation":

— Master piece and test piece may have the same radius of curvature, e.g. the imaging properties onto the detector are identical both times. Note that this can also be a disadvantage, since if the master piece has a radius different than the test piece, systematic errors will be introduced by the common path elements as described in the previous section.

— Influences of gravitational deformations, quasi-stationary temperature gradients, etc., are cancelled out to some extent.

— The reference piece may be made much thicker than the typical transmission sphere. If it is made more stable than the test piece, this might be helpful in some situations. For example, the "bending" of the test piece by the support will show in the result. In other cases, it may be possible to produce an almost identical master (but with a known error map) for calibration purposes.

— A greater flexibility exists for the design of the test configuration. For example, the reference surface does not have to be the last surface of a transmission sphere any longer. It could be a plane surface in front of the lens. In this way, "tilt" can be introduced into the test configuration, without causing uncompensated errors.

This last point is an important one, and will be explained later in greater detail.

5.2.5 Contrast as a function of the irradiance in test and reference arm: methods to attain equilibrium in both arms

It is well known (see ISO/TR 14999-1:2005, 3.2.1), that the intensity of two beam interference $I(x)$ is described by the formula

$$I(x) = I_1(x) + I_2(x) + 2\sqrt{I_1(x) I_2(x)}\,|\gamma| \cos\left[\frac{2\pi}{\lambda}\Delta l_{\mathrm{OPD}}(x)\right] \tag{2}$$

where

x is a vector describing the spatial coordinates, which may be on the detector or the surface under test;

Δl_{OPD} is the complete "optical path difference" of the two interfering beams, in principle from the light source all the way down to the detector (practically, in Fizeau interferometers, it is twice the "optical thickness" of the cavity, because of the (quasi) identical paths of the two beams through most parts of the interferometer, as was discussed before);

$|\gamma|$ is the magnitude of the complex degree of coherence; this quantity is close to unity in most cases when using a laser as the light source;

$I_1(x)$ and $I_2(x)$ are the intensities of the interfering wavefronts when measured separately.

The visibility V of the fringes is given by

$$V = \frac{2\sqrt{I_1 \cdot I_2}}{I_1 + I_2}|\gamma| \tag{3}$$

If we assume that $|\gamma| = 1$, then the visibility can be expressed as a function of the relative magnitude of the intensities of the two waves, e.g. as a function of

$$\eta = \frac{I_2}{I_1} \text{ with } I_2 \leq I_1 \tag{4}$$

Then, the visibility is

$$V = \frac{2\sqrt{I_1 I_2}}{I_1 + I_2} = \frac{2\sqrt{\frac{I_2}{I_1}}}{1 + \frac{I_2}{I_1}} = \frac{2\sqrt{\eta}}{1+\eta} \text{ for } 0 \leq \eta \leq 1 \tag{5}$$

Figure 2 gives the function $V = V(\eta)$. Table 1 provides related values.

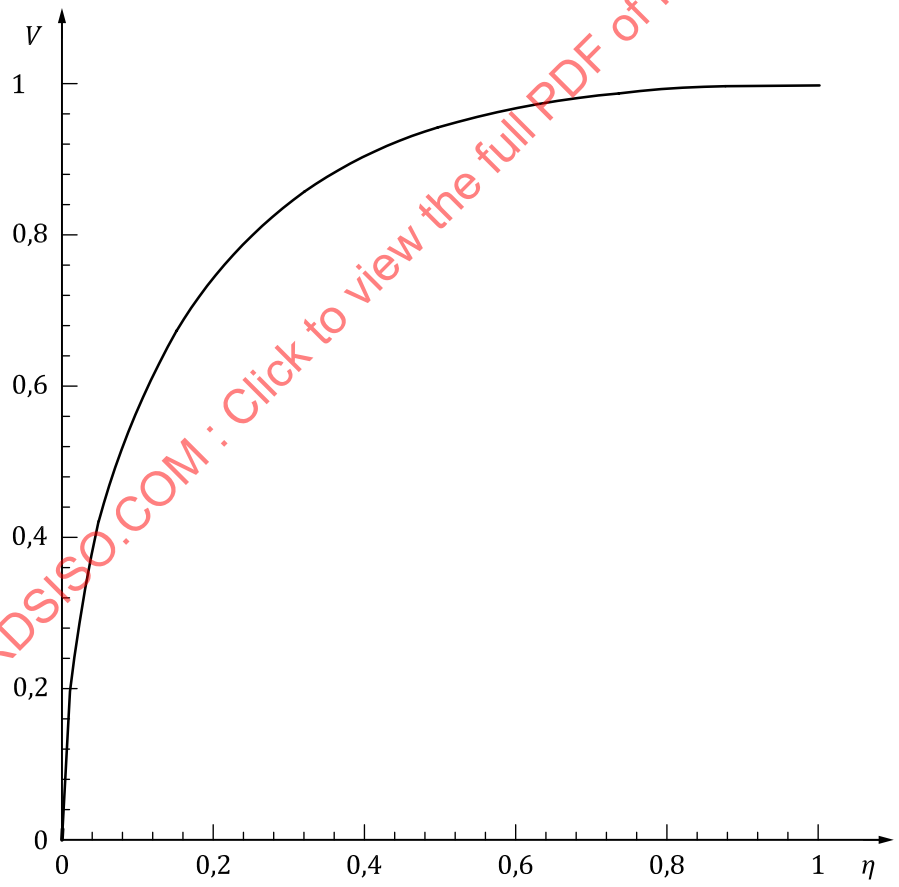


Figure 2 — Visibility of two-beam interference as a function of the ratio $\eta = I_2/I_1$ of the intensities of the two waves

Table 1 — Intensities I_1, I_2, intensity-ratio η , visibility V and number of bits lost in the detector's dynamic range when the intensities of the two-beam interference are not equal

I_1	I_2	η	V	bit
1	1	1,000	1,00	0,00
2	1	0,500	0,94	0,08
5	1	0,200	0,75	0,42
10	1	0,100	0,57	0,80
20	1	0,050	0,43	1,23
25	1	0,040	0,38	1,38
50	1	0,020	0,28	1,85
100	1	0,010	0,20	2,34
200	1	0,005	0,14	2,83
400	1	0,003	0,10	3,33

I_1	I_2	η	V	bit
1	1,00	1,000	1,00	0,00
1	0,90	0,900	1,00	0,00
1	0,80	0,800	0,99	0,01
1	0,50	0,500	0,94	0,08
1	0,10	0,100	0,57	0,80
1	0,06	0,060	0,46	1,11
1	0,04	0,040	0,38	1,38
1	0,03	0,030	0,34	1,57
1	0,02	0,020	0,28	1,85
1	0,01	0,010	0,20	2,34

It can be seen from Table 1 that the visibility drops to 38 % of its optimal value when the intensities differ by a factor of 25, e.g. when a highly reflective test surface is measured against an uncoated Fizeau reference plate. This corresponds to a 2,6× loss in dynamic range (or 1,38 bits). If the camera in the interferometer has a quantization into only 8 bits, which is the case in most systems, this might degrade the measurement precision unacceptably, whereas with cameras of 10 bits or 12 bits of quantization, the degradation might be acceptable. In any case, this effect can be compensated by raising the number of measurements to be included in one average by a factor of $N = 1/V^2 = (1 + \eta)^2/4\eta$, which is approximately seven in our example.

If this is not adequate, it will be necessary to adjust the intensities in the two beams, i.e., attenuate the reflectance of the higher reflecting surface. Possible ways to achieve this are discussed for two versions of a Twyman-Green configuration as well as for a Fizeau configuration. In the Twyman-Green configurations it is assumed that the beam splitter BS has no absorption and that the reflectance of the test-surface T and the reference surface R are ρ_T and ρ_R respectively. The absorbance for the attenuator is abbreviated by α_A, while η is the abbreviation for the ratio of the reflectances ρ_R / ρ_T .

Some possible interferometer configurations (two Twyman-Green, one Fizeau) are given in Figure 3. Figure 3 a) (hereafter referred to as "configuration 1") is a Twyman-Green type interferometer which includes an additional attenuation filter in the test arm. Figure 3 b) ("configuration 2") is a Twyman-Green interferometer using a beam splitter with unequal reflectivity and transmissivity (i.e. not 50 %). Figure 3 c) ("configuration 3") is a Fizeau interferometer that includes an attenuation filter between the test and reference surfaces.

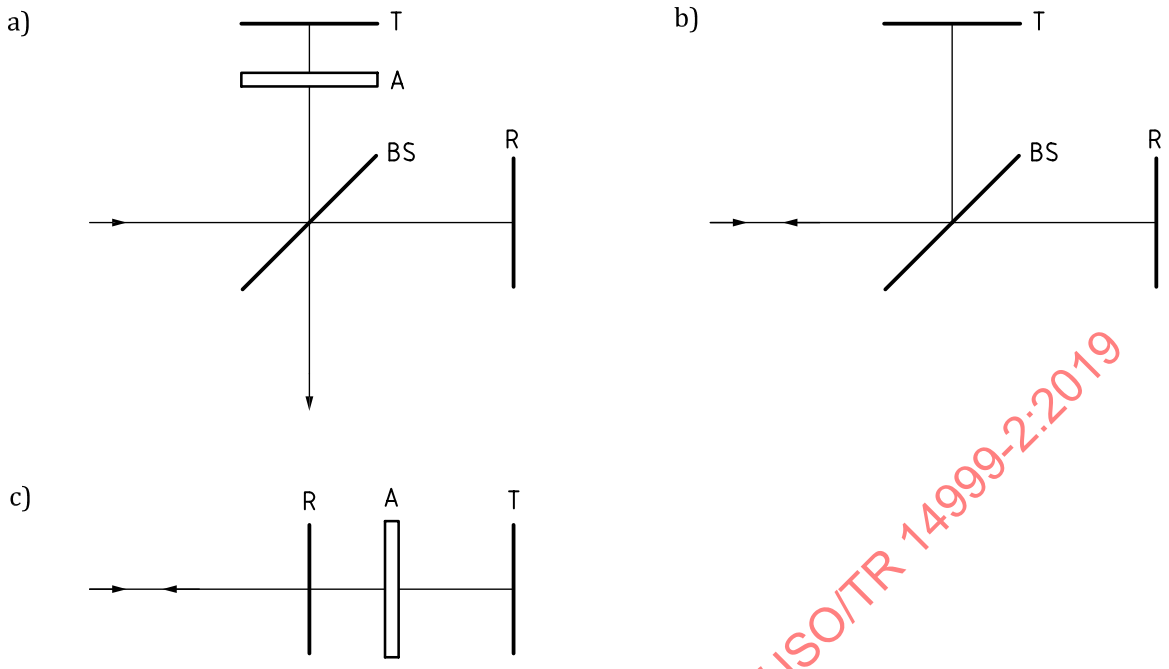


Figure 3 — Three configurations with unequal reflectance for the test surface T and the reference surface R; a) Twyman-Green with attenuation filter in the test arm; b) Twyman-Green using a beam splitter with unequal reflectivity and transmissivity; c) Fizeau interferometer with attenuation filter between the test and reference surfaces

The necessary attenuations and reflectance and transmittance values for the beam splitter are:

a) Configuration 1

$$\alpha_A = \sqrt{\frac{\rho_R}{\rho_T}} = \sqrt{\eta}$$

b) Configuration 2

$$\begin{cases} \rho_{BS} = \dfrac{\sqrt{\rho_T \rho_R} - \rho_R}{\rho_T - \rho_R} = \dfrac{\sqrt{\eta} - \eta}{1 - \eta} \\ \tau_{BS} = \dfrac{\rho_T - \sqrt{\rho_T \rho_R}}{\rho_T - \rho_R} = \dfrac{1 - \sqrt{\eta}}{1 - \eta} \\ \text{with } \eta = \dfrac{\rho_R}{\rho_T} \text{ and with } \rho_T \geq \rho_R \end{cases}$$

c) Configuration 3

$$\alpha_A = \sqrt{\frac{\rho_R}{\rho_T}} = \sqrt{\eta}$$

In the Fizeau configuration, the reflectance of the reference surface should not be higher than 4 %, to avoid the influence of multiple reflections on the fringe profile.

In the Twyman-Green configurations, only a pure two-beam interference will take place, whereas the Fizeau configuration suffers from the fact that multiple reflections within the cavity will occur. This effect can be neglected in most cases, provided that the reflectance of the reference surface (the so-called "transmission flat") is no greater than 4 %. For very high precision measurements, the true shape of the fringes should be taken into account for the evaluation formula used for phase-measuring analysis.

The attenuator plate A in Figure 3 should be of very high quality; normally an absorbing coating is applied onto a plane parallel glass plate or a pellicle. From Figure 3, it is clear that the attenuator plate A is an integral part of the cavity, and the influence of this plate (double transmission) to the transmitted wavefront will be part of the measured result, adding to the errors of the test piece. Thus it is important to verify that the phase variation introduced by the attenuator is sufficiently low to achieve the desired test accuracy, or that the phase variation can be calibrated and compensated in the measurement. For example, one such calibration method is to measure a test cavity with and without the attenuator in position, and use the difference between the measurements to compensate measurements made with the attenuator (the difference estimates the systematic phase error of the attenuator). Thus, it is possible to mathematically correct for the influence of the attenuator.

