TECHNICAL REPORT

ISO/TR 14999-1

First edition
2005-03-01

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

Part 1: Terms, definitions and fundamental relationships

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

Partie 1: Termes, définitions et relations fondamentales

Reference number
ISO/TR 14999-1:2005(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

In exceptional circumstances, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example), it may decide by a simple majority vote of its participating members to publish a Technical Report. A Technical Report is entirely informative in nature and does not have to be reviewed until the data it provides are considered to be no longer valid or useful.

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.

ISO/TR 14999-1 was prepared by Technical Committee ISO/TC 172, *Optics and photonics*, Subcommittee SC 1, *Fundamental standards*.

ISO 14999 consists of the following parts, under the general title *Optics and photonics — Interferometric measurement of optical elements and optical systems*:

— *Part 1: Terms, definitions and fundamental relationships* (Technical Report)

— *Part 2: Measurement and evaluation techniques* (Technical Report)

— *Part 3: Calibration and validation of interferometric test equipment* (Technical Report)

— *Part 4: Interpretation and evaluation of tolerances specified by ISO 10110*

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 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 is 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. which is based more in the field of physical optics than in the field of geometrical optics. As a consequence, only 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 there was discussion, 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 order to complement ISO 10110, the guideline adopted was to include what presently are common techniques used for the purpose of characterizing the quality of optical parts. Decision was made to complete a first Technical Report, and to then up-date it by supplementing new parts, as required. It is very likely that more material will be added in the near future as more stringent tolerances (two orders of magnitude) for optical parts and optical systems become mandatory when dealing with optics for the EUV range (wavelength range 6 nm to 13 nm) for microlithography. Also, testing optics with EUV radiation (the same wavelength as they are later used, e.g. at-wavelength testing) can be a new challenge, and is not covered by any current standards.

This Technical Report should cover 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 Technical Report 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, infrared radiation from CO_2-lasers in the range of 10,6 µm is used for testing rough surfaces after grinding or ultraviolet radiation from excimer-lasers in the range of 193 nm or 248 nm are used for at-wavelength testing of 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 1: Terms, definitions and fundamental relationships

1 Scope

This part of ISO/TR 14999 gives terms, definitions and fundamental physical and technical relationships for interferometric measurements of optical wavefronts and surface form of optical elements.

It explains why some principles of the construction and use of interferometers are important due to the wave nature of the wavefronts to be measured.

Since all wavefronts with the exception of very extended plane waves do alter their shape when propagating, this part of ISO/TR 14999 also includes some basic information about wave propagation.

In practice, interferometric measurements can be done and are done by use of various configurations; this part of ISO/TR 14999 outlines the basic configurations for two-beam interference.

The mathematical formulation of optical waves by the concept of the complex amplitude as well as the basic equations of two-beam interference are established to explain the principles of deriving the phase information out of the measured intensity distribution, either in time or in space.

Both random and systematic errors may affect the results of interferometric measurements and error types to be clearly differentiated are therefore described in this part of ISO/TR 14999.

2 Wave propagation and some topics on electromagnetic theory

2.1 Parameters, symbols, units and constants, operators and computational procedures

Basic parameters, symbols, units and constants are given in Table 1.

Operators and computational procedures are given in Table 2.

Table 1 — Parameters, symbols, units and constants

Parameters	Symbols	Recommended unit, constant
Electric field vector	E	V/m
Magnetic field vector	H	A/m
Electric displacement or electric flux density	D	$C/m^2 = As/m^2$
Magnetic induction or magnetic flux density	B	$T = Wb/m^2 = Vs/m^2$
Dielectric constant or permittivity [a]	ε	F/m = As/Vm
Dielectric constant in the vacuum	ε_0	$8{,}854 \times 10^{-12}$ F/m
Relative dielectric constant (relative permittivity)	ε_r	1
Magnetic permeability [b]	μ	H/m = Vs/Am
Magnetic permeability in the vacuum	μ_0	$1{,}257 \times 10^{-6}$ H/m
Relative magnetic permeability	μ_r	1
Velocity of the wave in the medium	c	m/s
Velocity of the wave in the vacuum	c_0	$2{,}997\ 924\ 58 \times 10^8$ m/s
Absolute refractive index	n	1

[a] Mathematical relationship: $\varepsilon = \varepsilon_0\, \varepsilon_r$.

[b] Mathematical relationship: $\mu = \mu_0\, \mu_r$.

Table 2 — Operators and computational procedures

Operator	Definition/Computational procedures	Name (type)
∇	$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$	Nabla (vector)
$\Delta\Psi$	$\equiv \left(\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} + \frac{\partial^2\Psi}{\partial z^2}\right)$ $\nabla \cdot \nabla\Psi = \nabla^2\Psi$	Laplacian (scalar)

2.2 Maxwell's equations

Maxwell's equations are the fundamentals for the electromagnetic wave propagation. Maxwell's equations for an electromagnetic wave propagating in a medium which does not involve any charge or current and has vanishing conductivity are expressed by:

$$\left.\begin{aligned} \nabla \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} &= 0 \\ \nabla \times \boldsymbol{H} - \frac{\partial \boldsymbol{D}}{\partial t} &= 0 \\ \nabla \cdot \boldsymbol{D} &= 0 \\ \nabla \cdot \boldsymbol{B} &= 0 \end{aligned}\right\} \qquad (1)$$

The mathematical relation between $\boldsymbol{D}$ and $\boldsymbol{E}$ as well as between $\boldsymbol{B}$ and $\boldsymbol{H}$ is given by:

$$\left.\begin{aligned}\boldsymbol{D} &= \varepsilon \boldsymbol{E} \\ \boldsymbol{B} &= \mu \boldsymbol{H}\end{aligned}\right\} \tag{2}$$

in a linear medium.

2.3 Electromagnetic fields in a medium

For media in which the dielectric constant ε and magnetic permeability μ are uniform, Equation (1) gives the following wave equations:

$$\left.\begin{aligned}\nabla^2 \boldsymbol{E} - \varepsilon\mu \frac{\partial^2 \boldsymbol{E}}{\partial t^2} &= 0 \\ \nabla^2 \boldsymbol{H} - \varepsilon\mu \frac{\partial^2 \boldsymbol{H}}{\partial t^2} &= 0\end{aligned}\right\} \tag{3}$$

2.4 Velocity of the wave

The velocity in an optically homogeneous and isotropic medium is given by:

$$c = \frac{c_0}{\sqrt{\varepsilon\mu}} \tag{4}$$

Analogously, in a vacuum the velocity is given by:

$$c_0 = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \tag{5}$$

2.5 Refractive index

The ratio of the propagation velocities in vacuum and in the medium with ε and μ

$$n = \frac{c_0}{c} \tag{6}$$

is called the refractive index of the medium or the absolute refractive index.

2.6 Scalar wave equation

As mentioned before, $\boldsymbol{E}$ and $\boldsymbol{H}$ are vectors. In many applications one deals with linearly polarized light, which can be fully described by one vector component. Equation (3) then reduces to the scalar wave equation. In the general form, the scalar wave equation may be written as:

$$\nabla^2 \psi - \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} = 0 \tag{7}$$

Equation (7) is in conformity with a second-order differential equation. ψ is called the light disturbance.

The basic problem of light propagation is thus simply the determination of the manner in which a wave propagates from one surface to another.

2.7 Amplitude, angular frequency, wavelength, wave number

We suppose a sinusoidal plane electromagnetic wave propagating in the z direction. The light disturbance ψ is determined as a function of position z and time t

$$\psi(z,t) = U \cos\left[\omega\left(t - \frac{z}{c}\right) + \delta\right] \tag{8}$$

where

U is the amplitude;

ω is the angular frequency;

δ is the phase constant of the wave.

The angular frequency ω is defined as $2\pi\nu$, where ν is the frequency, i.e. the number of waves per unit time. The wavelength λ is given from Equation (9):

$$\lambda = \frac{2\pi v}{\omega} = \frac{2\pi c_0}{n\,\omega} \tag{9}$$

The wave number k is defined as

$$k = \frac{2\pi}{\lambda} \tag{10}$$

In many applications of the concept of waves, as in diffraction or in interferometry, Equation (8) is used to define "wavefronts". In this case for a given position z the phase constant δ is a function of the lateral spatial coordinates, e.g. $\delta = \delta(x,y)$. This concept is useful, if $\delta = \delta(x,y)$ is measured, as in the case of interferometric measurements. Here $\delta = \delta(x,y)$ is the phase-difference of two interfering waves and x and y are the coordinates of the detector.

Sometimes it is more convenient to look for a "surface" in space $z = z(x,y)$, where the value for the phase $\delta = \delta[x,y,z(x,y)]$ remains constant. Such a surface defines the "shape" of a wavefront, which can be spherical, plane or aspherical, only to mention the most simple cases. This concept is used in 2.11 to 2.13 for discussing the question of the propagation of waves. Here it is shown, that the "shape" of the wavefront changes with z, with the only exception of a(n) (infinite) plane wave with constant amplitude.

In a more general case, also the amplitude U of the light stimulation might be a function of the lateral coordinates x,y, e.g. $U = U(x,y)$. If U is not constant with x,y, it is referred to as an "inhomogeneous wave". In practice, the variation of U is of minor importance, where the variation of δ is the quantity to be measured.

2.8 Complex notation, complex amplitude

The expression in Equation (8) can be written in complex form as

$$\psi(z,t) = \mathrm{Re}\left[u(z)\,\mathrm{e}^{\mathrm{i}(\omega t)}\right] \tag{11}$$

and

$$\begin{aligned} u(z) &= U(z)\,\mathrm{e}^{\mathrm{i}\,\phi} \\ \phi &= -\left(\omega\frac{z}{c}\right) + \delta \end{aligned} \tag{12}$$

where

$u(z)$ is called the complex amplitude of the wave;

U is the spatial dependent modulus (e.g. amplitude);

ϕ is the spatial dependent phase.

This complex notation is much more convenient than the notation in Equation (8) with real quantities. Nevertheless, only the real parts of Equations (11) and (12) have a physical meaning.

2.9 Irradiance

The irradiance I is given by the relationship

$$I = E/A \times \Delta t$$

where

I is the irradiance;

E is the energy;

A is the area;

Δt is the time interval.

A medium for direct recording of the field amplitude does not exist, since the frequency of the light stimulation is too high to be resolved. So the most common detectors register the irradiance which is proportional to the field amplitude absolutely squared:

$$I \propto |u|^2 = U^2 \tag{13}$$

The correct relation between the complex amplitude and the irradiance is given by:

$$I = \left(\frac{\varepsilon c_0}{2}\right) u \cdot u^* \tag{14}$$

where u^* is the conjugate complex amplitude.

2.10 Poynting vector

The vector $\boldsymbol{S}$ is known as the Poynting vector. $\boldsymbol{S}$ represents the amount of energy which crosses a unit area, normal to the directions of $\boldsymbol{E}$ and $\boldsymbol{H}$, per second

$$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H} \tag{15}$$

$\boldsymbol{S}$ can be interpreted as the density of the energy flow. The magnitude of the Poynting vector is a measure of the light intensity, and its direction represents the direction of propagation of the light.

The irradiance, I, is given by Equation (16) as follows:

$$I = \overline{S} \cdot n \tag{16}$$

where

n is the surface normal of the detector;

$\overline{S}$ is the time-averaged magnitude of the Poynting vector $\boldsymbol{S}$.

2.11 Propagation of plane waves

Waves, which have a constant phase at a fixed time t over each of the planes normal to the direction of propagation, are called plane waves (see Figure 1).

Let $\boldsymbol{r}\,(x,y,z)$ be a position vector of a point P in space and $\boldsymbol{n}\,(n_x,n_y,n_z)$ the normal unit vector of the wavefront in a fixed direction (see Figure 2). Wavefronts represent surfaces with constant phase. Any solution of Equation (7) of the form

$$\psi = \psi(\boldsymbol{r}\cdot\boldsymbol{n},t) \tag{17}$$

describes a plane wave since, at each instant of time, ψ is constant over each of the planes which are perpendicular to the normal unit vector $\boldsymbol{n}$:

$$\boldsymbol{r}\cdot\boldsymbol{n} = \kappa \tag{18}$$

where

κ is a constant.

The planes described by $\boldsymbol{r}\cdot\boldsymbol{n} = \kappa$ are defined as the wavefronts of the plane wave.

The disturbance $\psi\,(x,y,z,t)$ of a harmonic plane wave that propagates in the $\boldsymbol{n}$ direction is given by

$$\psi(x,y,z,t) = U\cos(2\pi\nu t - k\boldsymbol{n}\cdot\boldsymbol{r} + \delta) \tag{19}$$

The argument of the cosine function is termed the phase term; δ is the phase constant.