5.2.6 Contrast as a function of performance of the light source

5.2.6.1 General

The generalized interferometer may be considered as composed of

— an optical source,

— a collimating device,

— the interferometric device,

— a focusing device, and

— a detector.

The light emitted by the optical source is used to generate interference patterns and the detector collects the information, i.e., the interferogram.

Interferogram contrast depends on the following parameters:

a) spectral domain,

b) coherence, and

c) polarization of the light.

The light source needs to be chosen in accordance with the application.

5.2.6.2 Spectral domain

5.2.6.2.1 Source and detector

The user will choose the source to match its spectral emission range and the spectral sensitivity of the detector. The emission range should be close to the maximum sensitivity.

NOTE In the case of visual observation, a green (filtered mercury lamp) or a yellow source (sodium spectral lamp for close to null path difference) are the best choice (they correspond to the peak eye sensitivity); a red laser diode can be a poor choice for colour-defective observers.

5.2.6.2.2 Factors impacting transmission and reflection

Between the source and the detector, light will be subject to transmissions and/or reflections when travelling through the collimating, interferometric and focusing devices. These devices can distort the initial spectral emission range of the light passing through them.

NOTE For example, a gold-coated mirror has a poor reflectance in the blue and green spectral range; the reflectance increases in the yellow spectral range and is high in the red and infrared spectral ranges. With a silver coating, the reflection curve peaks in the yellow spectral range. With an aluminium coating, the visible spectral range is not affected.

Therefore, care should be taken when selecting the source.

5.2.6.2.3 Control requirements

If the sample is to be used with a pre-determined wavelength, the user has less freedom in the choice of source and detector, and will operate with the required wavelength. In some cases, the use of a substitute wavelength is possible, but in this case appropriate information about this should be included in the test report.

5.2.6.3 Temporal coherence

5.2.6.3.1 Pragmatic representation

Whatever the theoretical approach to describing the temporal coherence of the light, an elementary approach is to consider the beam as a light wave train. The train is limited in length equivalent to the coherence length.

The light wave train will travel from the source to the detector. Reaching the beam splitter, the train is divided into two “subtrains”, one going through the reference arm, and the other through the measurement arm of the interferometric device. After their separate paths, both subtrains reach the component used for the light wave train reconstruction (the beam splitter in some cases). The reconstructing optical component performs an addition of both subtrains, equal to the initial light wave train, if possible. The subtrains are fully superposed if the arm lengths are equal, or one subtrain is delayed if its arm is longer than the other one.

The interference contrast is maximum when the arm lengths are equal, but decreases as they become unequal.

NOTE The interference contrast obtained with a laser decreases when the path difference approaches the laser coherence length, but the coherence length can vary widely across laser types. Frequency-stabilized lasers can have tens of meters of coherence length; while others may have coherence lengths of centimeters or even sub-millimeter. Check the coherence length specifications of the laser, and ensure that it comfortably exceeds the path difference between the desired test and reference wavefronts.

The contrast obtained with a Na spectral lamp oscillates as the path difference varies, the period corresponds to a path difference of 0,6 mm due to the Na doublet.

The contrast obtained with a Hg low-pressure spectral lamp goes to zero, when the path difference becomes close to one centimetre.

The coherence obtained with a Hg high-pressure spectral lamp decreases very rapidly as the lamp temperature goes up, and is quickly reduced to less than one millimetre. Nevertheless, the user can maintain a longer coherence length by using a rheostat in the lamp power supply to cool the lamp during operation.

5.2.6.3.2 Sample surface contribution

The user needs to consider the question "Is the source coherence compatible with the peak to valley of the test piece?" although it is rare to be limited by the coherence when observing the defects of a test piece (the peak to valley would need to approach or exceed the coherence length — generally only an issue for very low coherence sources such as white light).

On the other hand, it is important to adapt the form of the comparison wavefront to the surface of the sample. This means, the user will adapt the shape of the wavefront to the sample shape, the goal being to keep the variations of path difference lower than the source coherence at all points.

5.2.6.3.3 Sample thickness contribution

When the measurement beam has to travel through the test piece, two effects need to be considered.

First, is the delay given to the "subtrain" going through the sample arm. Consider a sample with a thickness T and a refractive index n. It will introduce in the measurement arm a path difference of $2(n - 1)T$. This path difference can be compensated by moving the mirror of the reference arm away from the beam splitter.

Secondly, the sample can be considered as a thick window plate with parallel faces. This window plate gives an image from the mirror of the sample arm, translated from its initial position. The translation is $(n - 1)T/n$ in the direction of the beam splitter.

In total, the user wants to extend the reference arm, and the sample reduces the length of the measurement arm. The user needs to place the reference mirror in a compromise position, taking the coherence and the fringe visibility into account.

5.3 Test environment

5.3.1 General

The ideal interferometer system would be a system delivering the same correct results in any environment, independent of vibrations, temperature and temperature gradients, gravity and distribution of forces supporting the test piece. Such a "magic" interferometer could be thought of as having the ability to act as a filter, which completely blocks all environmental influences. At the same time, it would be completely responsive to all the surface form information of the test surface, sensing this to the highest precision, which might be in the range of sub-nanometres normally to the surface and more than 1 million measured points in the most demanding cases. Also, such an ideal instrument would not contribute any intrinsic errors or noise to the measurements.

In practice, it is impossible to block all environmental influences and intrinsic errors or noise by the instrument. But, still, there are a number of ways that a designer of an instrument can make it more (or less) robust against environmental influences.

It is also evident that the influence of the environment can be kept constant by keeping the environment itself constant. In a laboratory equipped with high-grade air conditioning, shielding against vibrations, acoustic noise, very stable and noise-free voltage supply, etc., and, ideally, if the tests were performed in a zero-gravity environment, it would be much easier to attain reproducible results. The quality of the laboratory environmental conditions will influence the precision of the result, but not necessarily accuracy; hence, generally the importance of a very good laboratory is overestimated.

However, besides the laboratory and the instrument, there is a third influence: the operator of the instrument. In most practical cases, the instrument and the laboratory will already be provided, and the operator will thus be the key person to influence the ultimate accuracy and precision within the given limitations.

5.3.2 Influence of vibrations

When testing optical surfaces, e.g. using a Fizeau interferometer as given in Figure 14 of ISO/TR 14999-1:2005, the two interfering waves are reflected back from the last surface V as well as the reference surface M. These two surfaces have to be mounted and this mount needs to provide the necessary stability. It is essential to ensure that the gap between the surfaces will not change unintentionally during the acquisition of the intensity data in a (conventional) phase measurement approach with temporal carrier as described in 6.4.

The reason for this is, that in the case of an additional change of the "interferometer cavity" due to vibrations of the mount carrying the parts, Formula (8) is no longer complete since $\varphi(t)$ now contains not only the wanted term as described in 6.4, but also an unwanted component describing the vibrations in the cavity. This "noise" term, φ_N, can, in the simplest case, be described by Formula (6).

$$\varphi_N = 2\left(\frac{2\pi}{\lambda}\right)\alpha \sin \omega_N t \tag{6}$$

where

α is the amplitude of the vibration;

ω_N is the (dominant) frequency.

Even if the noise might contain a complicated spectrum, in many cases it will be sufficient to restrict the consideration to one single frequency with the lowest mechanical eigenfrequency ω_N of the structure and the largest associated amplitude α.

Given a specific phase algorithm and given the frequency of the camera frames, it is possible to compute the amount of the phase errors due to α, under the assumption that α is small enough (on the order of $\lambda/10$) that the linear model can be applied. There are significant differences between the phase algorithms.

As interferometry can be used to measure deviations of wavefronts or surfaces as small as a few nanometres or even smaller, it is obvious that the test pieces need to be mounted in a stable manner. It is therefore common practice to install the interferometer on a passive or, even better, actively-controlled vibration isolated table, which acts as a low-pass filter to the vibrations of the floor of the laboratory. If possible, there should be no running machines in the vicinity of the laboratory. If an isolation table does not reduce the vibrations enough, it can sometimes be helpful to introduce a low-force mechanical contact between the reference and test parts (in a horizontal Fizeau set up). For example, a light plastic bar or piece of wood can be placed on the surfaces (or their mounts) to "bridge" them.

However, many circumstances exist where it is difficult to avoid sources of vibrations or to isolate well enough. In this case, it is useful to investigate the result of vibrations on the measurements in order

to optimize the given set-up. If the effect of vibration on the measurement is deemed too large, phase measurement devices and algorithms which are less sensitive to vibration can be considered. Two groups of algorithms can be distinguished here:

1) phase-measurement algorithms that do not rely on temporal phase shifting, such as spatial carrier frequency (see 6.5), and

2) algorithms which do not require *a priori* knowledge of the phase steps (the unknown differences between phase shifts are solved for).

The first approach is most suitable for cases where the vibration cannot be controlled, as the phase can be measured in a single fast exposure. The second approach is especially interesting in cases where already reasonably good starting values for the phase steps can be assured, which are then further refined in an iterative optimization process, i.e., when the vibration isolation is already very good but still the remaining disturbance is too high for the application.

It is very important for the user of a phase-measuring interferometer to know how to check the set-up for vibrations which might degrade the result. For a better understanding of the procedure described hereafter, it is useful to state that the phase errors $\Delta\phi(x,y)$ due to vibrations can be expressed by:

$$\Delta\phi(x,y)=\varepsilon\sin\left[2\phi(x,y)\right]+\ldots \text{ small higher-order terms} \tag{7}$$

where the amplitude ε of the sine function is dependent on a number of parameters as:

— the specific algorithm used;

— the amplitude of the vibration;

— the ratio of the dominant vibration frequency ω_N to the camera frame frequency ω_0, i.e., ω_N/ω_0;

— the exposure time of the camera.

It is interesting that the dominant error term in square brackets [ ] has as its argument twice the phase value to be measured, i.e., 2ϕ(x,y). Therefore, to decide whether ε is small enough to be ignored, the following procedure is useful:

1) Measure in a null configuration (e.g. <1 fringe in the interferogram) and save the phase measurement;

2) introduce fringes into the interferogram by tilting the object (e.g. ~5 fringes) and save a measurement in this configuration;

3) subtract the two measurements from each other;

4) remove a low-order polynomial from the difference (to remove the tilt and any contribution from air turbulence/thermal effects);

5) examine the fit residual for a sinusoidal-like pattern that looks like twice the number of fringes introduced by the earlier tilting step (e.g. ~10 periods)

6) the peak-valley of the sinusoidal pattern provides an estimate of the vibration error.

The above process can be repeated if a more statistically robust estimate of the vibration error is desired. Or other repeatability tests could be done. What is important is after you have an estimate of the vibration-induced error, you can decide whether it is at an acceptable level or whether vibration mitigation strategies are needed.

5.3.3 Influence of gravity and support of the test piece

The effect of the support of the test piece on the measuring result under the influence of gravity should be taken into consideration.

For many practical cases, the effect is quantitatively foreseeable by mathematical means. For practically edited details, see for example Reference [8]. Such an analysis is often a rather complicated procedure, which consists in solving differential equations or in FEM calculations. Fortunately, a detailed mathematical analysis is required only in cases where the highest measuring accuracy is required.

Some basic facts and rules of thumb are very useful to estimate the magnitude and help to avoid considerable measuring errors.

a) Optical materials vary in their stiffness by a factor of 10 (see Table 2). Optical glass BK7 has its place in the geometrical middle between the extrema. The sag varies between three times to one third of the value for BK7.

b) From mechanics it is known that the sag of a twofold supported beam (or straight-edge) is minimized if the support points are arranged at nearly one fifth of the beam length away from the beam ends (the so called minimum deflection points separated by 0,55 length). This rule is useful if a horizontally arranged standard surface is to be optimally supported. More detail on low deformation support design is presented in Reference [8].

c) Table 2 shows some numerically calculated values for the sag of uniformly loaded (gravity) circular plates supported continuously around their edges, using the formula in Reference [8] on p. 288. It shows the effect of different materials (A in Table 2), thickness according to the common rule "D/d = 8" (B in Table 2), and the required thickness to result in an approximately constant sag of 30 nm (C in Table 2).

Table 2 — Sag of uniformly loaded circular plates

Effect	Material	Diameter mm	Thickness mm	Sag nm
A	LF5	120	10	52,6
	SF59	120	10	112,7
	Sapphire	120	10	11,1
	BK7	120	10	30,5
B	BK7	20	2,5	0,4
	BK7	60	7,5	3,4
	BK7	120	15	13,5
	BK7	200	25	37,6
	BK7	300	40	74,4
	BK7	400	50	150,5
C	BK7	20	0,28	30,0
	BK7	60	2,5	30,5
	BK7	120	10	30,5
	BK7	200	28	30,0
	BK7	300	63	30,0
	BK7	400	112	30,0

NOTE The material properties used to calculate the values in Table 2 are shown in Table 3.