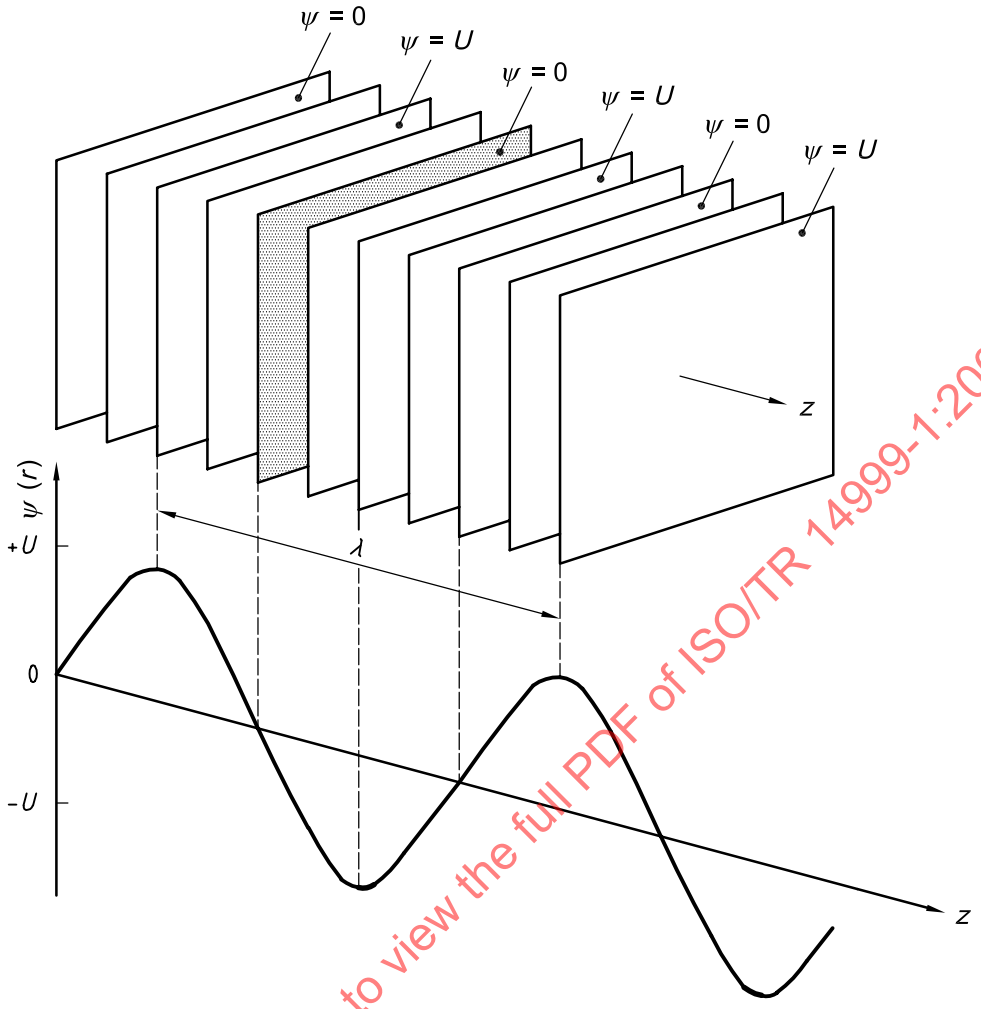


Figure 1 — Plane waves at the time $t = 0$

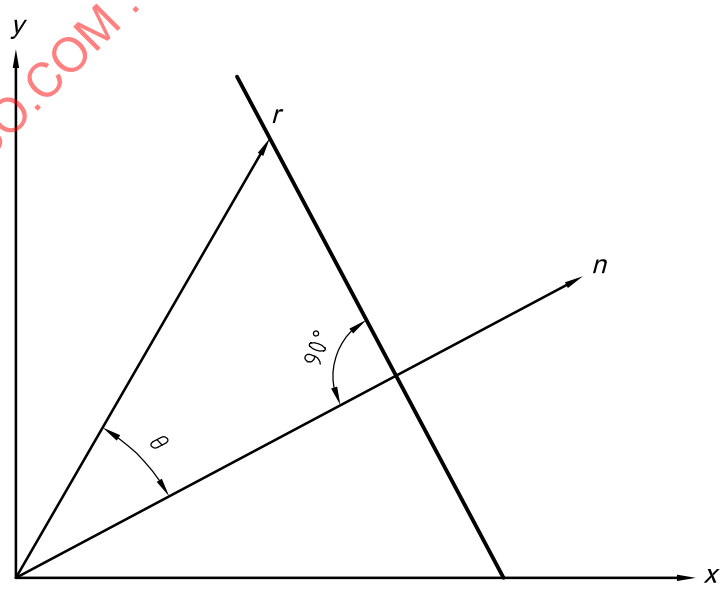


Figure 2 — Illustration of the condition $r \cdot n = \kappa$

2.12 Propagation of spherical wave

A spherical wave, illustrated in Figure 3, is emitted by a point source O.

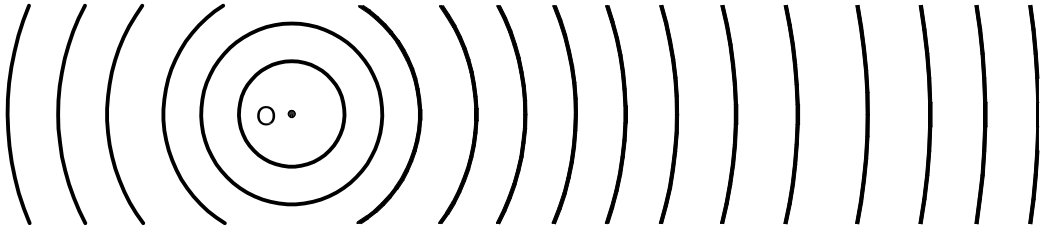


Figure 3 — Spherical waves

The complex amplitude representing a spherical wave should be of the form:

$$u = \frac{U}{r} e^{-ikr} \tag{20}$$

where r is the radial distance from the point source. The phase of this wave is constant for r equal to a constant, i.e. the phase fronts are spherically centred at the point source O. The r in the denominator of Equation (20) expresses the fact that the amplitude decreases as the inverse of the distance from the point source.

Consider Figure 4, where a point source is lying in the x_0,y_0-plane at a point of coordinates x_0,y_0. The field amplitude in a plane parallel to x_0,y_0-plane at a distance z will then be given by Equation (20) with

$$r = \sqrt{z^2 + (x - x_0)^2 + (y - y_0)^2} \tag{21}$$

where x, y are the coordinates of the illuminated plane. The approximation for the phase is carried out by a binomial expansion of the square root; when r is approximated by the first two terms of the expansion, the Fresnel approximation of the diffraction phenomena is obtained. In the amplitude factor [see Equation (21)], r may be replaced by z, because $(x - x_0)$, $(y - y_0) << z$. The complex amplitude of the field in the x,y-plane resulting from a point source at x_0, y_0 in the x_0,y_0-plane is then given by:

$$u(x, y, z) = \frac{U}{z} e^{-ikz} e^{-i(k/2z)[(x-x_0)^2 + (y-y_0)^2]} \tag{22}$$

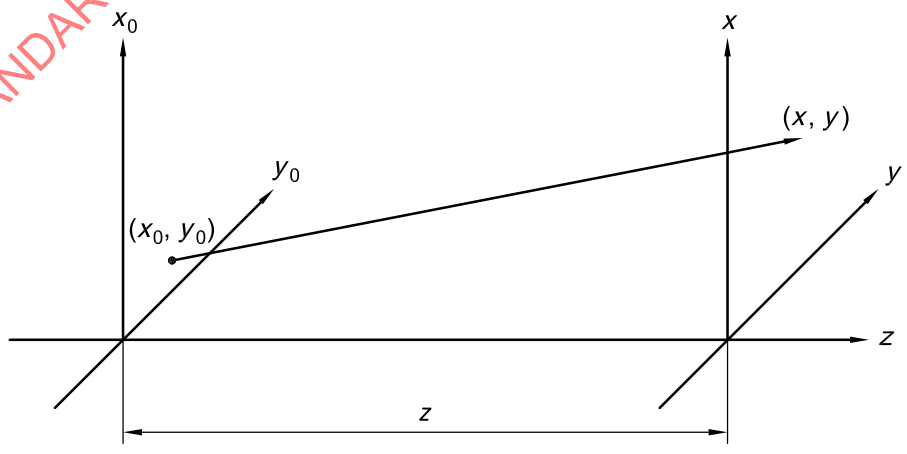


Figure 4 — A point source lying in x_0,y_0-plane

2.13 Propagation of waves with limited extent

The only type of wave that does not change its shape when propagating, is a plane wave with uniform amplitude. Such a wave is infinite in both lateral coordinates. A spherical wave, which is either converging to, or diverging from, its centre of curvature and which is defined on the complete solid angle of 4π is another example of a very special wave. In this case, the phase is altered, but any phase-distribution is similar to the adjacent one. Again, a uniform amplitude distribution should be demanded. These two special cases have been discussed in the previous subclauses.

All other wavefronts, especially plane or spherical wavefronts, which contain a boundary due to any limiting aperture, do alter their shape when travelling along their light path. One example is given in the next two figures, where the well-known Cornu's spiral is shown, which is a graphical representation of the solution of the Fresnel's integrals

$$\left.\begin{aligned} x &= \int_0^v \cos\frac{\pi v^2}{2}\,\mathrm{d}v \\ y &= \int_0^v \sin\frac{\pi v^2}{2}\,\mathrm{d}v \end{aligned}\right\} \tag{23}$$

x and y being the Cartesian coordinates of the plot and v being a parameter on the curve. Cornu's spiral can be used to visualize the complex amplitude of a plane wave diffracted at a knife-edge, as a most simple example.

Look at Figures 5 and 6 simultaneously. At the point of the geometrical shadow, point 0 on the abscissa in Figure 6 and the coordinate centre in Figure 5, the amplitude is dropped to 1/2 of the amplitude far away from the edge in the illuminated region. So, the intensity is dropped to 1/4. The complex amplitude for that point is visualized in Figure 5 by a straight line, joining points Z and 0. The point Z in Figure 5 remains the centre, where all vectors originate. Now going further into the illuminated region, the parameter v grows and with it the modulus of the complex amplitude, since the vectors grow larger. The extreme value is reached at b′, then dropping until a first relative minimum is reached at c′. At the same time the phase of the wavefront changes, since the direction of the vectors changes, when the point moves along the spiral. The gradients of the phase-change take large values in the vicinity of extreme points of the spiral, like points b′, c′, d′ and so on and are nearly zero when the line centred in Z is tangential to the Cornu's spiral.

Figure 5 shows the phase-change φ that the diffracted wave undergoes between points b′ and B′.

Since the phase is the quantity measured with interferometers, it is clear that the result of such a measurement is affected strongly by diffraction. That is one reason why precision measurements should be avoided where Fresnel diffraction can occur on the wavefronts to be measured. The only way that precision measurements can be performed is to carefully image any limiting aperture to the detector, where the two wavefronts interfere with each other (Fraunhofer diffraction!). This is dealt with in detail in 4.5.

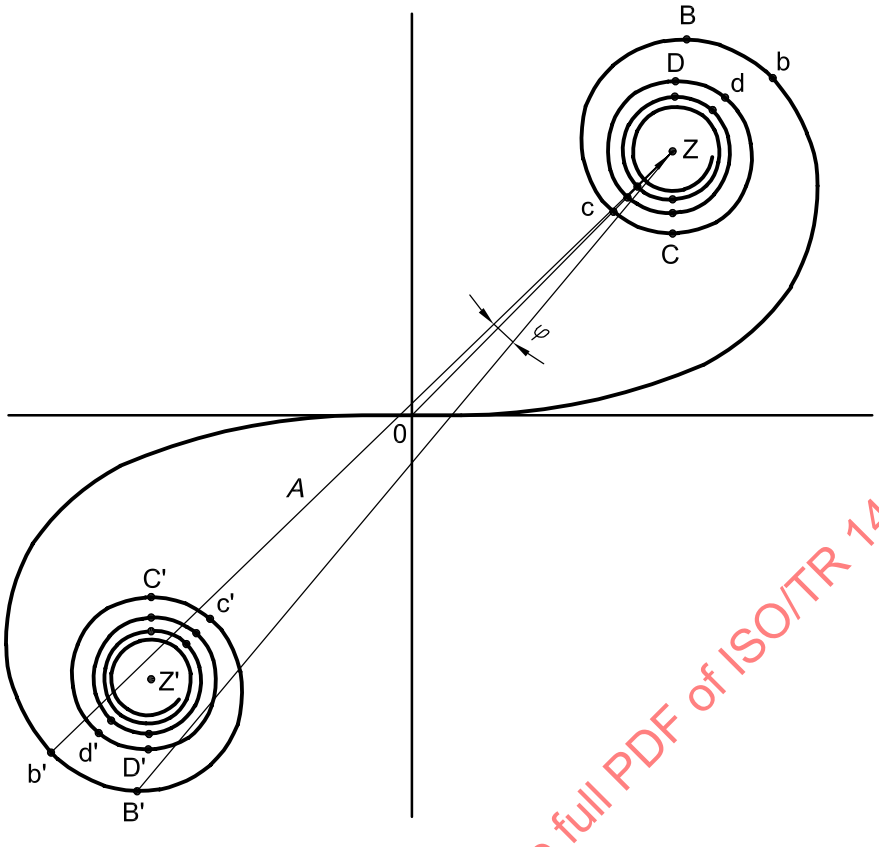


Figure 5 — Change of the phase-angle between the points b′ and B′ for the Fresnel diffraction on a knife edge

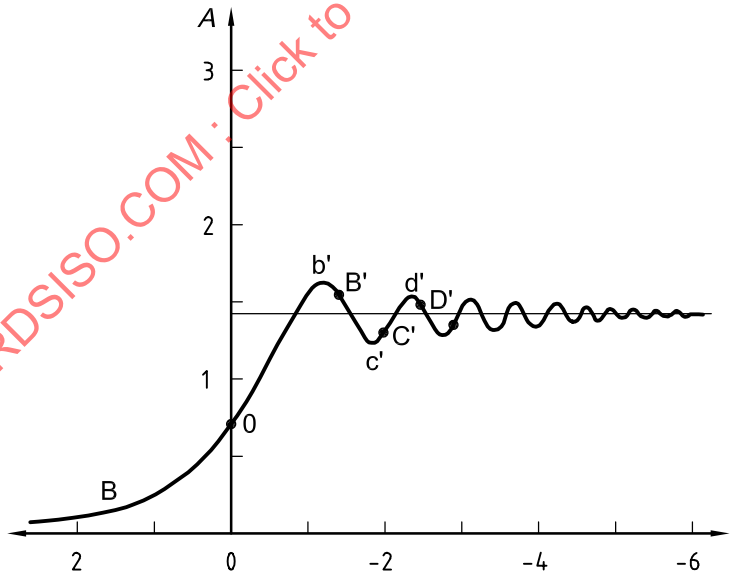


Figure 6 — Modulus A of the complex amplitude of the Fresnel diffraction pattern on a knife-edge

2.14 Propagation of aspherical waves

Any general wave which has either a more or less complicated function for the modulus or for the phase of the complex amplitude [(see Equations (11) and (12)], alters the values of both, modulus and phase, when propagating. In principle, the new wavefront after some distance z (see Figure 4) can be calculated with the

help of the Fresnel-Huygens principle, calculating the diffraction integral over the given distribution of the complex amplitude. For simplicity, this can be visualized with the help of the Huygens wavelets, where points on adjacent wavefronts have constant distances, measured along “light rays”, see Figure 7.

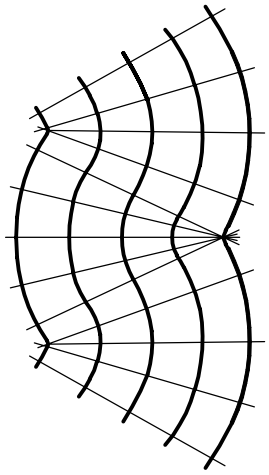
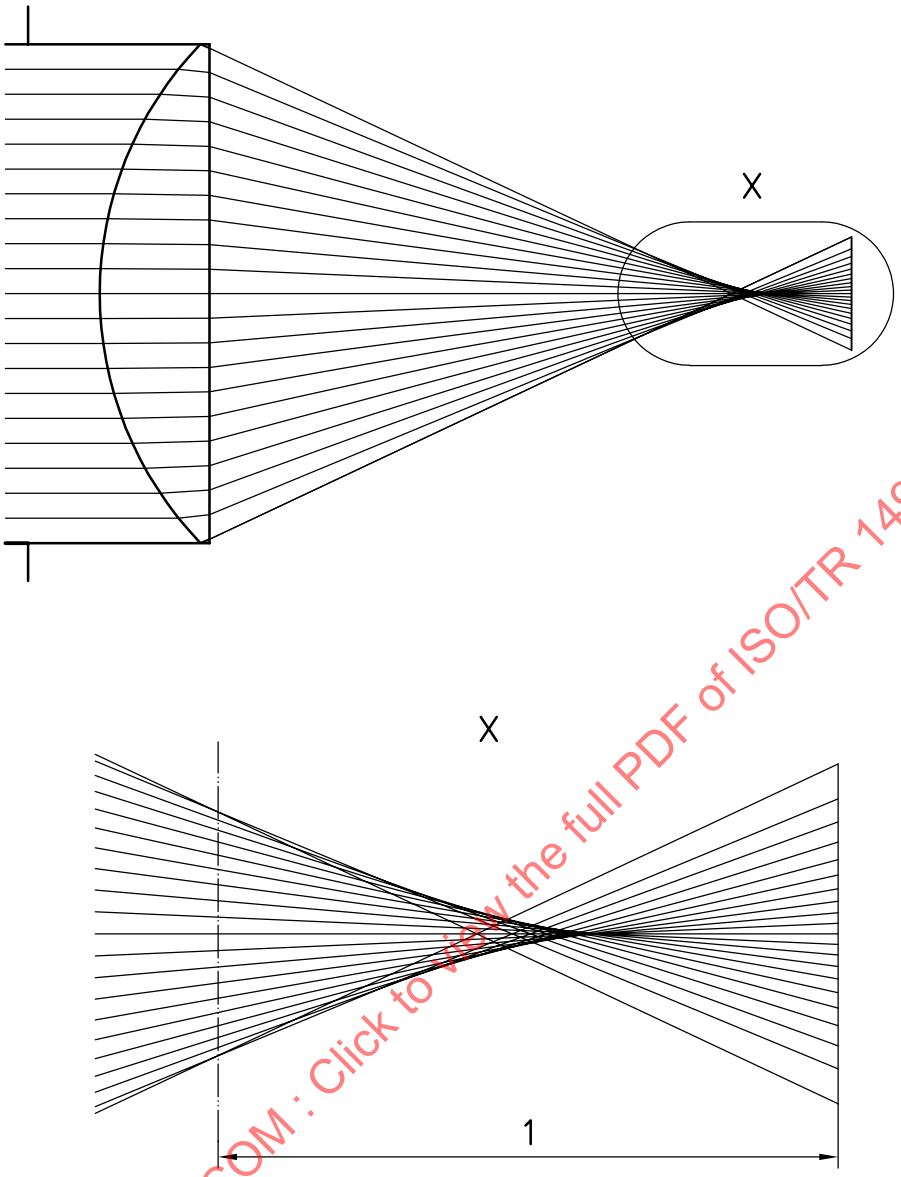


Figure 7 — Wavefront and light rays normal to each other

The same is true for any wavefront, which is not a sphere, a so-called aspherical wavefront. Such a wavefront can be characterized by the fact that the light rays do not come to a common centre point, but intersect along a region, called the caustic region. In Figure 8, such a caustic wave is shown.

If an aspherical wavefront is to be measured in an interferometer, the caustic region should be avoided, since the modulus of the complex amplitude varies greatly with the lateral coordinates and the correct mapping of the wavefront is very difficult. Also, the steadiness of the wavefront is not given in that region, leading to more difficulties. So, it is necessary to choose the region preceding the caustic for measurement of the phase-distribution. From the discussions before it is clear that the result for the measurement, e.g. for the phase distribution in the cross-section of the wavefront, where the measurement is carried out, depends very strongly on the axial position which is chosen for this cross-section. Since for practical reasons it will not be possible to locate the detector directly on that cross-section, the wavefront has to be imaged onto the detector. The result for the measured phase distribution depends then very strongly on the position, where the detector is placed in the region of the image of the wavefront under test. In other words, if the detector is shifted along the optical axis, the conjugate plane to the detector is shifted as well along the optical axis and a different cross-section of the aspherical wavefront is investigated. The result will be quite different, depending on the amount of asphericity of the wavefront under test. Special care has to be taken in order to have the imaging condition in the proper position.

The same is true for any deviation of a wavefront from the purely spherical or plane shape. If a surface is tested in reflection, that surface will alter the shape of the phase-distribution on the impinging wavefront. This phase-distribution of the wavefront is then measured, allowing to conclude from that result on the shape of the surface. Since the wavefront will alter its shape when propagating some distance from the surface, it is a requirement to properly image the surface under test onto the detector. This is especially necessary if the surface has errors with high spatial frequencies which give rise to high gradients on the phase-distribution. The shape of the wavefront can be altered significantly by moving more than a few micrometers away from such a surface. False results will be obtained.



Key

1 caustic region

Figure 8 — Due to the strong spherical aberration of a simple plano-convex lens, a caustic is produced

3 General description of interference and different types of interferometers

3.1 Interference between two waves

3.1.1 Intensity of two partial waves overlapping

Interference can occur when two or more waves overlap in space. Assume the simple case of two waves described by:

$$u_1 = U_1 e^{i\phi_1} \tag{24}$$

$$u_2 = U_2 e^{i\phi_2} \tag{25}$$

When u_1 and u_2 overlap, the resulting field simply becomes the sum, i.e.

$$u = u_1 + u_2 \tag{26}$$

The observable quantity is the intensity, which becomes:

$$I = |u|^2 = |u_1 + u_2|^2 = U^2 = U_1^2 + U_2^2 + 2U_1U_2\cos(\phi_1 - \phi_2)$$
$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\Delta\phi \tag{27}$$

where

$$\Delta\phi = \phi_1 - \phi_2 \tag{28}$$

3.1.2 Interference term of the intensity, destructive and constructive interference, mean value of the intensity

The resulting intensity does not become the sum of the intensities ($\neq I_1 + I_2$) of the two partial waves. The expression $2\sqrt{I_1 I_2}\cos\Delta\phi$ is called the interference term.

a) Case 1:

$$\Delta\phi = (2m+1)\pi \quad \text{for} \quad m = 0,1,2,...$$

$\cos\Delta\phi = -1$; $I = I_{min}$, i.e. the interference is destructive.

b) Case 2:

$$\Delta\phi = 2m\pi \quad \text{for} \quad m = 0,1,2,...$$

$\cos\Delta\phi = 1$; $I = I_{max}$, i.e. the interference is constructive.

The intensity distribution given in Equation (26) is sketched in Figure 9. The intensity varies between I_{max} and I_{min} with a mean value equal to:

$$I_m = \frac{I_{max} + I_{min}}{2} = I_1 + I_2 \quad . \tag{29}$$

For two waves of equal intensity, i.e. $I_1 = I_2 = I_0$, Equation (27) becomes

$$I = 2I_0[1 + \cos\Delta\phi] = 4I_0\cos^2\left(\frac{\Delta\phi}{2}\right) \tag{30}$$

where the intensity varies between 0 and 4 I_0.

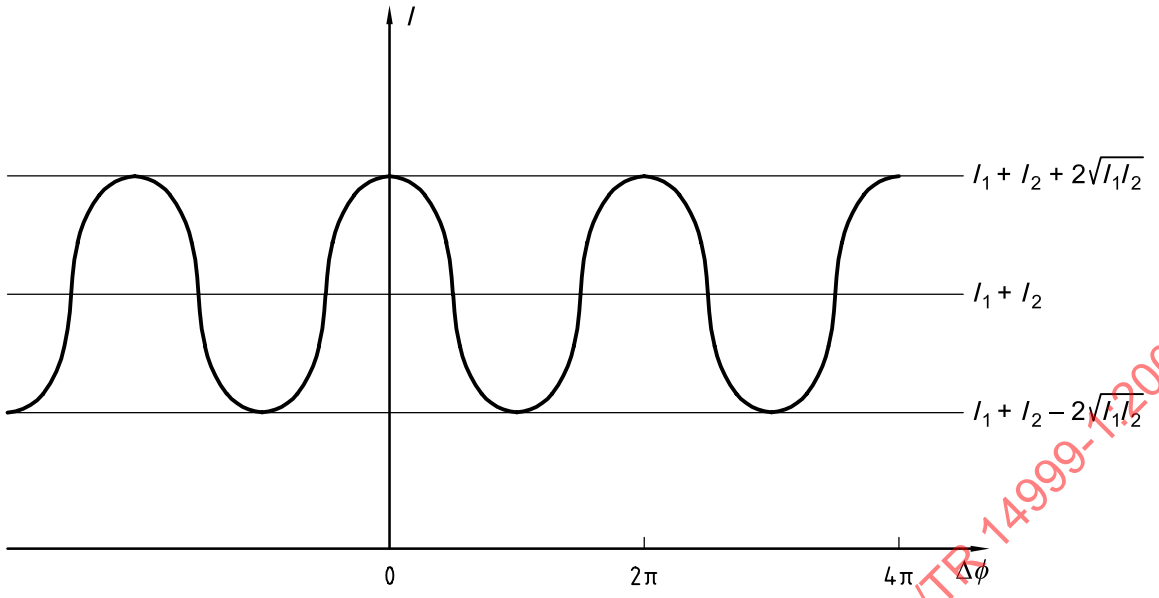


Figure 9 — Intensity distribution in the $x\,y$-plane from interference between two plane waves

3.2 Coherence

3.2.1 Temporal coherence: bandwidth, coherence time, coherence length

According to electromagnetic theory, the atoms of a light source do not continuously send out waves. Emission occurs through "wave trains", and there is a relationship between the length of the wave train and the spectral composition of the light emitted. The longer the wave train, the narrower the spectrum. At the theoretical limit, an infinite wave train consisting of monochromatic radiation of frequency ν_0 (the mean frequency emitted) would be emitted.