Table 3 — Material properties

Material	Density (g/cm^3)	Young's modulus E (GPa)	Poisson ratio μ
LF5	3,22	59	0,226
SF59	6,26	51	0,269
Sapphire	3,97	335	0,25

Table 3 *(continued)*

Material	Density (g/cm^3)	Young's modulus E (GPa)	Poisson ratio μ
BK7	2,51	81	0,208

6 Methods for evaluating the optical path difference

6.1 General

Various procedures exist for the accomplishment of interferometric measurements. This clause provides an overview of them, together with information-associated errors from random and systematic sources, as well as certain specific procedures for high-accuracy measurements.

The various methods for the evaluation of interferograms can be divided into the following.

a) Visual or manual procedures giving qualitative information about shape of the test wavefront (or surface) and simple characteristic values (e.g. PV value).

b) Phase measuring procedures yielding a complete quantitative description of the wavefront (or surface) under test.

6.2 Visual inspection of interferograms

6.2.1 General

An interferogram arises from a comparison of a test wavefront (surface) with a known reference wavefront (surface). The shape and orientation of the reference wavefront should be adequately chosen, so that the ideal test wavefront leads to a typical interference pattern (e.g. straight fringes) which is easy to interpret.

By observing the deviation of the actual interferogram from this ideal pattern, the various optical aberrations of wavefronts, as well as typical wavefront and surface deviations can be recognized and classified by their visual appearance.

In addition, it is important to gain information about the sign of the wavefront aberration, i.e., whether the deviation is a peak or a valley, or whether a test wavefront (surface) is convex or concave with respect to a reference wavefront (surface). This phase information is not contained in a static interference pattern or interferogram. To obtain it, it is necessary to ascertain and record the location of the zero order of the interference pattern, i.e., the location where the optical path difference between the interfering wavefronts is zero. Knowing the location of the zero order, and applying a variation of the optical path difference in a known way, it is possible to deduce the sign of the wavefront (surface) deviation by observation of the direction of motion of the fringes:

— If the optical path difference is decreased, the fringes move away from the direction of the zero order (for tilted wavefronts: to the thick portion of the wedge).

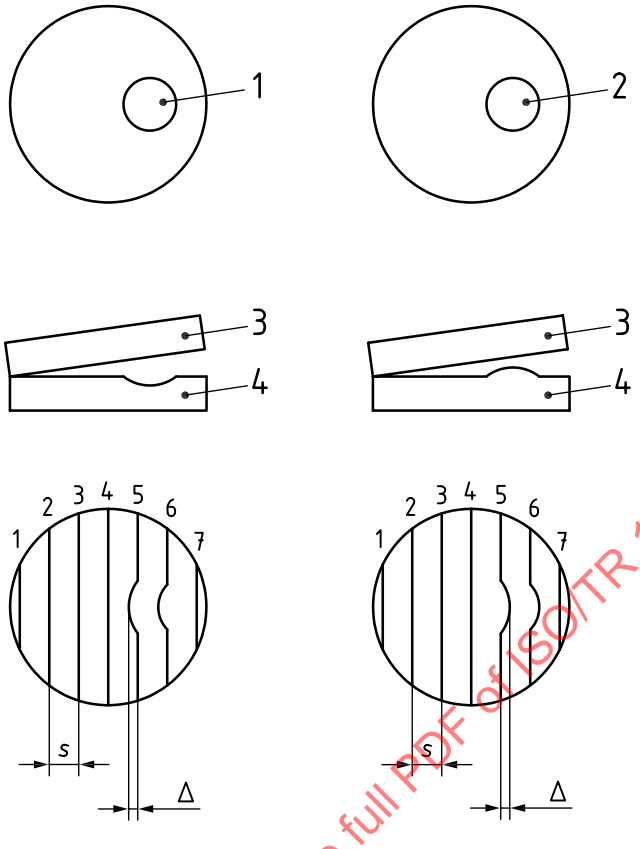
— If the optical path difference is increased, the fringes move toward the direction of the zero order (for tilted wavefronts: to the thin portion of the wedge).

— If the tilt between the wavefronts is increased, the fringe spacing is decreased and the fringes move toward the zero order fringe.

If possible, the interferometer should be adjusted "with tilt", so that the path difference contains an air wedge to obtain nearly straight fringes.

6.2.2 Example 1 — Fizeau interferometer

A hole will cause fringes to move toward the thin portion of the air wedge (Figure 4).

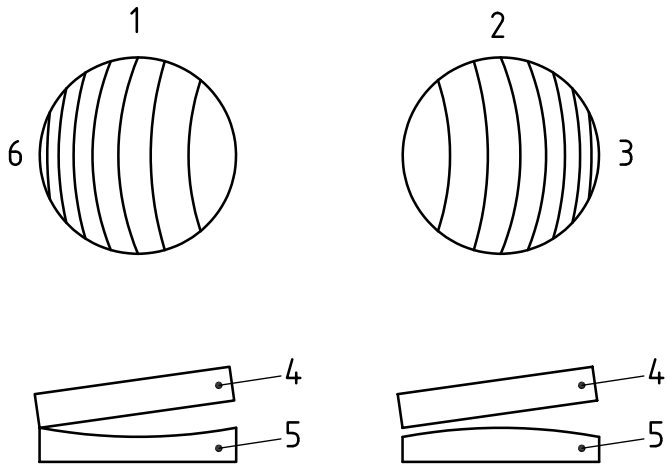


Key

1 hole
2 bump
3 reference surface
4 test surface
s fringe spacing in the best-fitting pattern
Δ maximum of the fringe pattern from the best-fitting pattern

Figure 4 — Movement of fringes in case of a hole and a bump, respectively

For a concave surface, the fringes at the edge of the pattern curve toward the thick portion of the wedge; for a convex surface the fringes curve toward the thin portion of the air wedge (Figure 5).



Key

1 concave sphere
2 convex sphere
3 thick part of wedge
4 reference flat
5 test surface
6 thin part of wedge

Figure 5 — Movement of fringes in case of a concave and a convex surface, respectively

Figure 5 assumes that the wedge angle between the two wavefronts being compared is larger than the wavefront errors, and thus no closed fringes are obtained. The wedge angles in the drawings are greatly exaggerated.

If the wedge direction is not known, it can be determined as described in Table 4.

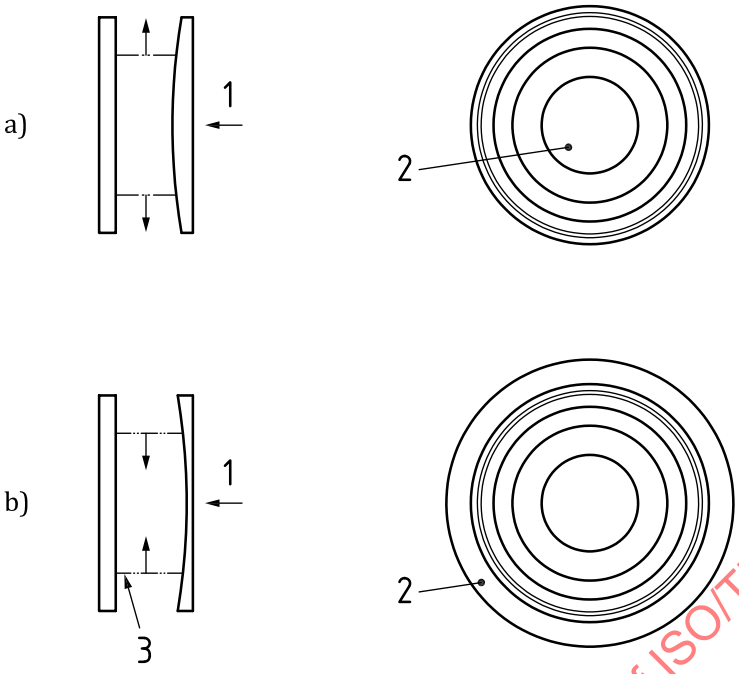
For the general case, where the fringes are neither circular nor straight, a combination of these facts applies. Firstly, the direction of the zero order as described for straight fringes needs to be determined. Secondly, the direction the fringes curve relative to the zero order needs to be observed to determine whether the wavefront/surface is convex or concave. This also applies to local deviations.

In the case of spherical wavefronts or surfaces, the sign of the optical path difference, Δl_{OPD}, indicates that the test wavefront is either concave or convex relative to the reference wavefront, which is important in the fabrication of optical surfaces as well as in the assembly of optical systems. Table 4 also contains the determination of the relative curvatures of test and reference surface (wavefront) in a wedge-free set-up (e.g. proof glass). If the test surface is convex relative to the reference surface, a decrease in air gap thickness causes the fringes to move out, away from the zero order [see Figure 6 a)]. In this case, the radius of curvature of the test surface is longer than the radius of the reference surface. If on the other hand, the test surface is concave relative to the reference surface, a decrease in air gap thickness causes the fringes to move in, i.e., away from the zero order, and the radius of curvature of the test surface is shorter than the radius of the reference surface [Figure 6 b)].

In these considerations, the expression “surface” may be exchanged by “wavefront”.

Table 4 — Determination of the sign of a wavefront deviation in a Fizeau interferogram

a) Interferometer set-up with tilt (containing air wedge)	
Wedge direction known	
Hole	Fringe deviation from straightness toward thin portion of wedge
Bump	Fringe deviation from straightness toward thick portion of wedge
Concave sphere	Fringes curve toward thick portion of wedge
Convex sphere	Fringes curve toward thin portion of wedge
Wedge direction not known	
— Push on edge of test piece to determine direction of wedge Number of fringes increase: pushing on thin portion of wedge Number of fringes decrease: pushing on thick portion of wedge	
— Pushing uniformly on test piece to decrease air gap thickness Closed fringes: fringes move toward low points Open fringes: fringes move from higher regions to lower regions	
— Pushing uniformly on test piece to increase air gap thickness Closed fringes: fringes move toward high points Open fringes: fringes move from lower regions to higher regions	
b) Interferometer set-up without tilt (air wedge), comparing spherical surfaces	
Test wavefront (surface) radius > reference radius	
— When decreasing air gap thickness: fringes move out, away from the zero order	
— Fringes are curved towards the zero order	
— Wavefront (surface) convex to reference	
Test wavefront (surface) radius < reference radius	
— When decreasing air gap thickness: fringes move in, away from the zero order	
— Fringes are curved away from the zero order	
— Wavefront (surface) concave to reference	



Key

1 push
2 zero order
3 constant distance corresponding to fringe

Figure 6 — Fringe movement

6.2.3 Example 2 — Twyman-Green interferometer

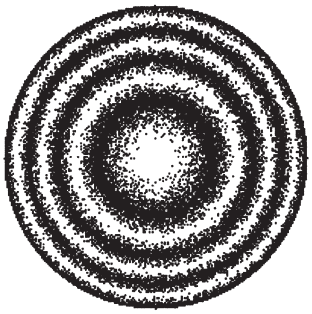
Table 5 outlines several common techniques for determining the sign of the error.

Table 5 — Determination of the sign of a wavefront deviation in a Twyman-Green interferogram

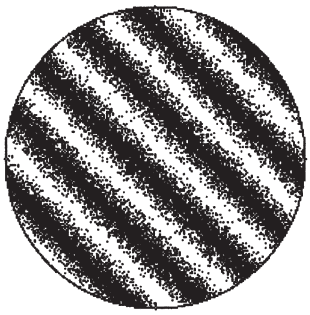
a)	**Push in on reference mirror** 1) Closed fringes — Surface test: fringes move toward high points — Transmission test: fringes move toward thin regions 2) Open fringes — Surface test: fringes move from low regions to high regions — Transmission test: fringes move from thick regions to thin regions
b)	**Push out on reference mirror** Opposite of a)
c)	**Push in on test mirror** Opposite of a)
d)	**Push out on test mirror** Same as a)
e)	**Put hot wire in test beam** 1) Closed fringes — Surface test: in vicinity of the wire fringes move toward lower regions — Transmission test: in vicinity of the wire fringes move toward thick regions 2) Open fringes — Surface test: in vicinity of the wire fringes move from high regions to low regions — Transmission test: in vicinity of the wire fringes move from thin regions to thick regions
f)	**Put hot wire in reference beam** 1) Closed fringes: opposite of e1) 2) Open fringes: opposite of e2)

Note that a change in spacing without a change in relative tilt between the wavefronts is nearly impossible to achieve.

When testing optical systems, the visual inspection of the interferogram allows (for simple cases and small aberrations) a rough statement about the nature of the aberration. Some typical cases are illustrated in Figures 7 and 8.



a) Pure defocus

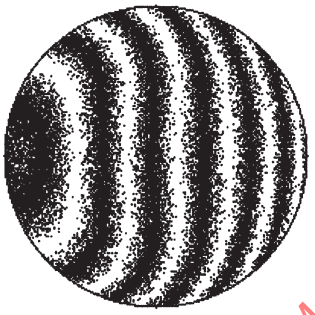


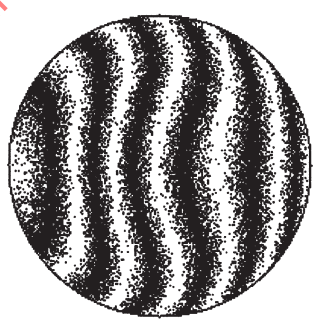
b) Pure tilt

c) Defocus + tilt

NOTE Interferograms show adjustment of interferometer.

Figure 7 — Interferograms for a system under test with only alignment errors (negligible aberrations)





a) Spherical aberration

b) Astigmatism

c) Coma

Figure 8 — Interferograms for a system under test with small aberrations

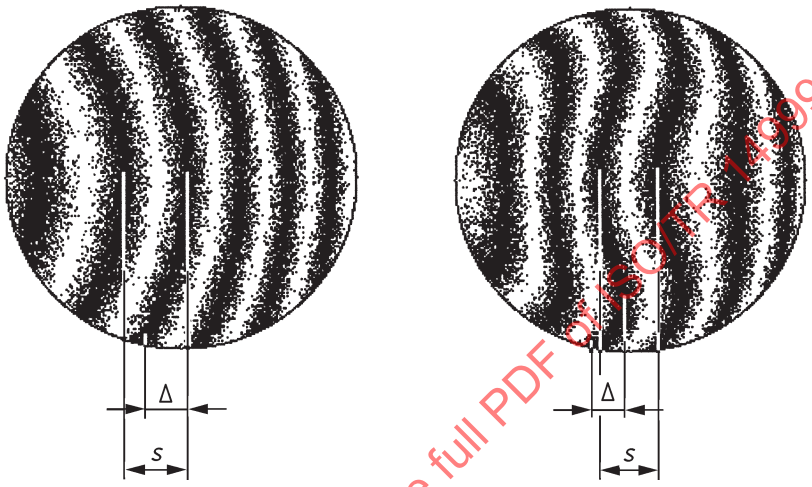
6.3 Manual evaluation of interferograms

To evaluate an interference fringe pattern, it is necessary to quantify the deviation of the fringe pattern from some ideal best-fitting pattern, which is expected for some specific (best-fitting) test wavefront.

Depending on this theoretical wavefront, several kinds of deviations can be extracted from the interferogram, i.e., wavefront sagitta error, irregularity and rotationally symmetric irregularity (see ISO 10110-5:1996[1)], Annex B). Ideally, the fit should be based on a least-squares computation.

In practice, try to adjust for straight fringes (minimum sagitta error), so that the estimation, respectively manual evaluation, can be done with the help of a ruler or specially designed hinged mechanical parallelogram. If this is not possible (e.g. when using proof glasses), the central fringe should not be compared to a straight line, but to the circular arc joining the two ends and the centre of the central fringe.

The deviation is usually presented as a fraction of the spacing between a pair of fringes in the ideal pattern. See Figure 9.



Key

s fringe spacing in the best-fitting pattern

Δ maximum of the fringe pattern from the best-fitting pattern

Figure 9 — Manual evaluation of interferograms

Quantitatively, Δ/s is the fractional deviation in the fringe spacing, where s is the fringe spacing in the best fitting pattern, and Δ is the maximum deviation of the fringe pattern from the best-fitting pattern. This corresponds to the maximum distance between the test and reference wavefront minus the minimum distance between the wavefronts, which is referred to as the peak-to-valley difference. In this case, where the wavefronts intersect, the minimum distance between the wavefronts is a negative number; take the sign into account in the computation of the peak-to-valley difference. If one of the wavefronts represents the theoretical planar wavefront, then the peak-to-valley difference between the wavefronts is called the "peak-to-valley deformation".

The deviation, in units of wavelengths, is equal to the product of Δ/s and the "interferogram scale factor".

Deviation = $(\Delta/s) \times$ (interferogram scale factor) $\times \lambda$

For example, when evaluating a mirror surface at normal incidence in a Fizeau or Twyman-Green interferometer, the scale factor is 1/2: one fringe spacing = $(\lambda/2)$.

Table 6 gives the interferogram scale factors for some common interferometric set-ups.

1) Withdrawn.

Table 6 — Interferogram scale factors

Measured parameter	Interferometric set-up	Interferogram scale factor
Reflective surface	Normal incidence, single pass	1/2
Reflective surface	Angle of incidence = Θ, single pass	$1/(2\cos\Theta)$
Reflective surface	Normal incidence, double pass	1/4
Reflective surface	Angle of incidence = Θ, double pass	$1/(4\cos\Theta)$
Single pass transmitted wavefront deviation	Single pass	1
Single pass transmitted wavefront deviation	Double pass	1/2
Double pass transmitted wavefront deviation	Single pass	2
Double pass transmitted wavefront deviation	Double pass	1

When characterizing a surface form deviation in units of fringe spacings, the word "fringe spacing" does not refer to the transverse distance between fringes, but to the fact that the number of fringe spacings visible in the interference pattern corresponds to the number of half-wavelengths of surface form deviation.

The transmitted wavefront quality of an optical element is judged by the deformation of a planar or spherical wavefront (depending on whether an infinite or finite conjugate is to be evaluated) transmitted twice through that element. This deformation is measured along the direction of propagation, that is, perpendicular to the nominal plane of the wavefront. The unit of measure is also a "double-pass fringe spacing", and it is equal to one wavelength of the light.

The rms measures of wavefront deviation cannot be determined by visual inspection or this elementary manual evaluation.

The analysis of fringe patterns is treated more fully in many textbooks, such as Reference [9].

6.4 Phase measurements with temporal carrier

6.4.1 General

The basic procedures for the analytical evaluation of interference patterns are described in 6.4.2 to 6.4.4, which use a temporally varying reference phase $\varphi(t)$ to measure the phase $\Phi(x,y)$ of the wavefront under test (without sign and phase ambiguity). The description is restricted to the most common case of two-beam interferometry.

The intensity $I(x,y)$ of the interference pattern, recorded on a uniform rectangular grid, obeys the following formula (see ISO/TR 14999-1:2005, 3.1.1):

$$I(x,y)=I_0(x,y)\left\{1+V(x,y)\cos\left[\Phi(x,y)-\varphi(t)\right]\right\} \tag{8}$$

where

$I_0(x,y)$ is the nearly uniform background intensity;

$V(x,y)$ is the visibility (contrast) of the fringes;

$\Phi(x,y)$ is the phase distribution to be measured;

$\varphi(t)$ is the reference phase.

Since there are three unknowns in the interference equation (I_0, V, Φ), a minimum of three intensity measurements for each location (x,y) is necessary to determine the phase $\Phi(x,y)$.

It is assumed that the relative orientation of both test and reference wavefront is such that the difference $\Delta\Phi$ between neighbouring points on the grid does not exceed π, so there are at least two

detector locations for each fringe (sampling according to the Nyquist frequency). Contrary to the manual evaluation methods, no other requirements exist for the relative adjustment of the wavefronts, since in the quantitative description of the test wavefront, tilt and defocus effects can be subtracted.

The temporal modulation of the reference phase is accomplished by continuously or stepwise changing the optical path of one of the interferometer arms. Among various possible devices, the most common are:

— moving reference mirror [mounted on a piezoelectrically driven transducer (PZT)];

— wavelength shifting;

— tilted glass plate;

— moving grating;

— rotating polarizers in conjunction with a polarizing dependent interferometer set-up;

— acousto-optical modulator.

These devices are used to either introduce a phase difference between the beams in discrete steps, or to vary the phase difference continuously, which corresponds to a frequency shift.

Depending on the phase variation, several methods can be distinguished (see References [10] and [11]):

— heterodyne interferometry ($\varphi = \omega\, t$);

— phase lock interferometry ($\varphi = a\sin\omega\, t + \varphi_0$);

— phase stepping and phase shifting interferometry [$\varphi = (r - 1)\varphi_0$].

NOTE The method described in 6.3 uses a spatially varying phase difference $\varphi = p_0 + p_1 x$, corresponding to a tilt between reference and test wavefront to produce wedge-type straight fringes of a certain spatial frequency. More sophisticated methods relying on spatial phase variation are described in 6.5.

6.4.2 Heterodyne interferometry

The phase difference ($\varphi = \omega\, t = 2\Delta f t$) between the interfering beams is varied by means of a continuous phase modulator, such as a rotating radial grating, acousto-optic modulator or rotating birefringent device. Since a continuously increasing phase difference corresponds to a second frequency, it is also possible to use a 2 frequency laser. The modulated fringe intensity is photoelectrically detected. At each location (x,y) of the interference pattern, the detector measures a signal at the beat frequency Δf, provided its bandwidth is high enough. The phase measurement can then be performed electronically by comparing two such signals.

One possibility is to use a detector at rest at the location (x_0,y_0) and another detector which is scanned over the whole interference pattern. For each position (x,y) of the scanning detector the phase difference $\Delta\Phi = \Phi\,(x,y) - \Phi\,(x_0,y_0)$ is determined by an electronic phasemeter. By integrating these differences over the whole interference pattern, the optical phase difference between the wavefronts can be measured without modulo-2π -ambiguity and including the sign information.

Another possibility is to use a second detector seeing the interference of light beams which are uninfluenced by the probe. The phase of this detector signal is fixed and serves as a reference in the electronic phase measuring.

6.4.3 Phase lock interferometry (PLI)

In phase lock interferometry, the phase difference between the interfering beams is modulated sinusoidally with an additional adjustable phase offset φ_0:

$$\varphi(x,y,t) = a\sin\omega t + \varphi_0(x,y,t) \tag{9}$$

A simple possibility to achieve this is a piezoelectrically driven reference mirror. The resulting intensity detected by a photoelectric sensor is

$$I(x,y,t) = I_0\left\{1 + V\cos\left[\Phi(x,y) - \varphi_0(t) - a\sin\omega t\right]\right\} \tag{10}$$

The formula for the interference intensity $I(t)$ at a given position (x,y) contains terms of the type $\cos[a\sin(\omega t)]$ and $\sin[a\sin(\omega t)]$. These terms can be expanded into a series with Bessel functions as coefficients:

$$I_{xy}(t) = I_0 + I_0 V\cos(\Phi - \varphi_0)\left\{J_0(a) + 2J_2(a)\cos(2\omega t) + \ldots\right\}$$

$$-I_0 V\sin(\Phi - \varphi_0)\left\{2J_1(a)\sin(\omega t) + 2J_3(a)\sin(3\omega t) + \ldots\right\} - \tag{11}$$

The frequency spectrum of the detected signal U contains components U_0, U_ω, $U_{2\omega}$, ... at the multiples of the modulation frequency ω. For small modulation amplitudes a, these components are:

$$U_\omega \approx I_0 V a\sin(\omega t)\cdot\sin(\Phi - \varphi_0) \tag{12}$$

$$U_{2\omega} \approx \frac{1}{4} I_0 V a^2\cos(2\omega t)\cdot\cos(\Phi - \varphi_0) \tag{13}$$

The signal component U_ω vanishes at the locations of a fringe maximum or minimum, whereas $U_{2\omega}$ reaches a maximum value:

$$U_\omega = 0\,;\; U_{2\omega} = \max. \tag{14}$$

U_ω can be separated by a suitable electronic filter. With a phase sensitive detector, the offset phase φ_0 can be steered in such a way that $U_\omega = 0$ is locked. So the change of Φ with varying positions (x,y) of the scanning detector is monitored by a variation of φ_0, since the electronic circuit maintains the condition $\varphi_0 = \Phi$.

The main features of this type of interferometry are as follows.

a) The phase detection is very sensitive to phase differences of the whole apparatus during the scanning procedure; the achievable rms accuracies are practically limited to $\lambda/30$.

b) The scanning region should be determined beforehand (with some type of interactive software) to avoid an interruption of the phase tracking if the rim of the interferometric aperture is crossed.

6.4.4 Synchronous detection and phase-shifting interferometry

6.4.4.1 General

Since the interference equation [see Formula (8)] contains three unknowns (I_0, V, Φ), a minimum of three discrete intensity measurements for each location (x,y) in the interference pattern is necessary to determine the phase $\Phi(x,y)$. For more than three measurements the technique corresponds to a least squares fitting procedure with enhanced accuracy.

The phase difference between the two beams needs to be stepped, preferably the reference phase is shifted such that R measurements are equally spaced over one or more modulation periods:

$$\varphi_r = \frac{(j-1)2\pi}{J} \text{ with } j = 1,\ldots,\mathrm{J} \tag{15}$$

To solve the resulting equation system, the interference equation [Formula (8)] is rewritten in the following form:

$$I(x,y) = L + M\cos\varphi + N\sin\varphi \tag{16}$$

where

L $= I_0$;

M $= I_0 V \cos\Phi$;

N $= I_0 V \sin\Phi$ multiplied by $\cos\varphi_r$ respectively $\sin\varphi_r$ and summed up over one period,

and where

φ_r $= (r-1)\cdot\varphi_0$ and $r = 1, \ldots, R$.