The coherence between two wave fields at one point in space is termed temporal or longitudinal coherence. One way of illustrating the light emitted by real sources is to describe it as sinusoidal wave trains of finite length with randomly distributed phase differences between the individual trains.

A source in an interference experiment, e.g. the Michelson interferometer is divided into two partial waves of equal amplitudes by a beam splitter where after the two waves are recombined to interfere after having travelled different paths.

In Figure 10, two successive wave trains of the partial waves are illustrated. The two wave trains have equal amplitude and length L_c, with an abrupt, arbitrary phase difference. Figure 10 a) shows the situation when the two partial waves have travelled equal path lengths. Although the phase of the original wave fluctuates randomly, the phase difference between the partial waves 1 and 2 remains constant in time. The resulting intensity is therefore given by Equation (27). Figure 10 c) shows the situation when partial wave 2 has travelled a path length L_c longer than partial wave 1. The head of the wave trains in partial wave 2 then coincide with the tail of the corresponding wave trains in partial wave 1. The resulting instantaneous intensity is still given by Equation (27), but now the phase difference fluctuates randomly as the successive wave trains pass by. As a result, cos ($\Delta\phi$) varies randomly between +1 and −1. When averaged over many wave trains, cos ($\Delta\phi$) therefore becomes zero and the resulting, observable intensity will be

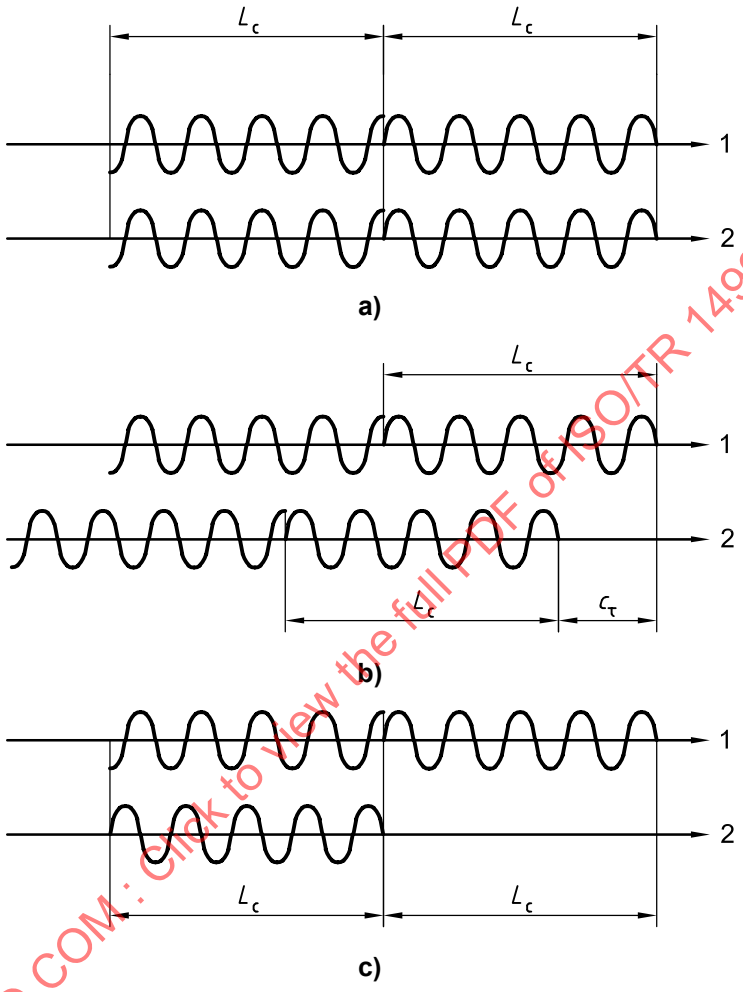
$$I = I_1 + I_2 \tag{31}$$

The length of the wave trains is called the coherence length l_c. Figure 10 b) shows an intermediate case where partial wave 2 has travelled a geometrical path difference Δl longer than partial wave 1, where $0 < \Delta l < l_c$. Averaged over many wave trains, the phase difference now varies randomly in a time period proportional to $\tau = \Delta l/c$ and remains constant in a time period proportional to $\tau_c - \tau$. If τ_c is the duration of each wave train and c is the speed of light, the coherence length is $l_c = c\,\tau$, where τ is called the coherence time.

The result is an interference pattern according to Equation (27) but with a reduced contrast. To account for this loss of contrast, Equation (27) can be written as

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\,|\gamma(\tau)|\cos\Delta\phi \tag{32}$$

with $\gamma(\tau)$ = complex degree of coherence or coherence function, where $|\gamma(\tau)|$ means the absolute value of $\gamma(\tau)$.



Key

1,2 partial waves

Figure 10 — Two successive wave trains of the partial waves

3.2.2 Coherence function, contrast

To see clearly that this quantity is related to the contrast of the pattern, we introduce the definition of contrast or visibility K:

$$K = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \tag{33}$$

The contrast is a function of τ:

$$K = \frac{2\sqrt{I_1 I_2}\,|\gamma(\tau)|}{I_1 + I_2} \tag{34}$$

For two waves of equal intensity, $I_1 = I_2$, and Equation (34) becomes

$$K = |\gamma(\tau)| \tag{35}$$

Three cases are identifiable:

$$|\gamma(0)| = 1 \tag{36}$$

$$|\gamma(\tau_c)| = 0 \tag{37}$$

$$0 \leqslant |\gamma(\tau)| \leqslant 1 \tag{38}$$

Equations (36) and (37) represent the two limiting cases of complete coherence and incoherence respectively, while inequality (38) represents partial coherence.

Of more interest is to know the value of l_c where $|\gamma(\tau)| = 0$.

In the case of a two-frequency laser this happens when

$$l_c = \frac{c}{\Delta\nu} = c\tau_c \tag{39}$$

where $\Delta\nu$ is the spectral bandwidth, the difference between the two frequencies.

It can be shown that this relationship applies to any light source with a frequency distribution of width $\Delta\nu$. l_c is termed the coherence length and τ_c the coherence time.

Sources of finite spectral width will emit wave trains of finite length. This is verified by the relation

$$\Delta\nu = \frac{c\Delta\lambda}{\lambda_0^2} \tag{40}$$

where

$\Delta\lambda$ is the bandwidth of the radiation;

λ_0 is the mean wavelength for the bandwidth of the radiation.

EXAMPLE 1 For a mercury vapour lamp with an interference filter with the width of the spectral lines $\Delta\lambda$ = 10 nm by the wavelength λ_0 = 546 nm the coherence length is:

$$l_c = \lambda_0\left(\frac{\lambda_0}{\Delta\lambda}\right) \approx 30\ \mu\text{m} \quad \text{and} \quad \tau_c = \frac{l_c}{c} = 10^{-13}\,\text{s} \tag{41}$$

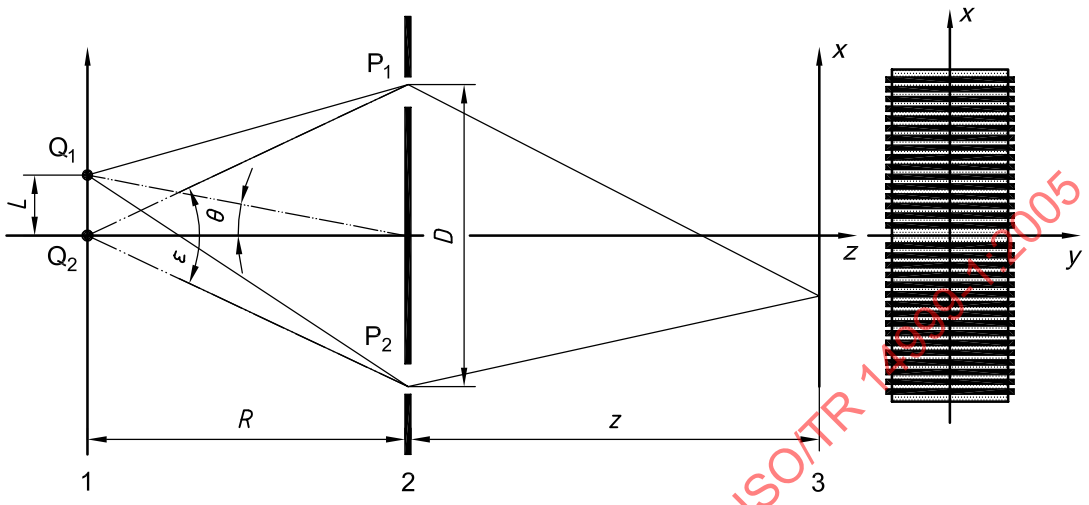
EXAMPLE 2 For a Helium-Neon-laser with λ = 632,8 nm and $\Delta\nu_{1/2}$ = 14 MHz [half width of the distribution $I_\nu(\nu)$], all vibrational modes are included, and the coherence length is:

$$l_c = c\tau_c \approx 20\ \text{m} \quad \text{and} \quad \tau_c = \frac{1}{\Delta\nu_{1/2}} = 0{,}7 \times 10^{-7}\,\text{s} \tag{42}$$

3.2.3 Spatial coherence

It is also possible to measure the coherence of a wave field at two points separated laterally in space. This phenomenon is called spatial or transverse coherence and can be analysed by the classical Young's double slit (or pinhole) experiment.

The incident wavefront is divided by passing through two small holes at P_1 and P_2 in a screen S_1. The emerging spherical wavefronts from P_1 and P_2 will interfere, and the resulting interference pattern is observed on the screen S_2 (see Figure 11).



Key

1 source
2 double slit S_1
3 screen S_2

Figure 11 — Young's double slit interference experiment

The geometric path length difference Δl of the light reaching an arbitrary point x on S_2 from P_1 and P_2 is found from Figure 11. When the distance z between S_1 and S_2 is much greater than the distance D between P_1 and P_2

$$\frac{D}{z} << 1 \tag{43}$$

then a good approximation is obtained as

$\frac{\Delta l}{D} = \frac{x}{z}$ and, after rearranging, gives

$$\Delta l = \frac{D}{z} x \tag{44}$$

The phase difference therefore becomes

$$\Delta\phi = \frac{2\pi}{\lambda}\Delta l = \frac{2\pi D}{\lambda z} x \tag{45}$$

which can now be inserted into the general expression for the resulting intensity distribution. Equation (27) gives

$$I(x) = 2I_0\left[1 + \cos\left(2\pi\frac{D}{\lambda z}x\right)\right] \tag{46}$$

We get interference fringes parallel to the y-axis with a spatial period a:

$$a = \frac{\lambda z}{D} \tag{47}$$

which decreases as the distance between P_1 and P_2 increases.

The distance $l_c = \lambda/\theta$ between P_1 and P_2 for which interference fringes disappear is called the spatial coherence length. The spatial coherence length l_c is inversely proportional to the diameter of the aperture θ in analogy with the temporal coherence length, which is inversely proportional to the spectral width.

The condition [Inequality (43)] therefore becomes:

$$D << l_c \text{ and } l_c = \frac{\lambda}{\theta} \tag{48}$$

Suppose the fringes on S_2 are formed, when the double slit is illuminated by a point source located at Q_2, on the symmetry axis between P_1 and P_2. From symmetry considerations constructive interference takes place on S_2 at $x = 0$, since the optical path length from Q_2 to that point on S_2 is equal in both cases, with the light passing either through P_1 or through P_2. At $x = 0$, $\Delta\phi = 0$ and therefore a bright fringe with a fringe order number of zero (the zero-order fringe). As already stated, adjacent to that zero-order fringe are further bright fringes at spatial distances given by Equation (47), which may be numbered with the help of order numbers $m = +1, +2, \ldots$ and $m = -1, -2, \ldots$, depending on the sign convention for $\Delta\phi$.

Now suppose the point source Q_2 is shifted by distance L to the position Q_1. Since the optical path difference from Q_1 to P_1 gets shorter and the optical path difference from Q_1 to P_2 gets longer, the zero-order fringe on S_2 is shifted down, as shown in Figure 11. All other fringes are shifted together with the zero-order fringe by the same amount. If Q_2 is shifted laterally a distance L, then the geometric path length difference Δl is equal to DL/R, as can be concluded from the quite similar consideration of Equation (44).

If the distance L is chosen such that the fringes are shifted exactly by an amount of one fringe period, the intensity at $x = 0$ has undergone one full period and the geometric path length difference Δl is equal to λ. When the intensity would be integrated during the shifting of the point source from position Q_2 to position Q_1, a mean intensity $2I_0$ would be measured. The same is true for all other points on the screen.