It follows:

$$M \propto \sum_{r=1}^{R} I_r \cos\varphi_r$$

$$N \propto \sum_{r=1}^{R} I_r \sin\varphi_r \tag{17}$$

where I_r being the intensity value belonging to the reference phase value φ_r.

The following formula for the phase Φ holds:

$$\Phi(x,y) = \arctan\frac{N}{M} = \arctan\frac{\sum_r I_r(x,y)\sin\varphi_r}{\sum_r I_r(x,y)\cos\varphi_r} \quad (\mathrm{mod}\ \pi) \tag{18}$$

The multiplication with $\sin\varphi_r$ and $\cos\varphi_r$ corresponds to a synchronous detection of the cosine type signal $I_r(\varphi_r)$.

Depending on the actual number of phase steps, there are several common schemes.

6.4.4.2 Four-step technique

With $\varphi_r = 0, \pi/2, \pi, 3\pi/2$, Formula (18) simplifies to:

$$\Phi(x,y) = \arctan\frac{I_2(x,y) - I_4(x,y)}{I_1(x,y) - I_3(x,y)} \tag{19}$$

The fringe visibility can be written:

$$V(x,y) = \frac{\sqrt{(I_1 - I_2)^2 + (I_1 - I_3)^2}}{2I_0} \tag{20}$$

For each detector element, the value $V(x,y)$ gives information whether the phase measurement $\Phi(x,y)$ is reliable or not. In an automated measuring procedure it can be used for the determination of the rim of the interference aperture or detection of dust particles, which do not exhibit the sinusoidal φ_r-dependence.

6.4.4.3 Four-bucket technique

Concerning the measurement time, the phase stepping technique is not the fastest possible procedure, because the phase stepper (for instance a piezo-driven mirror) requires a certain time to settle down after each step. In a modified scheme the phase is shifted continuously during the detector's integration time. The phase change φ_r is divided into R sections in which the relative phase changes by an amount $\Delta = 2\pi/R$. For each section the integrated intensity is:

$$I_r(x,y) = \frac{1}{\Delta}\int_{\varphi_r - \Delta/2}^{\varphi_r + \Delta/2} I_0(x,y)\left\{1 + V\cos\left[\Phi(x,y) - \varphi(t)\right]\right\} \mathrm{d}\varphi(t) \tag{21}$$

After integrating this expression, the recorded intensity is:

$$I_r(x,y) = I_0(x,y)\left\{1 + V\,\mathrm{sinc}\left(\frac{1}{2}\Delta\right)\cos\left[\Phi(x,y) + \varphi_r\right]\right\} \tag{22}$$

where $\mathrm{sinc}\left(\frac{1}{2}\Delta\right) = \sin\frac{1}{2}\Delta \Big/ \frac{1}{2}\Delta$.

So the only difference of this method to phase stepping is a reduction in the detected fringe modulation. For example, the visibility is reduced by a factor of 0,9 for $\Delta = \pi/2$ (e.g. for 4 buckets).

NOTE The Formulae (21), (22) assume the detector is exposed for the entire duration of the phase shifting. This is not possible for all detectors; e.g. a shutter may be engaged during data is read out, thus part of the phase shift ramp will not be integrated, and the integration range would be less than Δ.

6.4.4.4 Three-step technique

There are several ways of implementing a three-step technique, the most common use phase steps of 60°, 90°, or 120°.

$$\Delta\varphi_r = 60° = \pi/3 \tag{23}$$

$$\Phi(x,y) = \arctan\left[\frac{2I_1 - 3I_2 + I_3}{\sqrt{3(I_2 - I_3)}}\right] \tag{24}$$

$$\Delta\varphi_r = 90° = \pi/2 : \varphi_r = \pi/4, 3\pi/4 \text{ and } 5\pi/4 \tag{25}$$

$$\Phi(x,y) = \arctan\left(\frac{I_3 - I_2}{I_1 - I_2}\right) \tag{26}$$

$$V = \frac{\sqrt{(I_1 - I_2)^2 + (I_2 - I_3)^2}}{2I_0} \tag{27}$$

$$\Delta\varphi_r = 120° = 2\pi/3 : \varphi_r = 0, 2\pi/3 \text{ and } 4\pi/3 \tag{28}$$

$$\Phi(x,y) = \arctan\left(\sqrt{3}\frac{I_3 - I_2}{2I_1 - I_2 - I_3}\right) \tag{29}$$

$$V = \frac{\sqrt{3(I_1 - I_2)^2 + (2I_1 - I_2 - I_3)^2}}{2I_0} \tag{30}$$

For a phase difference other than π/2 or 2π/3, the phase can be calculated using

$$\Phi(x,y) = \arctan\left(\frac{1-\cos\varphi}{\sin\varphi}\frac{I_1 - I_3}{2I_2 - I_1 - I_3}\right) \tag{31}$$

where phase shifts $-\varphi$, 0 and φ are assumed.

6.4.4.5 Carré technique

In the previous technique, the phase difference is known. With four steps, it is also possible to determine the unknown but constant phase difference φ_r

$$\tan\left(\frac{\Delta\varphi_r}{2}\right) = \sqrt{\frac{3(I_2 - I_3) - (I_1 - I_4)}{(I_2 - I_3) + (I_1 - I_4)}} \tag{32}$$

The phase Φ at each point is

$$\Phi(x,y) = \arctan\left\{\tan\left(\frac{\varphi_r}{2}\right) \times \frac{(I_2 - I_3) + (I_1 - I_4)}{(I_2 + I_3) - (I_1 + I_4)}\right\}$$

$$= \arctan\sqrt{\frac{\left[3(I_2 - I_3) - (I_1 - I_4)\right]\left[(I_2 - I_3) + (I_1 - I_4)\right]}{(I_2 + I_3) - (I_1 + I_4)}} \tag{33}$$

For this technique the intensity modulation is

$$V = \frac{1}{2I_0}\sqrt{\frac{\left[(I_2 - I_3) + (I_1 - I_4)\right]^2 + \left[(I_2 + I_3) - (I_1 + I_4)\right]^2}{2}} \tag{34}$$

Here φ is assumed to be near π/2. If the phase difference is off by ±10 % the estimation of V will be off by ±10 %.

The technique works with a phase difference that does not have to be calibrated. The phase $\Phi(x,y)$ at each point in the interferogram can be determined modulo 2π, even in the case of varying phase calibration differences across the beam.

6.4.4.6 Calibrating the phase shifting device

The accuracy of the above methods depends on the linearity of the phase shifting device. For the typical phase shift from the piezoelectric transducer (PZT), the linearity depends on the actual driving voltage due to hysteresis and other nonlinear response effects. Therefore, a calibration of the phase-shifting equipment is necessary.

One possibility is given by the Carré technique already mentioned. A set of four intensity values is used to calculate $\Delta\varphi$. The phase shift controller should be adjusted so that the average phase shift is at the desired value and the spread in calculated phase shifts is small.

A simpler formula can be used when taking five intensity measurements with a constant phase shift between them (see Reference [12]):

$$\Delta\varphi_r = \arccos\left[\frac{1}{2}\left(\frac{I_5 - I_1}{I_4 - I_2}\right)\right] \tag{35}$$

With some prior knowledge of φ, the sign ambiguities of the cosine function can be resolved.

With a separate small interferometer fed by light split off from the reference arm, and with a separate detector, the phase can also directly be measured.

With the measured values, corrected values for the phase shifting device can be obtained (for instance a lookup table for the input voltages of the piezo driving high voltage amplifier, if a computer controls the phase shifting device).

6.5 Phase measurements with spatial carrier

6.5.1 Fringe analysis by Fourier transform operations

If the spatial frequency v_0 of the interference pattern obeys the condition $v_0 > \max|\mathbf{grad}\Phi|$, no closed fringes are present in the interferogram and one-dimensional Fourier transforms can be used to obtain the phase Φ as described in Reference [13].

The interference Formula (8) can be rewritten as follows:

$$I(x,y) = a(x,y) + c(x,y)\boldsymbol{e}^{j2\pi v_0 x} + c^*(x,y)\boldsymbol{e}^{-j2\pi v_0 x} \tag{36}$$

where $c(x,y) = \frac{1}{2} I_0 V e^{j\Phi(x,y)}$.

After a one dimensional Fourier transform the following holds:

$$F(I) = A(v,y) + C(v - v_0, y) + C^*(v + v_0, y) \tag{37}$$

where A, C and C^* are complex Fourier amplitudes.

By means of digital filtering, one side-band of the intensity distribution is filtered out and undergoes an inverse fast Fourier transformation (FFT). In this way the quantity $c(x,y)$ is obtained, from which the phase can be calculated:

$$\Phi(x,y) = \arctan\frac{\operatorname{Im} c}{\operatorname{Re} c} \quad \bmod \pi \tag{38}$$

or by using the complex logarithm:

$$\log\left[c(x,y)\right] = \log\left[\frac{1}{2}(I_0 V)\right] + j\Phi(x,y)$$

If the sampling theorem is fulfilled (more than two detector elements per period of the cosine function), phase ambiguities can also be eliminated (see 6.4.4).

A two-dimensional approach of the FFT technique (see Reference [14]) can also deal with interference patterns, where the fringe density is smaller than the maximum slope of the wavefront: $v_0 < \max|\mathbf{grad}\Phi|$. For example, closed fringes are an obvious violation of this condition.

In the two-dimensional algorithm, the formula for $F(I)$ becomes

$$F(I) = A(\nu,\mu) + C(\nu,\mu) + C^*(\nu,\mu) \tag{39}$$

The filtering of the complex quantity C can now be done in two orthogonal directions (with different filters in the ν and μ directions). Ambiguities can be resolved when making two separated calculations of Φ with band-pass filters oriented in orthogonal directions.

Another procedure is based on the evaluation of two interference patterns where the second pattern is shifted by about $\alpha = \pi/2$ compared to first pattern. The respective Fourier coefficients $c_1(x,y)$ and $c_2(x,y)$ can be used to calculate

$$\alpha(x,y) = \arctan\frac{\operatorname{Re} c_1 \cdot \operatorname{Im} c_2 - \operatorname{Im} c_1 \cdot \operatorname{Re} c_2}{\operatorname{Re} c_1 \cdot \operatorname{Re} c_2 + \operatorname{Im} c_1 \cdot \operatorname{Im} c_2} \tag{40}$$

α is expected to be nearly constant apart from the sign indication, which can be used to determine the sign of $\Phi(x,y)$ unambiguously.

6.5.2 Spatially synchronous fringe analysis

At the locations in the interference pattern where closed fringes appear, the local slope of the wavefront is steeper than the mean relative tilt between test and reference wavefront. At such locations, the Fourier coefficients $C(\nu, \mu)$ and $C^*(\nu, \mu)$ of the equation for $F(I)$ partially overlap and the sign is undetermined (change of sign of the calculated phase). To remove this difficulty, the relative tilt of the reference wavefront may be increased, which corresponds to a higher spatial frequency v_0.

The underlying information about the local phase Φ (x,y) of the wavefront under test can also be extracted by multiplying the interference pattern of Formula (8) directly with cosine and sine functions that have a linear phase dependence with the same mean spatial carrier frequency as the interference pattern (as described in References [15] and [16]). Since the interference pattern is usually scanned, a

one-dimensional representation can be used. Let f_0 be a rough estimate of the mean spatial frequency $\bar{f}$ of the interference pattern. Then

$$I(x,y)\cdot\cos(2\pi f_0 x)=a(x,y)\cos(2\pi f_0 x)+\frac{b}{2}\cos\left[\Phi(x,y)+2\pi f_0 x\right]+\frac{b}{2}\cos\left[\Phi(x,y)-2\pi f_0 x\right] \tag{41}$$

and

$$I(x,y)\cdot\sin(2\pi f_0 x)=a(x,y)\sin(2\pi f_0 x)+\frac{b}{2}\sin\left[\Phi(x,y)+2\pi f_0 x\right]+\frac{b}{2}\sin\left[\Phi(x,y)-2\pi f_0 x\right] \tag{42}$$

If the difference between the test wavefront represented by $\Phi(x,y)$ and the "reference wavefront" $2\pi f_0 x$ is small, the third term of these expressions represents a low spatial frequency component which can be extracted by a low-pass filter with a window function $H(x,y)$ (preferably of the Hanning type).

The resulting low spatial frequency Moiré can be written:

$$M_1=\frac{1}{2}I_0 V\cos\left[\Phi(x,y)-2\pi\left(\bar{f}-f_0\right)x\right] \tag{43}$$

$$M_2=\frac{1}{2}I_0 V\sin\left[\Phi(x,y)-2\pi\left(\bar{f}-f_0\right)x\right] \tag{44}$$

and

$$\Phi(x,y)-2\pi f_0 x=\arctan\frac{M_2(x,y)}{M_1(x,y)} \tag{45}$$

Since the convolution window should be several fringes wide, a high fringe density is needed in order for each phase measurement to be located to a region much smaller than the overall size of the pupil. This also keeps the loss of phase information small where the convolution window overlaps the edge of the pupil.

The simplest implementation of the quadrate multiplicative Moiré (QMM) involves a pure one-dimensional reference containing pure tilt, and is applicable to interferograms to which a large tilt component has been added, eliminating closed fringes. In this case, M_1 and M_2 reduce to:

$$\overline{M_1}=\sum_{i=-N}^{N} I(x_i,y)\cos(2\pi f_0 x_i)H(x-x_i,y) \tag{46}$$

$$\overline{M_2} = \sum_{i=-N}^{N} I(x_i, y)\sin(2\pi f_0 x_i) H(x - x_i, y) \tag{47}$$

and

$$\Phi(x, y) = \arctan \frac{\overline{M_2}}{M_1}$$

In another version, the high frequency component of the equation is extracted by filtering with a filter frequency f_0. Then:

$$F_1(x, y) = \sum_{i=-N}^{N} I(x_i, y)\cos\left[2\pi f_0 (x - x_i)\right] H(x - x_i, y) \tag{48}$$

and

$$F_2(x, y) = \sum_{i=-N}^{N} I(x_i, y)\sin\left[2\pi f_0 (x - x_i)\right] H(x - x_i, y) \tag{49}$$

is calculated, resulting in the phase

$$\Phi(x, y) = \arctan \frac{F_2(x, y)}{F_1(x, y)} \tag{50}$$

This method is called "sinusoidal window algorithm".

A third method makes use of a complex exponential window to filter out the phase information of a single side lobe of the Fourier transform

$$h' = h(x, y) e^{j2\pi f_0 x} \tag{51}$$

resulting in a function which real and imaginary parts are identical to F_1 and F_2 of the sinusoidal window method. The phase is then given by the arctangent of the ratio of real and imaginary parts.

6.6 Removal of phase ambiguities (phase unwrapping)

The above formulae allow the determination of the phase modulo π. In order to determine the phase difference between the interfering wavefronts without ambiguity, perform the following additional steps (see Reference [9]).

a) In a first step, the determination of phase modulo 2π requires to take the signs of the numerator ($\sin\Phi$) and the denominator ($\cos\Phi$) in the arctangent expression of the phase into account. The phase is calculated modulo $\pi/2$ using absolute values in the nominator and denominator, then the signs of these quantities are used according to Table 7 to calculate the phase modulo 2π.

   For the Carré technique, two other quantities instead of ($\sin\Phi$) and ($\cos\Phi$) have to be examined in the same way:

$$(I_1 - I_3) = \left[2 I_0 V \sin\varphi\right]\sin\Phi \tag{52}$$

   and

$$(I_2 + I_3) = (I_1 + I_4) = \left[2 I_0 V \cos\varphi \sin^2\varphi\right]\cos\Phi \tag{53}$$

b) Provided that the sampling requirement is fulfilled and the phase difference between adjacent pixels exceeds π a phase jump will occur and a phase of 2π should be added with the sign of the phase of the previous pixel. This procedure should be carried out in two dimensions, first for one diameter and then in the perpendicular direction from this diameter proceeding over the rest of the aperture.

In this way, a continuous phase function can be obtained from the raw phase data.

Table 7 — Determination of the phase modulo 2π

Numerator ($\sin \Phi$)	Denominator ($\cos \Phi$)	Adjusted phase	Range of phase values
positive	positive	Φ	$0-\frac{\pi}{2}$
positive	negative	$\pi - \Phi$	$\frac{\pi}{2}-\pi$
negative	negative	$\pi + \Phi$	$\pi-\frac{3\pi}{2}$
negative	positive	$2\pi - \Phi$	$\frac{3\pi}{2}-2\pi$
0	anything	π	π
negative	0	$\frac{\pi}{2}$	$\frac{\pi}{2}$
positive	0	$\frac{3}{2}\pi$	$\frac{3}{2}\pi$

6.7 Registration of wavefronts; coordinate systems, coordinate system definition

In direct phase measuring interferometers, wavefront phase or surface figure data are recorded on a detector. At each x_p, y_p pixel location on the detector where there is phase modulated intensity data, a phase or height value is recorded. A right-handed coordinate system is preferred to represent this data: positive x to the right, positive y to the top, positive z coordinate corresponding to a "taller" surface height (shorter OPD/cavity for a Fizeau). Some measured objects will fill the entire detector array, while others will fill only part of it. When the entire array contains data, the centre of the array is the centre or origin of the processed data set. The ratio of the number of pixels vertically to horizontally needed to describe a circle is the aspect ratio.

When only part of the array is filled with phase data, the origin of the processed data is the centre of gravity (or centroid) of the data set. For data from a circular object, the coordinate system of the object being measured will be shown how it is related to the pixel coordinates. Some detectors have pixels that are not square but rectangular. However, it is desirable that a circular object be displayed and analysed as circular. Both the centring and the aspect ratio of the pixels are accounted for by the following method. Using all the modulated phase data, the mean x_p and y_p are found by summing the x and y locations of every modulated data pixel. This mean, x_0, y_0, of the $x_p$and y_p, gives the centre of the data set in pixel coordinates. The coordinates (x_{pc}, y_{pc}) of the centred pixel data set is then

$$x_{pc} = x_p - x_0 \text{ and } y_{pc} = y_p - y_0 \tag{54}$$

The centred pixel data set are then used to determine the statistics of the data by using various processing algorithms described below.

Another issue regarding coordinate systems is the mapping of a curved surface onto the plane detector. For a plane surface the lateral (x,y) coordinates of the object are mapped linearly onto the detector. For a curved surface, however, particularly a surface with a high numerical aperture, there is a question of how the surface is mapped onto the detector. For most interferometers, the object under test is mapped onto that object's pupil plane, a plane at the object's vertex. The mapping is given by rays that appear to come from infinity so they are parallel to the optical axis. Thus the x,y coordinates of the surface under test are mapped one to one to the pupil plane while the component of sag in the direction of the optical axis is suppressed. The pupil plane is mapped linearly onto the detector just as in the case of a plane surface.

There are two consequences that follow immediately from this.

— A plane grid of straight lines perpendicular to the optical axis placed between the object and the interferometer will be imaged on the detector with pincushion distortion.

— Also, the x,y coordinates in the pupil plane are the same x,y coordinates that appear in all the analysis of the phase data

6.8 Polynomial and other representations of wavefronts

6.8.1 Representation of phase data

Once the phase data is transformed or mapped from detector or pixel coordinates into analysis coordinates, it can be represented in basically two different ways. It can be left as discrete data points or it can be interpolated with an analytic function into a global representation expressed by a number of coefficients. Each method has advantages and disadvantages depending on how the data are to be used. Neither method is optimum in all cases. In some methods of analysis, it is best to represent the low spatial frequency part of the data by analytic functions and the high spatial frequency part by the raw pixelated data.

A common use of analytic functions in interferometry is alignment error removal. The interference fringes are typically "nulled" by adjusting the test part position (via rigid body motions: translations and tilts/rotations). The nulling will not be perfect, so any part of the measured wavefront attributable to misalignment should be removed mathematically. Tilt alignment error is defined in ISO 14999-4:2015, 3.2.2 (and the wavefront with the tilt removed is defined in ISO 14999-4:2015, 3.2.4). This tilt function could be represented with the first 3 terms of the polynomial series given in Annex A (Z, L, B functions); e.g. Z(0, 0), Z(1, 1), and Z(1, −1) for Zernike polynomials. Twist alignment error for cylindrical wavefronts is defined in ISO 14999-4:2015, 3.2.3, and could be represented with Z(2, −2), L(1, 1), or B5.

Power deviation for spherical wavefronts (from ISO 10110-5 indication 3/A which is defined by ISO 14999-4:2015, 3.2.6 and 3.3.1) is often treated in a similar way. A parabolic approximation is effective in cases where the spherical wavefront numerical aperture (NA) is low (small ratio of part diameter to radius), e.g. Z(2, 0) for Zernike polynomials. This approximation, however, can introduce significant amounts of spherical aberration for higher NA wavefronts; in such cases the function $\left(1-\sqrt{1-r^2NA^2}\right)$ is a better approximation. But unfortunately this function cannot be represented by a single term from the annex polynomials. Further note that this fitting function depends on the part prescription (from the NA), while the low NA approximation (and the tilt and twist terms given above) do not.

NOTE 1 The more accurate power fitting function depends on the part prescription (from the NA), while the low NA approximation (and the tilt and twist terms given above) do not.

NOTE 2 The power approximation is similar to that given in ISO 10110-5:2015, Annex A.

Leaving the phase data as discrete data points, or pixelated data, has the advantage that it is essentially raw data. It has the maximum intrinsic information about the optic under test. As with any sampled data set, there is no, or too little information about the object on a spatial scale of less than two pixels, the Nyquist limit. But there is detailed information about the surface up to the Nyquist limit.

The disadvantage of leaving the data pixelated is that many useful aspects of the data are not obvious from this large number of essentially unrelated data points. Also to derive most statistical parameters from the data, the entire data set should be processed to determine each different parameter.

If the pixelated data is fitted to 2-D analytic functions, many properties of the data set are obvious from inspection or can be calculated simply from the coefficients representing the function. The entire data set is no longer needed to get most of the surface information out of it. Also, if certain analytic functions are used, many properties of how well an imaging system will work can be determined directly from the functional coefficients.

The disadvantage of fitting to the raw data is that the fitting process acts as a filter, particularly of high spatial frequency components, i.e., it is not possible to recover the raw data from the analytic function.

Analytic functions for describing surfaces with either circular or rectangular boundaries are defined in 6.8.2 to 6.8.6.

6.8.2 Zernike polynomials for a circular boundary

Because most optical systems have circular pupils, a particularly useful set of analytic functions for representing optical surface and wavefront data are the Zernike polynomials. This particular set of polynomials is complete and orthogonal over the unit circle. In addition, with appropriate normalization, the coefficients of the Zernike polynomials are the rms values of the individual optical aberrations present in the phase data and these directly relate to the imaging properties of optical systems.

It should be pointed out immediately that the useful properties of Zernike polynomials apply only to circular apertures (or their close relatives) and that very misleading results will be obtained by applying Zernike fitting to non-circular data sets. For square or rectangular data sets, see 6.8.5, A.3.2, and A.3.3. For situations where the clear aperture is elliptical, it is usually this way because a circular beam is incident on a mirror at a 45° (or some other large) angle of incidence. In this case, the correct amount of astigmatism over the elliptical aperture is the equivalent of power at normal incidence. For this reason, in 6.8.3, it is suggested to rescale the aperture coordinates to make the pupil circular, to evaluate the error with the regular Zernike polynomials and then to scale the aperture back to its elliptical shape. The notation used is that of Reference [17] with indices given by (n, m), where “n” is the radial order and “m” and the angular order of the functions. See A.3.1 for more detail.

A useful initial operation to perform with the Zernike representation is to remove the piston, tip and tilt present in almost every set of interferometric data. Similarly, higher order coefficients can be set to zero to see the effect of removing higher order aberrations. The Zernike normalization and nomenclature related to preparation of drawings is defined in ISO 14999-4. See A.2 for further detail about normalization.

6.8.3 Use of Zernike polynomials for an elliptical boundary

Elliptical apertures generally occur in optical systems because a collimated beam of light is reflected from a plane mirror to fold the beam path. As in many optical systems, a small amount of power in the optical path is of little consequence because it can usually be focused out in the image plane. However, if a plane mirror has a little power and is used in non-normal incidence the mirror will introduce astigmatism, which will degrade the optical performance of the system.

On the other hand, if the plane mirror has astigmatism aligned with the plane of incidence then this astigmatism will appear as power in the system and may be focused out. Thus, the main use for Zernike polynomials in the analysis of elliptical apertures is to determine the magnitude and orientation of power and astigmatism in the mirror surface.

The following procedure will permit the analysis of phase data over an elliptical aperture. First, assume the major axis, a, of the aperture is aligned along the x-axis. Let a and b be the lengths of the semi-major and minor axes of the ellipse. Multiply the x coordinates of all the data locations by the ratio (b/a). The pupil will now be circular. Fit Zernike polynomials as usual.

If the optical system will allow power to be focused out then the Z(2,0) coefficient times 2 times the cosine of the angle of incidence at which the mirror will be used is a measure of the peak-to-valley power introduced by the mirror. Any astigmatism or other aberration terms will degrade the wavefront and should be considered as irregularity. After power (Z(2,0)) is removed, the peak-to-valley magnitude over the remaining irregularity needs to be multiplied by the cosine of the angle of incidence in the system to determine its effect on the wavefront. The comments above assume the mirror surface is tested at normal incidence. If it is tested at the angle in which it will be used the test will give a direct measure of the irregularity.