If the light source has a spatial extent equal to L, the integration over all points of the light source should be carried out in space rather than in time but with the same result, i.e. no fringes are visible on S_2. Therefore for a light source of extent L, the spatial coherence is zero. From this consideration, it can be concluded that the spatial coherence function can be found by integrating the equation for the intensity of the two-beam interference, with the source being the region of integration. In the case of a slit-like source of width l, a sinc-function results, having its first zero position at $l = L$. In the general case, the van Cittert-Zernike theorem states that the degree of coherence between two points is the modulus of the scaled and normalized Fourier transform of the source intensity distribution.

The limits between "coherent" and "incoherent" are somewhat arbitrary. In the discussion of the source of extent L a value was chosen where the variation of the path length difference Δl equals 1 λ. In this case, no fringes are visible at all. If it still desirable to see fringes, but on a lower contrast level, take the limit for the source size more stringently. It is the convention to define the limit for the spatial coherence such that the path length difference is $\Delta l = \lambda/2$.

Since $R >> D$, D/R might be replaced by the angle ε, under which the two points P_1 and P_2 appear when observed from the source. Equally, since $R >> L$, L/R may be replaced by θ.

By assuming an extended source with incoherent light, the condition for spatial coherence is given by:

$$\frac{DL}{R} \leqslant \frac{\lambda}{2} \qquad L\varepsilon \leqslant \frac{\lambda}{2} \qquad D\theta \leqslant \frac{\lambda}{2} \tag{49}$$

3.3 Different arrangements of interference between two beams

3.3.1 General

In order to produce interference phenomena, arrangements called interferometers are used. Most interferometers consist of the following elements:

a) light source;

b) means for shaping the waves emerging from the light source;

c) element for dividing the light into two partial waves;

d) two different propagation paths, where the partial waves experience different phase-contributions;

e) element for combining the partial waves;

f) means for imaging the waves onto the detector;

g) detector for observation of the interference.

The elements a) and b) are called the “illumination part” of the interferometer, the elements c), d) and e) are called the “interference part” of the interferometer and the elements f) and g) are called the “imaging part” of the interferometer. Some arrangements may not include all seven elements, and some of the elements may not be clearly distinguished. In general, the most challenging part is the interference part since any disturbance which the waves encounter in this part is included in the quantity which is measured by an interferometer. This quantity is the phase-difference between the two wavefronts in the two different propagation parts (the two “arms”) of the interferometer. Nevertheless, for a thorough treatment of the phase differences between the two interfering waves, it is not sufficient to investigate only the interference part, the illumination and the imaging part have to be taken into account as well. The reason is that the light beams which are superimposed onto each other at the detector, might have travelled slightly different paths in the illumination and in the imaging part of the interferometer. This path difference might influence the measured phase difference. This is especially true when some tilt is introduced between the interfering waves at the detector (“non-nulled” interferogram).

A general differentiation of interference arrangements can be done by the manner in which the light from the light source is divided [see c) in the above list] to serve the two channels. It is necessary to distinguish between wavefront division and amplitude division of wavefronts. In the case of geometrical division, the wavefront is split into a right part and into a left part which do not overlap each other at the stage of wavefront splitting. Looking from the light source towards the optical axis of both wavefronts, an angle ε is established between them. This imposes a constraint on the source-size L, as can be concluded from the condition for spatial coherence, Equation (49).

Physical or amplitude division of a wavefront can be done by a semitransparent mirror, a so-called beam-splitter. The same functionality might be obtained by a grating, using different diffraction orders for the two channels. In this case, the angle ε can be kept small or even zero, therefore no constraints are imposed on the extent of the source. For other reasons (beam collimation with little divergence), the source size is also restricted in practical instruments.

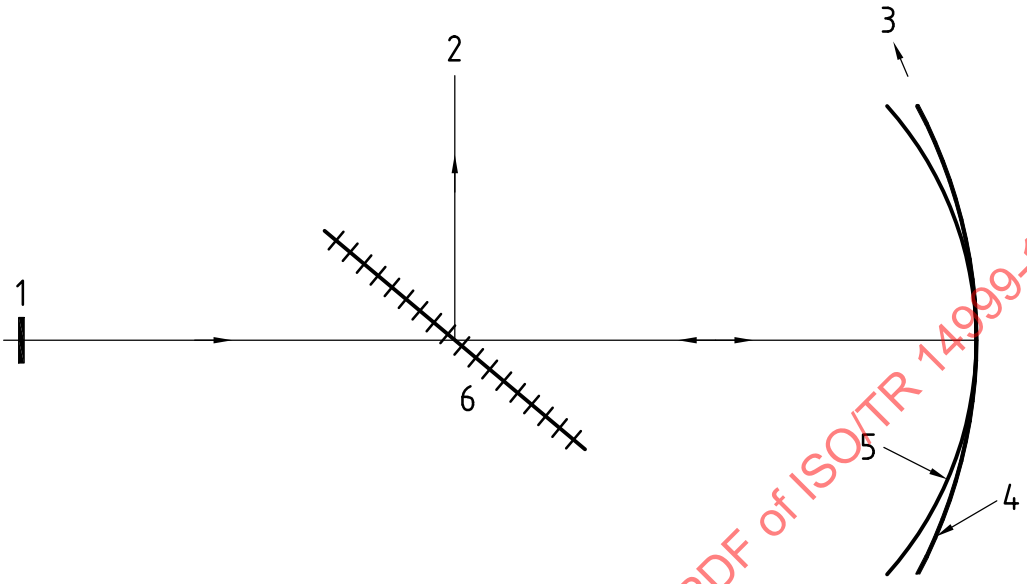
3.3.2 Simple interferometer structures: Fizeau, Haidinger, Newton

3.3.2.1 General

Newton, Fizeau, and Haidinger interferometers are the simplest and most powerful tools available to a working optician. With very little effort these interferometers can be set up in an optical workshop for routine testing of optical components to an accuracy of a fraction of the wavelength of light. Even though these instruments are simple in application and interpretation, the physical principles underlying them involve a certain appreciation and application of physical optics.

3.3.2.2 Newton interferometer

A Newton interferometer is an arrangement of two surfaces M and V in contact and illuminated by a monochromatic sources of light as seen in Figure 12. Fringes obtained with a wedge depend on the thickness of the gap. The fringes lie along the lines where the gap width is constant (fringes of equal thickness).



Key

1 source
2 observation
3 air gap
4 surface under test V
5 test surface M
6 beam splitter

Figure 12 — Newton interferometer

Newton's rings, a special type of fringe pattern, result when two spherical surfaces, a polish surface under test and a test glass of opposite curvature, are put together. Light is incident from the top.

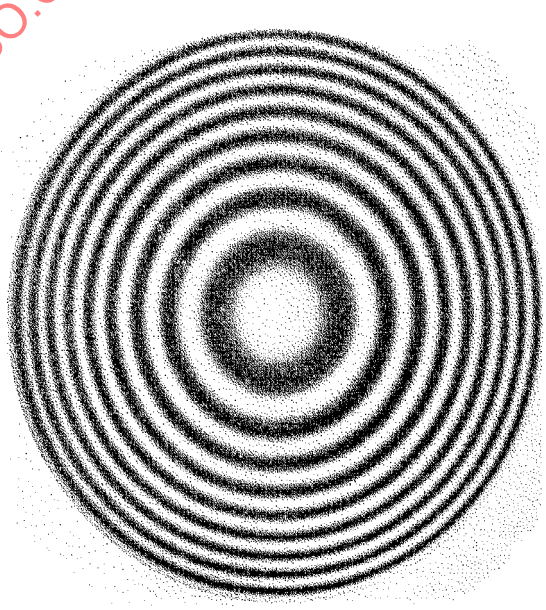


Figure 13 — Newton's rings

The fringes now extend in circles at gaps of equal thickness. Testing surfaces of small lenses while they are being polished is one of the most common applications of the Newton interferometer. A test glass is useful not only to detect surface irregularities but also to check the deviation of the radius of curvature from the desired value. For reflected light, there is a dark spot in the middle; for transmitted light, there is a bright spot (see Figure 13). The qualitative description of the interference pattern is similar to that given for the wedge-shaped gap. For reflected light, a phase shift of π corresponding to a path difference of $\lambda/2$ is obtained at the centre, resulting in destructive interference. For transmitted light, there is no phase difference and constructive interference is observed.

If m dark fringes are obtained at the lens surface, the deviation in the radius r of curvature is given by:

$$\Delta r = \frac{4m\lambda r^2}{D^2} \tag{50}$$

where D is the diameter of the surface.

EXAMPLE For $r = 40$ mm, $D = 40$ mm, $\lambda = 500$ nm, $m = 2$, the deviation in the radius is $\Delta r = 4$ µm and the relation $\Delta r/r = 10^{-4}$. For $k = 1$, $\Delta r = 2$ µm; for $\Delta r < 2$ µm equal thickness fringes will not be observed.

3.3.2.3 Fizeau interferometer

In the Fizeau interferometer the air gap between the two surfaces M and V is much larger than in the Newton interferometer. The angular size of the source to be used depends on the air gap. The optical path difference, Δl_{OPD}, between two reflected rays at an air gap is given by the expression

$$\Delta l_{\mathrm{OPD}} = \frac{2d}{\cos\beta} - 2d\tan\beta\sin\beta \tag{51}$$

where

d is the thickness at an air gap between M and V;

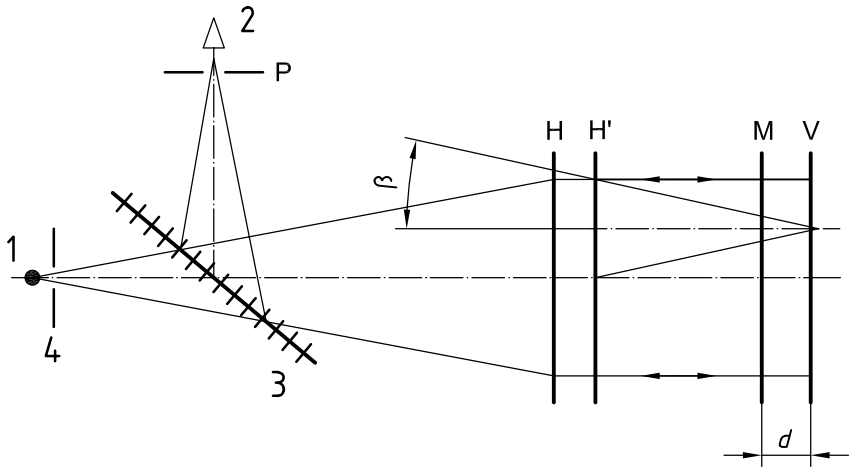
β is the semi-angle of incidence of the light source.

In order to get high contrast fringes, it is necessary to restrict the semi-angle β of the illumination source to:

$$d\beta^2 \leqslant \frac{\lambda}{4} \tag{52}$$

It is seen that it is necessary to have a collimating system in a Fizeau interferometer. Figure 14 shows the schematic arrangement of a Fizeau interferometer using a lens for collimation. To view the fringes, a beam splitter is located close to the pinhole. When the air wedge is very large, two distinct images of the pinhole can be seen originating from the two surfaces M and V in the plane P in Figure 14. By making use of screws provided to tilt the flat under test, the movement of the image through the pinhole is observed and can stop when it coincides with that of the reference flat.

Then the observer sees the fringes due to variation in the air gap thickness. The interpretation of the Fizeau fringes is exactly the same as that for Newton fringes. These fringes are called fringes of equal thickness.



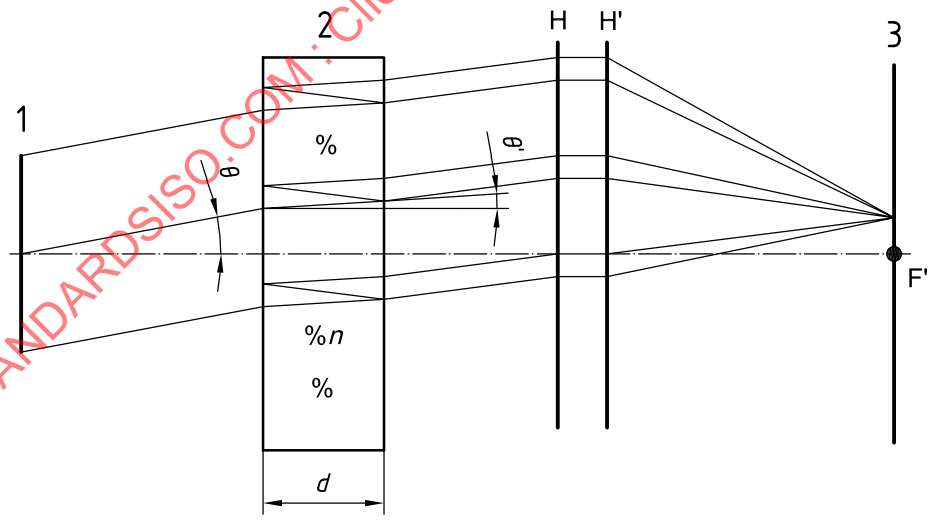
Key

1 source
2 observation
3 beam splitter
4 pinhole
d air gap

Figure 14 — Fizeau interferometer

3.3.2.4 Haidinger interferometer

In this case, the fringes are referred to as fringes of equal thickness. The thickness of the air gap is uniform and it is illuminated by a source of large angular size. The Haidinger fringes are called fringes of equal inclination. These fringes are formed at infinity and a lens HH′ can be used to focus them on its focal plane (see Figure 15).



Key

1 monochromatic source of light
2 parallel plate of glass
3 image plane

Figure 15 — Haidinger interferometer

If the air gap is replaced by a very good parallel plate of glass, then the following simple relation:

$$2nd\cos(\theta') = m\lambda \tag{53}$$

where

n is the refractive index of the glass plate;

d is the thickness of the glass plate;

θ' is the angle of refraction inside the glass plate.

For small values of θ' the Equation (53) can be approximated as:

$$2nd - \left(\frac{d}{n}\right)\theta'^2 = m\lambda \tag{54}$$

The Haidinger fringes may be used for testing a nearly parallel plate. The wedge angle may be found either with the Fizeau or with the Haidinger method. With the Haidinger interferometer the observer looks for the stability of the concentric fringes as he moves his line of sight across the plate with a small aperture. If d is slowly varying the centre of the circular fringe system also appears to change. As the observer moves towards the thinner side of the wedge, d decreases. In this case, the circular Haidinger fringes appear to expand from the centre. Conversely, the fringes appear to converge to the centre if the observer moves toward the thick side of the wedge. The wedge angle can be estimated in the same manner as for the Fizeau method if the observer notes how many times the centre of the fringe system has gone through bright and dark cycles.

3.3.2.5 Twyman-Green interferometer

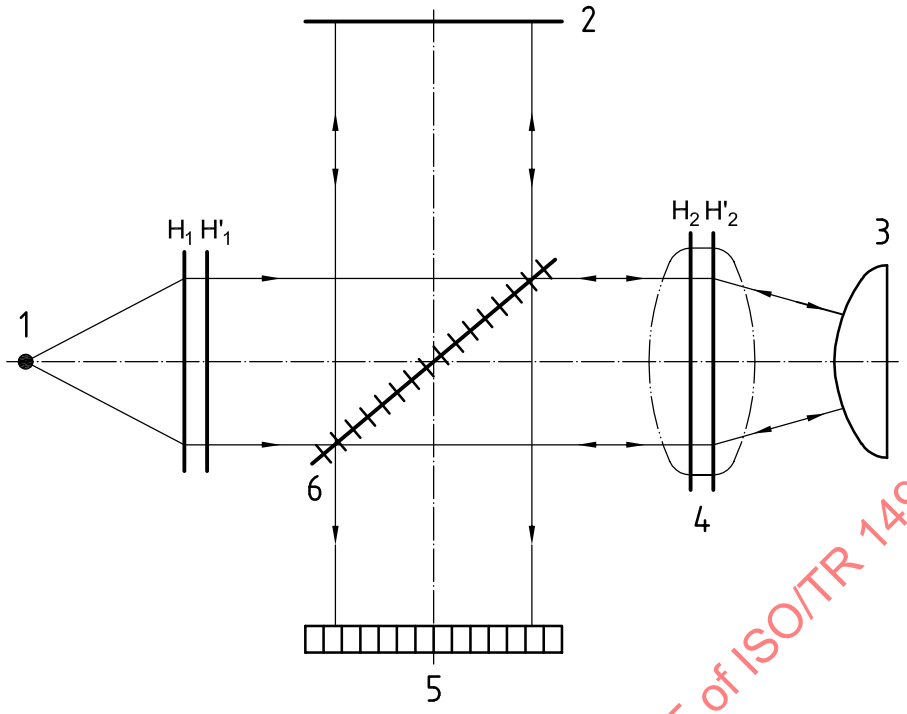
The Twyman-Green interferometer is a modification of the Michelson interferometer used to test optical elements. The basic Twyman-Green configuration is illustrated in Figure 16.

The system is illuminated with a quasi-monochromatic (or monochromatic) point light source at the focus of a well-corrected lens system $H_1H'_1$. So the light is collimated by means of the first lens system in order to form plane wavefronts.

The amplitude of the wavefront is then divided by means of a beam splitter plate. In Figure 16, a convergent optical test element $H_2H'_2$ with long focal length, is inserted in arm M_2. After reflection, light from both mirrors, the flat mirror M_1 and the convex spherical mirror M_2 with its centre of curvature at the focus of the lens, impinges again on the beam splitter. If the optical test element is perfect, the returning wavefront is a plane. If the flat mirror M_1 is inclined at a suitable small angle α, parallel equidistant fringes are visible on the observing screen.

The lens under test can be tilted about an axis perpendicular to its optical axis to allow tests at different field angles, and a mechanical linkage ensures that the centre of curvature of M_2 remains in the focal plane as the lens is tilted. There might be some uncertainty, whether the focal plane of the lens is maintained during the tilt by this mechanical means.

The wavefront aberrations introduced by the lens $H_2H'_2$ might result from residual design errors or from fabrication errors of several components of the lens. These wavefront aberrations have to be located somewhere; it is common to locate them in one of the pupils of the lens, for further computation of quality numbers like rams wavefront aberration or computation of the point-spread-function or the modulation-transfer-function. In the case of the Twyman-Green interferometer, the lens is tested in double passage. Therefore the entrance pupil of the lens is imaged by the mirror M_2 to some other location and then imaged again by the lens $H_2H'_2$ to another location. This latter location will normally not coincide with the position of the entrance pupil of the lens, so it is necessary to decide which plane should be imaged on the detector (see Clause 4 for details). Ideally, the entrance pupil should be imaged with a magnification of $+1$ by the lens together with the mirror M_2; in this case, the image would coincide with the pupil itself. In the testing of high numerical aperture microscopic lenses with a Twyman-Green interferometer, a special arrangement consisting of a half-sphere of radius R_1 made of glass with an index of refraction of n and a concave mirror with a radius of $R_2 = R_1\,[n/(n-1)]$ replaces the simple convex mirror M_2 of Figure 16. Using this configuration, Dyson [13] has managed to image the exit pupil onto itself.



Key

1 light source
2 flat mirror M_1
3 convex spherical mirror M_2
4 convergent lens under test
5 observing screen
6 beam splitter

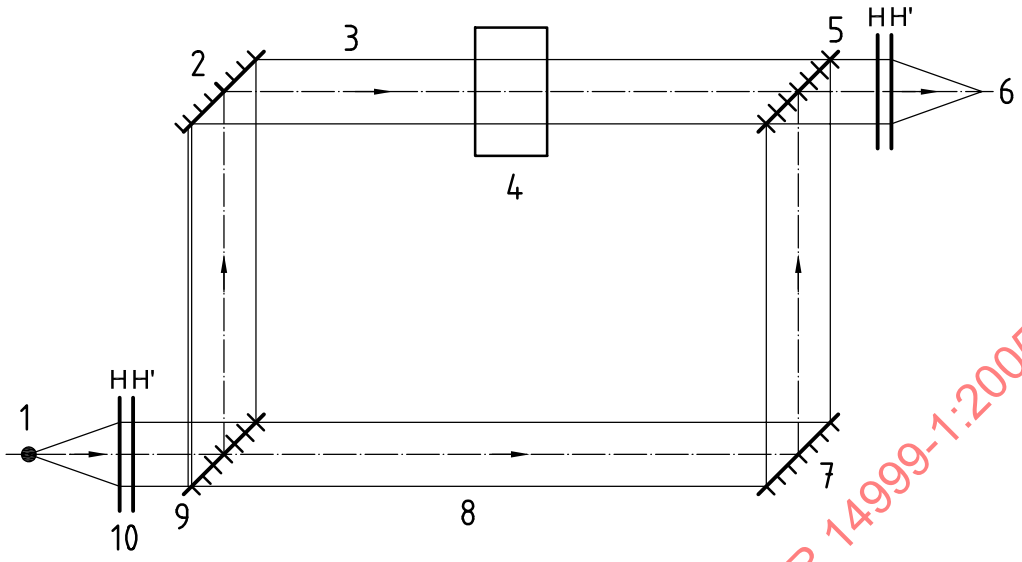
Figure 16 — Twyman-Green interferometer

The Twyman-Green interferometer arrangement is equivalent to the Fizeau interferometer, but with the beams having clearly separated paths (see Annex A).

3.3.2.6 Mach-Zehnder interferometer

The Mach-Zehnder interferometer as shown in Figure 17 is a variation from the Twyman-Green configuration, which has some advantages with respect to the Twyman-Green configuration. If a sample has large aberrations, it is better to pass the beam of light through the sample under test only once. An additional advantage is that this interferometer is automatically compensated because it has two identical beam splitters.

A Mach-Zehnder interferometer can be used to test the wave front quality of the light source[14][15], to test off-axis paraboloids[16], to align equilateral hyperbolic zone plates[17] or to analyse the errors that result from a test section misalignment.



Key

1 light source
2 flat mirror
3 test beam
4 sample under test
5 beam splitter
6 observer
7 flat mirror
8 reference beam
9 beam splitter
10 collimator

Figure 17 — Mach-Zehnder interferometer

3.4 Characteristic features for interferometer structures

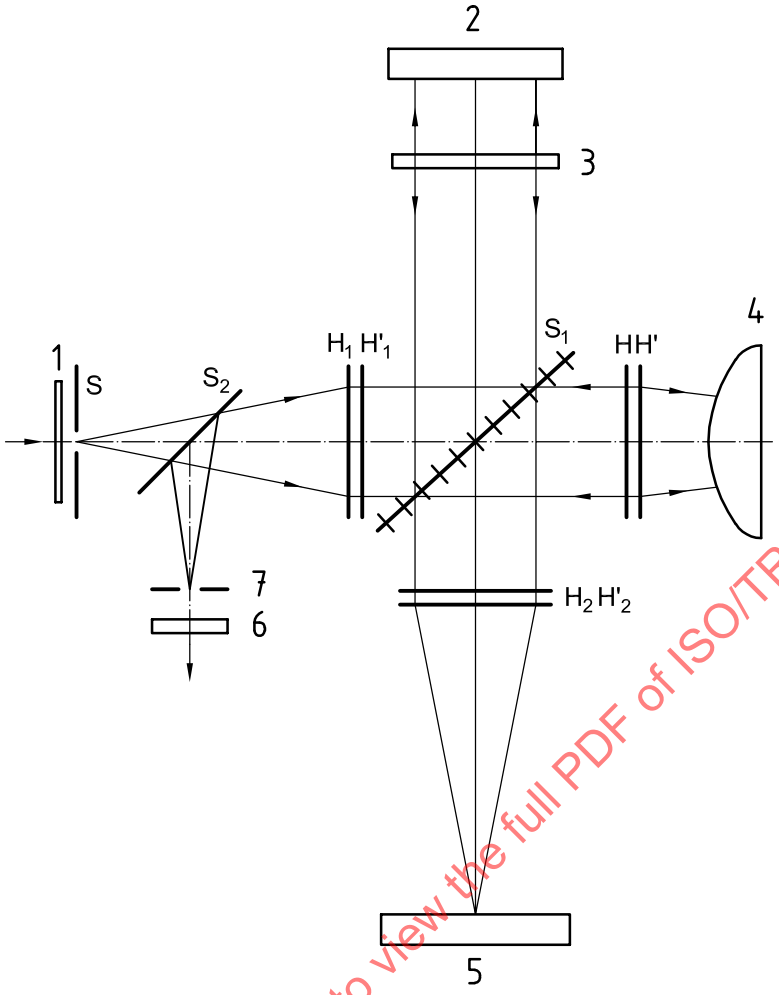
3.4.1 Double-pass and multiple-pass

3.4.1.1 General

Double- or multiple-pass interferometers offer definite advantages for some testing applications. In these interferometers one or more of the wavefronts are sent back and make two or more traverses of either the whole system or a part of it. The interference pattern which is obtained with a lens in a Twyman-Green interferometer gives a contour map of the wavefront leaving the lens aperture. The problem is the separation of the individual aberrations, when many aberrations are present. This problem can be simplified if the Twyman-Green interferometer is used in a double-pass configuration (see References [4] and [12]) or in a multiple-pass configuration (see References [4] and [18]).

3.4.1.2 Double-pass configuration

This configuration is shown in Figure 18. The beams emerging from the interferometer through the lens L_2 are reflected back through it by the plane mirror M_3 which is placed at its focus. The double-pass beams emerging from L_1 are brought to a focus at the eye stop by the auxiliary beam splitter S_2. The two images formed at the eye stop move off the axis in opposite directions, if the source is shifted very slightly sideways. So it is possible to view either the fringes produced by the double-pass beams or the normal interference pattern.