Once Zernike polynomials have been fitted to the circle aperture data, a contour map of the surface over the elliptical aperture can be made by multiplying the x coordinates of the circular data by the ratio (a/b) and plotting the results.

There may be other occasions for fitting Zernike polynomials to data from an elliptical clear aperture. The approach suggested above may also be used in this case provided there is a clear understanding of how scaling the aperture coordinates in one direction affects the results of the fitting.

6.8.4 Zernike-Tatian polynomials for a circular shape with a central hole (annular)

In the case of an interferogram of an annular aperture, such as might be obtained when testing a Cassegrain type telescope, the annular aperture can be fitted with modified Zernike polynomials described by Mahajan (References [18] and [19]). These polynomials are a function of ε, the ratio of the inner radius of the annulus to the unit outer radius of the annulus. In all other attributes, they are the same as the Zernike polynomials for the full circular aperture.

6.8.5 Legendre polynomials for a rectangle boundary

There are times when it is convenient to represent a rectangular aperture with a 2-D analytic function, such as when testing cylindrical optics. When this is necessary, Legendre polynomials are an appropriate choice, and are detailed in A.3.2. In some cases polynomials with a rotationally symmetric character are desired over a rectangular aperture (for example, when a power-like ($x^2 + y^2$) term is desired). In such cases, polynomials based on Zernikes can be employed, see A.3.3 for details.

6.8.6 Orthogonal functions on "unusual areas"

It is possible to compute orthogonal polynomials, given certain desired characteristics, over an arbitrarily shaped area, including a weight function, if desired. Mathematically, this is not a new method but it does not seem to be applied in the optics world. The term orthogonal *function*, and not *polynomial*, is used in this text as this method applies to absolutely any base of functions. For example, the functional basis could include $\left(1-\sqrt{1-r^2NA^2}\right)$, where r is the normalized coordinate (1 corresponding to the numerical aperture NA) to accurately model the wavefront from a longitudinal shift or radius difference in a spherical surface test in a Fizeau interferometer (i.e. defocus).

Shown in Annex A are a few examples of orthogonal polynomials, without showing the details of the computation. This computation can be very simple using matrix notation, enabling the generation of the polynomials one after the other, in a deterministic way (no iterations or approximations).

7 Test reports and calibration certificates

7.1 General

The results of each test, calibration, or series of tests or calibrations carried out by a laboratory should be reported accurately, clearly and objectively, and in accordance with any specific instructions in the test or calibration procedures in accordance with policies outlined in ISO/IEC 17025.

The results should be reported, usually in a test report or a calibration certificate, and should include all the information requested by the client and necessary for the interpretation of the test or calibration results. This information typically provided is described by 7.2, and 7.3 or 7.4

7.2 Content of test reports and calibration certificates

Each test report or calibration certificate should include at least the following information, unless the laboratory has valid reasons for not doing so:

a) a title (e.g. “Test report” or “Calibration certificate”);

b) the name and address of the laboratory, and the location where the test or calibration was carried out if different from the address of the laboratory;

c) unique identification of the test report or calibration certificate, such as a serial number, on each page;

d) the name and address of the client;

e) a brief description of the method of measurement used;

f) a brief description, including unambiguous identification, of the item(s) tested or calibrated;

g) the date(s) of performance of the test or calibration;

h) the test or calibration results with the units of measurement;

i) the name, function and signature of the person authorizing the test report or calibration certificate.

Hard copies of test reports and calibration certificates should also include the page number and the total number of pages, e.g. Page 1 of 3.

It is recommended that laboratories should also include a statement specifying that the test report or calibration certificate should not be reproduced except in full.

7.3 Test reports

In addition to the requirements listed in 7.2, test reports should, where necessary for the interpretation of the test results, include the following:

a) deviations from, additions to, or exclusions from the test method and information on specific test conditions, such as environmental conditions;

b) where relevant, a statement of compliance/non-compliance of the test item(s) with requirements or specifications;

c) where applicable, a statement on the estimated uncertainty of measurement (information on uncertainty of measurement is needed in test reports when it is relevant to the validity or application of the test results, when a client’s instructions so require, or when the uncertainty affects compliance to a specification limit);

d) where appropriate and needed, opinions and interpretations based on the measurement results;

e) additional information which may be required by specific methods or clients.

7.4 Calibration certificates

7.4.1 Basics

In addition to the requirements listed in 7.2, calibration certificates should include the following, where necessary for the interpretation of calibration results:

a) the conditions (e.g. environmental) under which the calibrations were made that have an influence on the measurement results;

b) the uncertainty of measurement and/or a statement of compliance with an identified metrological specification or clauses thereof;

c) evidence that the measurements are traceable to national standards or physical constants (such as a particular wavelength of light; for example BIPM has a statement on the Helium Neon laser 633 nm wavelength for traceability to the metre).

7.4.2 Specification

The calibration certificate should relate only to quantities and the results of functional tests. If a statement relating to compliance with a specification is made, this should identify which clauses of the specification are met or not met.

When a statement of compliance with a specification is made omitting the measurement results and associated uncertainties, the laboratory should record those results and retain them for possible future reference.

When statements of compliance are made, the uncertainty of measurement needs to be taken into account.

7.4.3 Adjustment or repair

When an instrument/device for calibration has been adjusted or repaired, the calibration results before and after adjustment or repair, if available, should be reported.

7.5 Opinions and interpretations

When opinions and interpretations are included, the laboratory should document the basis upon which the opinions and interpretations have been made. Opinions and interpretations should be clearly marked as such in a test report.

Opinions and interpretations included in a test report may comprise, but not be limited to, the following:

— an opinion on the statement of compliance/non-compliance of the results with requirements;

— recommendations on how to use the results;

— guidance to be used for improvements.

In cases where it is appropriate to communicate opinions and interpretations by direct dialogue with the client, that dialogue should also be written down.

7.6 Electronic transmission of results

In the case of transmission of test or calibration results by telephone, telex, facsimile or other electronic means, the requirements of this document should be met.

7.7 Format of reports and certificates

The format should be designed to accommodate each type of test or calibration carried out and to minimize the possibility of misunderstanding or misuse.

Attention should be given to the layout of the test report and calibration certificate, especially with regard to the presentation of the test and calibration data and ease of assimilation by the reader.

The headings used within the report or certificate should be standardized as far as possible.

7.8 Amendments to test reports and calibration certificates

Material amendments to a test report or calibration certificate after issue should be made only in the form of a further document, or data transfer, which includes a statement indicating that it is a "Supplement to test report, serial number (XXXX)" or the equivalent for a calibration certificate.

When it is necessary to issue a completely new test report or calibration certificate, this should be uniquely identified and should contain a reference to the original that it replaces.

8 Data format

Information concerning data formats used in interferometric computational processes is not standardized, and so the software documentation of the interferometer manufacturer will need to be referred to. The format will usually be binary, though the interferometer software may also support export to a human-readable (text) format as well. The data file will typically include some metadata in a header (e.g. instrument serial number, date of measurement, wavelength of light used), as well as a matrix of values corresponding to the measured wavefront on a uniformly sampled grid (in X and Y).

Annex A
(informative)

Orthogonal polynomials

A.1 General

Various polynomial sets are described here, to be used on the two most widely used area shapes in optics:

a) Circular area (extendible to the elliptical area);

b) Square area (extendible to the rectangular area).

Each set is designated by a combination of letters:

— Area shape: C (circle/ellipse); S (square/rectangle);

— Generation basis: P (polar); R (rectangular);

— Form of expression: Z (Zernike), P (polar); R (rectangular).

NOTE 1 In this annex, the terms "polynomial" and "function" are sometimes used interchangeably. For example, a Zernike *function* may be obtained by multiplying a Zernike *polynomial* by a *harmonic* function [namely, $\cos(\phi)$, $\sin(\phi)$]. However, expressed in Cartesian coordinates, Zernike functions become *polynomials* with respect to (x,y).

NOTE 2 The orthogonal sets described here are orthogonal on continuous areas of the appropriate shape (circle or square). Using these sets for discrete, which is always the case with real data, results in non-orthogonality of the said functions. It is up to the user to determine the degree to which this non-orthogonality perturbs the coefficient computation. However, it can be said that for metrology devices common today, the discrepancy is negligible for the lower order functions.

CAUTION — Any orthogonal set is to be used with caution. When applied to *real* data (as opposed to, e.g. studying the *mathematics* of aberrations), the orthogonal functions are convenient for representing arbitrary topographies over the defined aperture shape, but individual terms do not generally have any particular physical meaning beyond that.

The reason why the orthogonal sets described here are useful for *practical* analysis is the form of first few terms, namely "piston", "tip", "tilt", "power", which describe changes in conjugates, and also some useful lower order aberrations: "coma", "astigmatism", "spherical aberration".

The universally used Zernike set (A.3.1), as well as the alternative set (A.3.3), include these useful terms. The Legendre set (A.3.2) does not include all these terms (in particular, it has no "power" term).

However, above these lower orders, the terms of any orthogonal set do not typically have any physical meaning in *practice* (beyond a convenient representation of an arbitrary surface). For example, optical components rarely exhibit a pure Z20, L17 or B29 shape.

Table A.1 — Orthogonal set summary

No.	Reference	Name	Area	Symmetry	Coordinates	Form (see notes)	Note
Table A.2	C-P-P	Circular/ Zernike	Circle (ellipse)	Polar	Polar	$\{P_{n,m}(r)\cos(m\theta)\}$ $\{P_{n,m}(r)\sin(-m\theta)\}$	Zernike { Z }
Table A.3	S-R-R	Square/ Legendre	Square (rectangle)	Rectangular	Rectangular	$\{L_p(x)\ L_q(y)\}$	Legendre { L }
Table A.4	S-P-Z	Square/ Polar	Square (rectangle)	Polar	Polar	Combination of Zernike functions	New set { B }
Table A.5	S-P-P	Square/ Polar	Square (rectangle)	Polar	Polar	$\{B_{n,m}(r, \theta)\}$	New set { B }

NOTE 1 $P_n(r)$ denotes polynomials for the variable "r".

NOTE 2 n, p, q are non-negative integers; m is positive for cosine terms, negative for sine terms. See "A.3.1.2 Definitions" below.

NOTE 3 More information may be obtained, e.g. the references and the worldwide web.

A.2 Coefficient computation

When fitting a given set of orthogonal functions $\{Q_0, Q_1, Q_2, ...\}$ to a function "f", coefficients $\{a_0, a_1, a_2, ...\}$ of the orthogonal functions are computed as follows:

$$a_i = \frac{\iint_A Q_i(u,v) f(u,v) \mathrm{d}A}{\iint_A Q_i(u,v) Q_i(u,v) \mathrm{d}A} = \frac{1}{A\, s_{rms,i}^2} \iint_A Q_i(u,v) f(u,v) \mathrm{d}A$$

where

(u,v) are the space coordinates;

A is the area of integration and $\mathrm{d}A$ is the area increment;

$s_{rms,i}^2$ is the mean-squared value of orthogonal polynomial (Q_i) (s_{rms} is the root mean square, or rms).

$s_{rms,i}^2$ values are tabulated in the tables below. However, because the orthogonality of the sets over discrete areas is not strict, it is advised to perform the actual computation of the s_{rms}^2 values when they are required, rather than using the values tabulated here.

NOTE 1 The default normalization of the orthogonal polynomials given here is unity at the edge (polynomial terms achieve a maximum value of one at the edge). In some cases, however, using an orthonormal basis (polynomial terms have unit rms) is convenient. For example, orthonormal polynomial coefficients can be root sum squared together to obtain a total rms of those polynomials.

NOTE 2 The edge-normalized polynomials can be converted to rms-normalized by dividing by s ($\sqrt{s_{\mathrm{rms}}^2}$, where s_{rms}^2 is found in the tables). For example, Z(2,0) goes from $(2r^2 - 1)$ edge normalized to $\sqrt{3}$ $(2r^2 - 1)$ rms normalized.

NOTE 3 Edge-normalized coefficients can be converted to rms-normalized by multiplying by s. For example, if a wavefront fit with edge normalized Zernikes obtained a coefficient of 10 nm for Z(2,0), then the rms-normalized Z(2,0) is $(10\ \mathrm{nm}) \times \sqrt{\frac{1}{3}}$ or approximately 5,77 nm.

A.3 Definitions, tables and illustration of orthogonal functions

NOTE $C_n^i = \mathrm{Comb}(n,i) = \dfrac{n!}{i!(n-i)!}$

A.3.1 Circular areas: Zernike functions ("C-P")

A.3.1.1 General

This is the classical Zernike set, which is widely used in optics. It is orthogonal over the circular area (extendible to the elliptical area).