Key

1	polarizer
2	M_1-reference mirror
3	$\lambda/4$ plate
4	M_2-test mirror
5	M_3-additional mirror
6	analyser
7	eye stop B_1
HH′	test lens
$H_1H'_1$	lens L_1
$H_2H'_2$	lens L_2
S	source
S_1, S_2	beam splitters

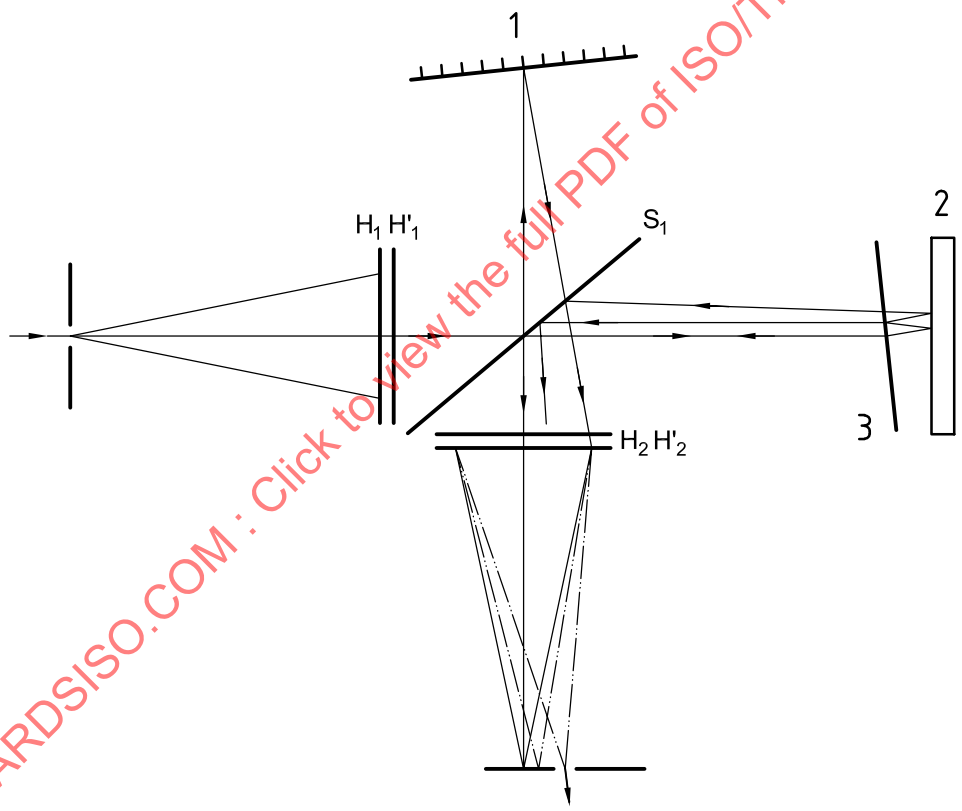
Figure 18 — Double-pass Twyman-Green interferometer for separation of symmetrical and antisymmetrical wavefront aberrations

In classical interferometry, the accuracy with which measurements can be made with a conventional two-beam interferometer is determined by the precision of locating any fringe displacement ΔQ. The displacement ΔQ can be estimated as a fraction of Q, where Q is the average spacing of the fringes. This precision is limited by the sinusoidal irradiation distribution in two beam fringes. One method to obtain greater accuracy is the multiple-beam interferometry (see 3.4.2) in order to get sharper fringes. At the present time, phase measurement schemes are used for determining the phase-distribution with very high precision.

3.4.1.3 Multiple-pass interferometry

In multiple-pass interferometry, the direct visual presentation of wavefront errors with increased sensitivity can be obtained. The fact that the total deformation of the wavefront is proportional to the number of times it is reflected from or transmitted through the optical system under test is utilized to increase the ratio of the fringe displacement ΔQ to the average fringe spacing Q, for the same error given by the system under test.

In Figure 19, a typical configuration of the Twyman-Green interferometer[18] is shown schematically. In one arm of the interferometer an additional beam splitter is inserted. The beam splitter makes a slight wedge angle with respect to the mirror under test. So the multiple reflected beams formed between the beam splitter and the mirror give rise to a series of laterally separated images of the light source in the rear focal plane of lens L_2. This technique permits selecting a beam that has undergone any desired number of reflections at the mirror under test by an aperture in the focal plane of lens L_2. Two beam interference fringes are obtained between the reference wavefront and a wavefront that has undergone n reflections at the mirror under test, if the reference mirror is tilted so that the light reflected by it also passes through the same aperture. Therefore the wavefront exhibits a deformation of $2n\Delta t$, where Δt is the deviation of its surface from flatness.



Key

1	reference mirror
2	test surface
3	auxiliary beam splitter S_2
S_1	beam splitter
$H_1H'_1$	lens L_1
$H_2H'_2$	lens L_2

Figure 19 — Multiple-pass Twyman-Green interferometer

For the ratio of the fringe displacement ΔQ to the average interfringe, spacing Q is given by the following equation:

$$\Delta Q/Q = 2n\Delta t/\lambda \tag{55}$$

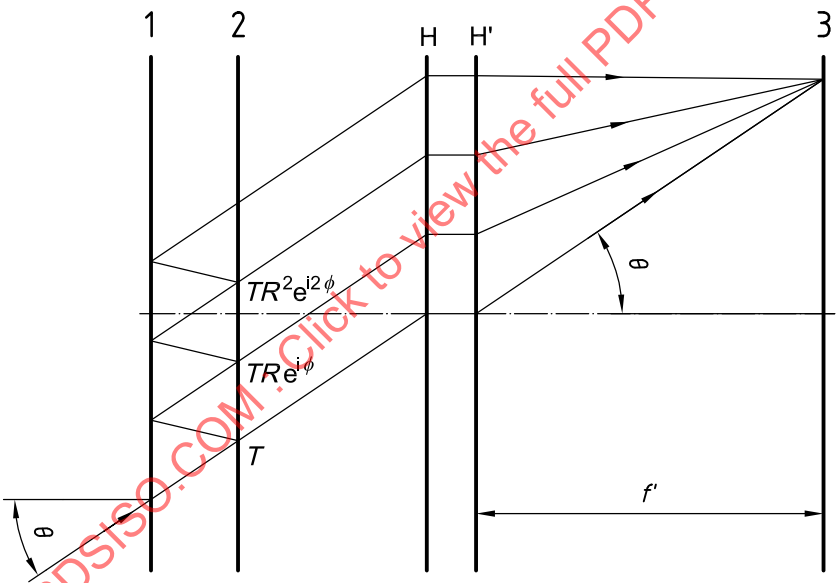
Therefore the sensitivity is n times greater than that obtained with a normal Twyman-Green interferometer.

3.4.2 Two-beam interference and multiple-beam interference

Multiple-beam interferometry is a method of obtaining high spatial precision. One of the most compact multiple-beam interferometer is the Fabry-Perot interferometer. It consists of two glass or quartz plates with plane surfaces.

The inner surfaces coated with partially transparent films of high-reflectivity are parallel, so that they enclose a plane parallel plate of air. In the original form of this interferometer, one plate was fixed and the other was mounted on a screw-controlled carriage to allow continuous variation of the plate separation. So a single incident plane wave of light is incident upon the transparent mirror 1 at angle θ and produce a series of waves by multiple reflection.

All of these waves can be superposed to form ideal Airy fringes at the focal plane of a lens, as seen in Figure 20. In contrast, multiple-beam fringes are extremely sharp and can have local finesse as high as 300 with surfaces of about 99 % reflectance.



Key

1 mirror 1
2 mirror 2
3 focal plane
HH′ lens

Figure 20 — Fringe formation by a multiple-beam Fabry-Perot interferometer

The Fabry-Perot interferometer may be used for high resolving power spectroscopy when it is illuminated by quasi-monochromatic light or the plates act as a wavelength filter when a parallel beam of white light is incident on the plane parallel plates having surfaces of high reflectivity.

The irradiance distribution, $I(\phi)$, is given by:

$$I(\phi) = \left| \sum_{n=0}^{\infty} TR^n e^{in\phi} \right|^2 = \left(\frac{T}{1-R} \right)^2 \frac{1}{1+\left(4N_R^2/\pi^2\right)\sin^2\phi/2} \tag{56}$$

where

N_R is the ideal reflective finesse as follows:

$$N_R = \frac{\pi\sqrt{R}}{1-R} \tag{57}$$

T, R are the intensity transmittance and reflectance, respectively, of both mirrors;

ϕ is the total effective phase or optical path delay between any two consecutive wavefronts.

In practice the total effective number of interfering beams is not infinite because of the fact that the energy carried by the nth transmitted beams, $T^2 R^{2n}$, becomes negligible for sufficiently large n.

An additional fact of practical interest in multiple-beam Fabry-Perot fringe formation is that the phase delay ϕ between any two consecutive wavefronts is constant. Hence, once ϕ has been determined, the entire series of multiple reflected beams combines to form the appropriate part of the ideal Airy fringe.

The Fabry-Perot fringes are fringes of equal inclination. These fringes become much narrower when the reflectivity of the surfaces is increased. Increased reflectivity has a similar effect on the intensity distribution of the Fizeau fringes of equal thickness given by a thin film. The result is that finer details of the thickness contours of the film are revealed.

For simplicity, multiple-beam Fizeau interferometry can be considered a thin film in the form of a wedge with plane surfaces inclined at a small angle θ, illuminated with a collimated wave of monochromatic or quasi-monochromatic light. Because of multiple reflections at the surfaces, the transmitted light consists of a set of plane waves propagated in different directions. The multiple-beam interference pattern may also be observed in reflection. When there is no absorption at the reflecting surfaces, this pattern is complementary to the transmitted pattern in the sense that the sum of the intensities in the two is at each point equal to the incident intensity.

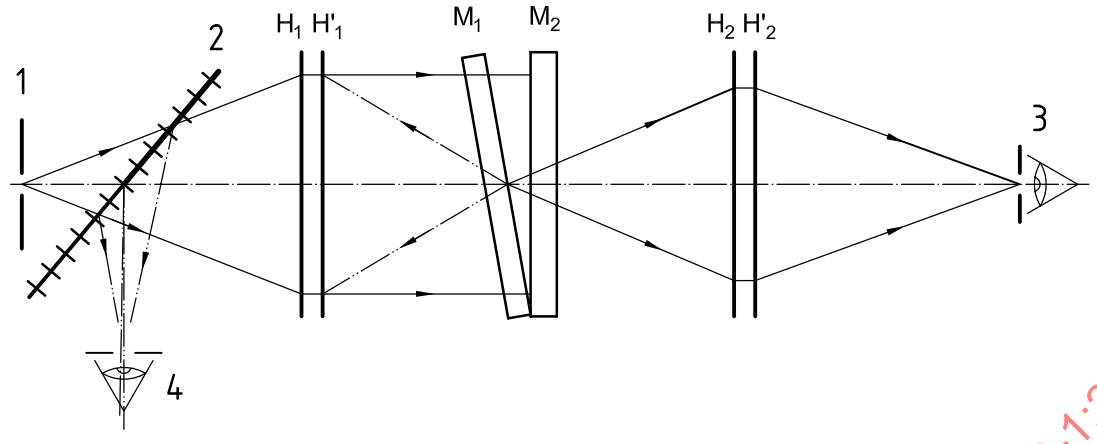
The fringes in transmission are sharp and bright, in very high contrast against an almost dark background, and those in reflection are exactly complementary, they are dark fringes against a bright background.

Multiple-beam Fizeau fringes are used in optical workshops to test the figure of high-grade optical flats such as are required for the Fabry-Perot interferometer. The surface contour fringes show the total optical path variation of both flats.

Practical differences between a Fabry-Perot interferometer and a Fizeau interferometer are the following.

a) The phase delay, between the constructive wavefronts produced by the wedged Fizeau mirrors, is a progressively increasing quantity, rather than a constant one as in a parallel Fabry-Perot.

b) The spatial walk-off of the beams with multiple reflection in the Fizeau interferometry cannot be compensated for perfectly by any focusing or imaging device.



Key

1	light source
2	beam splitter
3	observing transmission
4	observing reflection
M_1, M_2	plane surfaces

Figure 21 — Essential components of Fizeau interferometric arrangement

3.4.3 Common path

In so-called common-path interferometers, the reference and the test beams traverse the same general path. In the general type of interferometer, where the reference and test beams follow widely separated paths, mechanical shocks and temperature fluctuations affect both paths differently, giving rise to disturbances of the interferogram. These problems are particularly relevant when optical systems of large aperture are being tested. Most of these difficulties can be avoided in common-path configurations, since disturbances act similarly on both beams and therefore cancel out in the measured phase-difference of both beams. At the same time, the common-path interferometers do not require almost perfect optical components of dimensions equal to those of the system under test for producing the reference beam. Furthermore, the use of white light is possible, hence the path difference between the two beams in the centre of the field of view is naturally zero. In most common-path interferometers the reference and the test beams are affected by the system aberrations, and interference is produced by shearing one beam with respect to the other. Hence the information obtained is implicit, and some computations are needed to determine the shape of the aberrated wavefront.

Beam splitting can be realized by amplitude division with the help of a partially scattering surface, such as a doubly refracting crystal, or a semi-reflecting surface.