Its definitions are given in A.3.1.2. s_{rms}^2 values are given in Table A.2 and an illustration of the orthogonal set up to N = 10 (Z0–Z35) is given in Figure A.1.

NOTE 1 The indexing of Zernike polynomials is standardized in ISO 14999-4, and uses a double index (n, m). The use of a single index (often termed "fringe" ordering), however, is commonly used in the optics industry (e.g. in lens design and metrology analysis software packages). This ordering is given in the "Z" column of Table A.2, while the equivalent standardized dual indices (n, m) are given in the "n, m" column.

NOTE 2 The fringe ordering index given in Table A.1 starts at 0; e.g. Z(0,0) = Z0. Another convention of the fringe ordering starts at 1; e.g. Z(0,0) = Z1. This notation is not standardized, so use caution when interpreting Zernikes that use a single index.

A.3.1.2 Definitions

Indices: $(n + |m|) = N$: Order of the function;

$(n - |m|) \geq 0$: $|m|$ is not greater than n;

$(n - |m|)$ is even.

NOTE 1 For the purpose of compact notation, the sign of "m" denotes whether that term uses sine or cosine (m=0 is rotationally symmetric, with neither a sine nor a cosine factor), rather than having a third index for sin/cos. This deviates from conventional mathematical practice, where m is nonnegative.

Area of definition: Circle, radius equal to 1 (area = π).

Zernike polynomial:

$$P(n,m,r)=\sum_{i=0}^{n-|m|}-1^{i}\left[\frac{\frac{(n-i)!}{i!}}{\left(\frac{n+|m|}{2}-i\right)!\left(\frac{n-|m|}{2}-i\right)!}\right]r^{n-2i}$$

Principal term:

$$\left[\frac{n!}{\left(\frac{n+|m|}{2}\right)!\left(\frac{n-|m|}{2}\right)!}\right]r^{n}=C_{n}^{\left(\frac{n+|m|}{2}\right)}\ r^{n}=C_{n}^{(N/2)}\ r^{n}$$

Zernike function (orthogonal series for use in optics):

$$Z(n,m,r,\theta)=P(n,m,r)\times\begin{vmatrix}\cos(m\theta);m>0\\ \sin(-m\theta);m<0\end{vmatrix}$$

Function mean-square value $\left(s_{\text{rms}}^{2}\right)$:

$$\left\|Z(n,m,r,\theta)\right\|^{2}/A=1/(n+1)\times\begin{vmatrix}\frac{1}{2};|m|>0\\ 1;m=0\end{vmatrix}$$

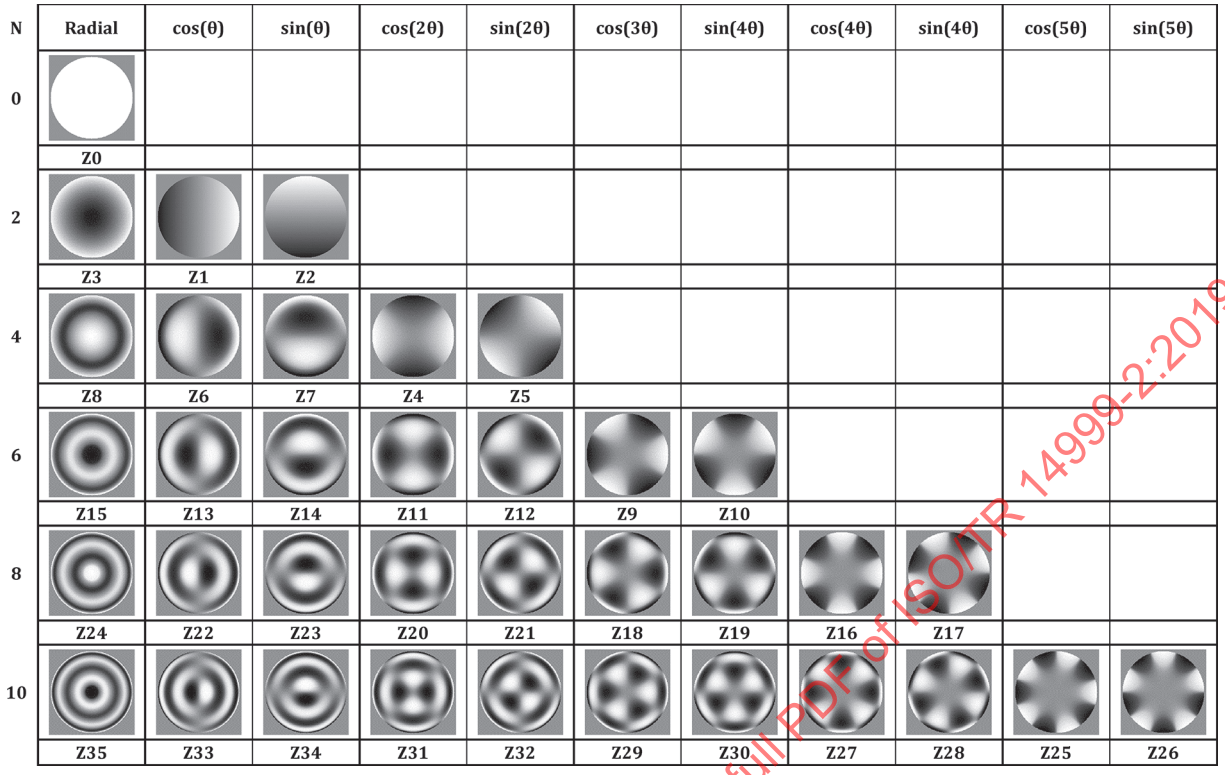
NOTE 2 The mean-square value of the Zernike *functions* is modified by the cosine or sine functions. For these functions ($|m| > 0$), the mean-square value is half that of the corresponding Zernike *Polynomial*. For the rotationally-symmetrical functions ($m = 0$), the mean-square values are equal to those of the corresponding polynomial.

Table A.2 — Circular — Zernike ("C-P-P")

Orthogonal on the circle (extendible to the ellipse) **Polar variables:** (r, θ) **Expressed in polar coordinates:** $\sum\{P_n(r)\cos(m\theta)\}$ for $m > 0$ **or** $\Sigma\{P_n(r)\sin(-m\theta)\}$ for $m < 0$; $N = n + \lvert m\rvert$; $n \geq \lvert m\rvert$				
Z	N	n, m	**Formula**	s_{rms}^{2}
Z0	0	0, 0	1	1
Z1	2	1, 1	$r\cos\theta$	1/4
Z2	:	1, –1	$r\sin\theta$	1/4
Z3	2	2, 0	$2r^2 - 1$	1/3
Z4	4	2, 2	$r^2\cos2\theta$	1/6
Z5	:	2, –2	$r^2\sin2\theta$	1/6
Z6	:	3, 1	$(3r^2 - 2)\, r\cos\theta$	1/8
Z7	:	3, –1	$(3r^2 - 2)\, r\sin\theta$	1/8
Z8	4	4, 0	$6r^4 - 6r^2 + 1$	1/5
Z9	6	3, 3	$r^3\cos3\theta$	1/8
Z10	:	3, –3	$r^3\sin3\theta$	1/8
Z11	:	4, 2	$(4r^2 - 3)\, r^2\cos2\theta$	1/10
Z12	:	4, -2	$(4r^2 - 3)\, r^2\sin2\theta$	1/10
Z13	:	5, 1	$(10r^4 - 12r^2 + 3)\, r\cos\theta$	1/12
Z14	:	5, –1	$(10r^4 - 12r^2 + 3)\, r\sin\theta$	1/12
Z15	6	6, 0	$20r^6 - 30r^4 + 12r^2\text{-}1$	1/7

Table A.2 *(continued)*

Orthogonal on the circle (extendible to the ellipse) **Polar variables:** (r, θ) **Expressed in polar coordinates:** $\sum \{P_n(r) \cos(m\theta)\}$ for $m > 0$ **or** $\Sigma \{P_n(r) \sin(-m\theta)\}$ for $m < 0$; $N = n + \lvert m \rvert$; $n \geq \lvert m \rvert$				
Z	N	n, m	**Formula**	s^2_{rms}
Z16	8	4, 4	$r^4\cos 4\theta$	1/10
Z17	:	4, −4	$r^4\sin 4\theta$	1/10
Z18	:	5, 3	$(5r^2 - 4)\, r^3\cos 3\theta$	1/12
Z19	:	5, −3	$(5r^2 - 4)\, r^3\sin 3\theta$	1/12
Z20	:	6, 2	$(15r^4 - 20r^2 + 6)\, r^2\cos 2\theta$	1/14
Z21	:	6, −2	$(15r^4 - 20r^2 + 6)\, r^2\sin 2\theta$	1/14
Z22	:	7, 1	$(35r^6 - 60r^4 + 30r^2 - 4)\, r\cos\theta$	1/16
Z23	:	7, −1	$(35r^6 - 60r^4 + 30r^2 - 4)\, r\sin\theta$	1/16
Z24	8	8, 0	$70r^8 - 140r^6 + 90r^4 - 20r^2 + 1$	1/9
Z25	10	5, 5	$r^5\cos 5\theta$	1/12
Z26	:	5, −5	$r^5\sin 5\theta$	1/12
Z27	:	6, 4	$(6r^2 - 5)\, r^4\cos 4\theta$	1/14
Z28	:	6, −4	$(6r^2 - 5)\, r^4\sin 4\theta$	1/14
Z29	:	7, 3	$(21r^4 - 30r^2 + 10)\, r^3\cos 3\theta$	1/16
Z30	:	7, −3	$(21r^4 - 30r^2 + 10)\, r^3\sin 3\theta$	1/16
Z31	:	8, 2	$(56r^6 - 105r^4 + 60r^2 - 10)\, r^2\cos 2\theta$	1/18
Z32	:	8, −2	$(56r^6 - 105r^4 + 60r^2 - 10)\, r^2\sin 2\theta$	1/18
Z33		9, 1	$(126r^8 - 280r^6 + 210r^4 - 60r^2 + 5)\, r\cos\theta$	1/20
Z34	:	9, −1	$(126r^8 - 280r^6 + 210r^4 - 60r^2 + 5)\, r\sin\theta$	1/20
Z35	10	10, 0	$252r^{10} - 630r^8 + 560r^6 - 210r^4 + 30r^2 - 1$	1/11
Z36	12	6, 6	$r^6\cos 6\theta$	1/14
Z37	:	6, −6	$r^6\sin 6\theta$	1/14
Z38	:	7, 5	$(7r^2 - 6)\, r^5\cos 5\theta$	1/16
Z39	:	7, −5	$(7r^2 - 6)\, r^5\sin 5\theta$	1/16
Z40	:	8, 4	$(28r^4 - 42r^2 + 15)\, r^4\cos 4\theta$	1/18
Z41	:	8, −4	$(28r^4 - 42r^2 + 15)\, r^4\sin 4\theta$	1/18
Z42	:	9, 3	$(84r^6 - 168r^4 + 105r^2 - 20)\, r^3\cos 3\theta$	1/20
Z43	:	9, −3	$(84r^6 - 168r^4 + 105r^2 - 20)\, r^3\sin 3\theta$	1/20
Z44	:	10, 2	$(210r^8 - 504r^6 + 420r^4 - 140r^2 + 15)\, r^2\cos 2\theta$	1/22
Z45	:	10, −2	$(210r^8 - 504r^6 + 420r^4 - 140r^2 + 15)\, r^2\sin 2\theta$	1/22
Z46	:	11, 1	$(462r^{10} - 1\,260r^8 + 1\,260r^6 - 560r^4 + 105r^2 - 6)\, r\cos\theta$	1/24
Z47	:	11, −1	$(462r^{10} - 1\,260r^8 + 1\,260r^6 - 560r^4 + 105r^2 - 6)\, r\sin\theta$	1/24
Z48	12	12, 0	$924r^{12} - 2\,772r^{10} + 3\,150r^8 - 1\,680r^6 + 420r^4 - 42r^2 + 1$	1/13



NOTE This figure depicts the same Zernike orthogonal polynomial set shown in ISO 14999-4:2015, Figure B.1, but with 2 key differences. First, the grey scale used in this figure has a higher contrast, to better highlight the structure in the polynomial terms. Second, the labels in the terms use the fringe ordering rather than the dual index standardized in ISO 14999-4 (to convert the fringe index to the corresponding dual index, see Table A.2, columns 1 and 3).

Figure A.1 — Example of visualisation of Zernike orthogonal set (Z0 – Z35)

A.3.2 Square areas — Legendre functions ("S-R")

A.3.2.1 General

Legendre polynomials are 1D polynomials. By multiplying two such polynomials, one for each direction (x and y), one obtains a 2D set of polynomials (functions) orthogonal on a square (extendible to the rectangular area).

The definitions are given in A.3.2.2. s^2_{rms} values are given in Table A.3 and an illustration of the orthogonal set (L0 – L20) is given in Figure A.2.

A.3.2.2 Definitions

Indices: $(p + q) = N$: Order of the function;

p, q: Non-negative integers.