Burch's scatter plate interferometer, used for testing a concave mirror is schematically shown in Figure 22. A lens $H_1H'_1$ forms an image of a small source S at S′ on the mirror. The beam splitter (the scattering plate) is located at the centre of curvature C of the mirror with a radius R. By photographing and bleaching a speckle pattern, two identical scatter plates R_1 and R_2 can be made. The mirror M forms an image of R_1 at R_2. This is placed the other way round so that there is a point-for-point coincidence between R_2 and the image of R_1. A semi-reflecting plate B directs the light returning from the mirror M onto R_2. A part of the light incident on the scatter plate R_1 passes through it without scattering and arrives at S′. This beam acts as the reference beam, and it is not affected by the errors of the mirror surface.

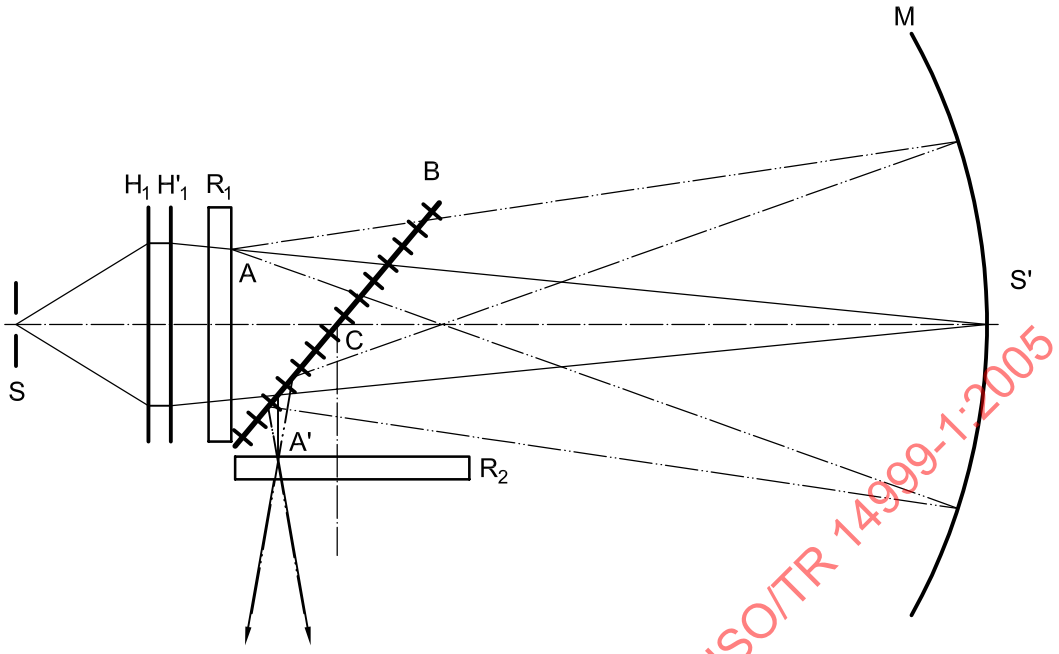


Figure 22 — Burch's scatter plate interferometer

Some of the incident light is scattered by R_1 and fills all of the aperture of M. This beam is the test beam which picks up the errors of the mirror. One beam is directly transmitted by R_1 and scattered by R_2, and the second is scattered by R_1 and transmitted by R_2. So two mutually coherent beams emerging from R_2 overlap, and an observer looking at the mirror surface through R_2 can see an interferogram between two beams. If the mirror is free of any errors in the region of S', the interferogram will provide explicit information.

3.4.4 Phase conjugation

Phase conjugation is a mathematical operation by which the complex amplitude of a wave is converted into its conjugate complex amplitude. For example, the complex amplitude of a spherical wave and its conjugate complex amplitude can be written in the following form:

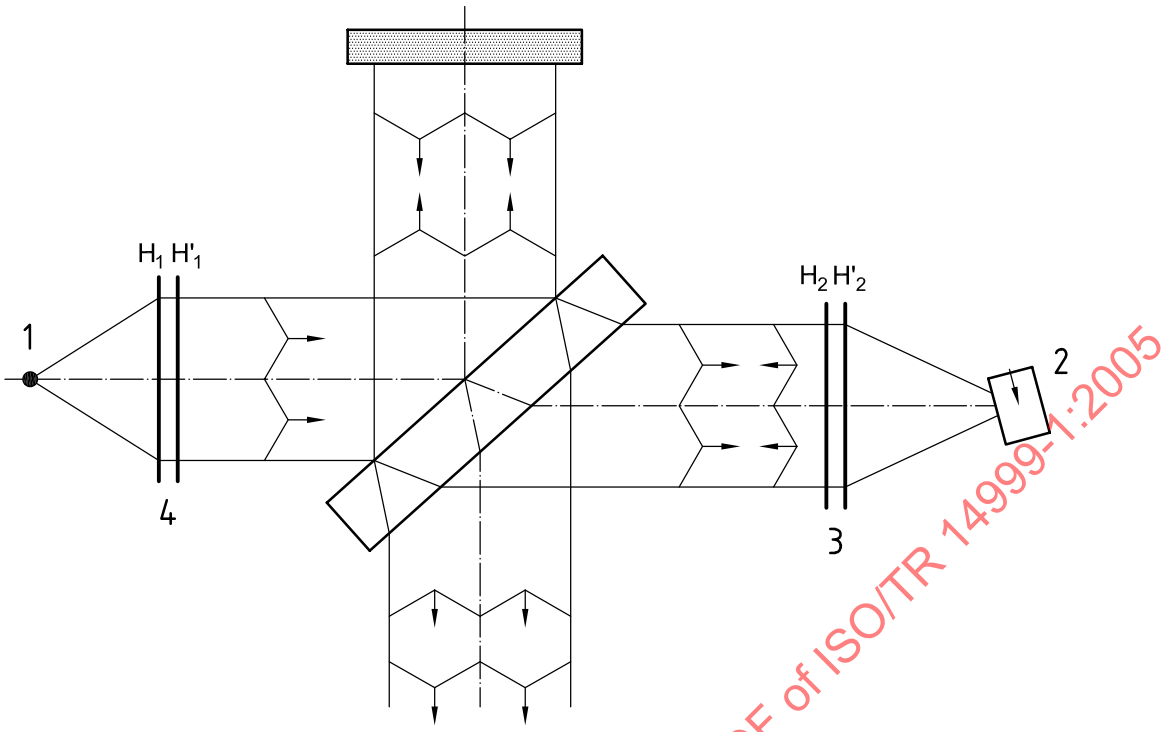
$$u = \frac{U}{r}\mathrm{e}^{\mathrm{i}kr} \quad \text{and} \quad u^* = \frac{U}{r}\mathrm{e}^{-\mathrm{i}kr} \tag{58}$$

If a uniform plane wave falls normally on a perfect plane mirror, then the reflected wave is identical in form to the incident wave but travelling in the reverse direction. The form is that of the incident wave, but for a reversal of the time coordinate.

A phase-conjugating interferometer using phase-conjugating mirrors eliminates the need for a perfect reference wavefront.

EXAMPLE A Twyman-Green interferometer as shown in Figure 23, using a phase-conjugating mirror, has been described by [19] and [20]. The phase-conjugating mirror is formed by a $BaTiO_3$ crystal, with the C-axis parallel to one of its edges and inclined 20° with respect to a plane perpendicular to the optical axis.

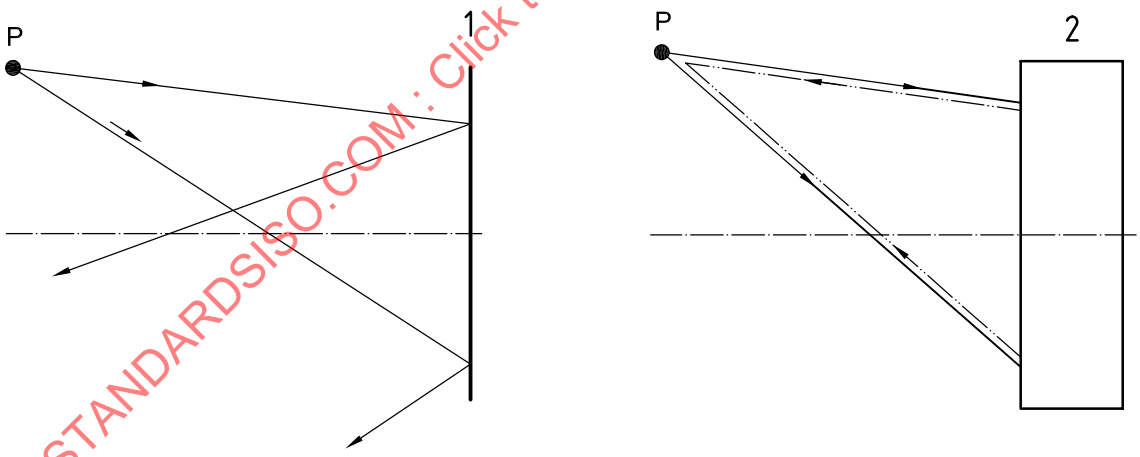
The phase conjugation is obtained by four waves mixing, which are automatically self-generated from a 30 mW argon laser with $\lambda = 514{,}5$ nm incident beam, by reflection at the crystal faces. Thus, it is a self-pumped phase-conjugating mirror. The wavefront incident on this phase-conjugating mirror is reflected back along the same ray direction that the incident wavefront has. Hence the wavefront deformations change sign as shown in Figure 24. The quality of the focusing lens $H_2H'_2$ is not important, because the returning rays have the same direction as the incident rays.



Key

1 light source
2 $BaTiO_3$ crystal
3 focusing lens
4 lens under test

Figure 23 — Twyman-Green interferometer with a phase-conjugating mirror



a) Reflection from a standard mirror **b) Reflection from a phase-conjugating mirror**

Key

1 standard mirror
2 phase-conjugating mirror
P point source

The reflection of a spherical wave from point source P by the standard mirror a) produces the divergent reflected wave whereas for the phase-conjugating mirror b) produces the convergent reflected wave, i.e. the wave from a point source P reflected from the mirror is the *time-reversed* form of the incident wave.

Figure 24 — Comparison of properties of a phase-conjugating mirror (PCM) with a standard mirror

The quality of the light source collimator is important. The wavefront distortions produced by this collimator will appear duplicated in the final interferogram, because the wavefront is not tested against a flat reference but against another wavefront, with deformations opposite in sign. The lens under test is the collimator, and the sensitivity is the same as that in the common Twyman-Green interferometer, but with only a single pass through the lens.

Advantage: No perfect lenses are necessary.

Disadvantage: An argon laser is required.

4 Coupled ray-paths in interferometers

4.1 Aperture stops and field stops; telecentric imaging

The term “field stop” is used in optics for stops, which are positioned in planes where also the object to be imaged or the images of the object are located. In the case of two-beam interference the object to be imaged is the wavefront under test. This might be the surface of a test piece or the pupil of an optical system. The field stop therefore defines the area of the object (entrance stop) or the area of the image (exit stop). Entrance stop and exit stop are images of each other. Only one needs to be a physical stop.

Pupils are images of the physically realized aperture stop, and are used to restrict the cone of rays, defining an angle β. By them, the amount of energy emitted from each element of the object or from the object as a whole transported through the optical system is defined. The entrance pupil is related to the object space, the exit pupil to the image space; entrance pupil and exit pupil are images of one another.

As the Abbe-Theory of imaging states, the ability to properly image fine details in an object is dependent on the angles β and β' of the cones of rays emerging from an object-point and impinging on an image-point. In order for an object detail of size d to be resolved as an image of size d', it is necessary that a cone of rays with a cone-angle β' emerge from d and be converted into the corresponding converging cone of rays with a cone-angle of β' impinging on d'. The quantitative relation for coherent imaging (see 4.3) is:

$$d = \frac{1{,}22\lambda}{\sin\left(\frac{\beta}{2}\right)} \text{ and } d' = \frac{1{,}22\lambda}{\sin\left(\frac{\beta'}{2}\right)} \tag{59}$$

The centre point of the pupils is the centre for the perspective. If the entrance pupil is located at infinity (the physical stop is then located in the back focal plane of the optical system), the object is imaged with parallel perspective. In this case, the size of the image is independent of the distance of the object from the optical system. If the exit pupil is located at infinity, the image size is independent of the distance of the image plane from the optical system. In this case, it is possible to shift the image plane for proper focusing without altering the image size.

If both pupils are located at infinity, both properties are realized at the same time. In that case, the imaging system is an afocal system, also named Kepler telescope. Here, two optical systems are combined in such a way, that they have a common “inner” focal point.

The physically realized aperture stop is located at that common focal point. If b is the diameter of that aperture stop, then the cone angles β and β' are related to the diameter b via the focal length f_1 and f_2 of the two lens-systems combined for the telescope. The relation is:

$$b = 2f_1 \sin\left(\frac{\beta}{2}\right) = 2f_2 \sin\left(\frac{\beta'}{2}\right) \tag{60}$$

Combining Equations (59) and (60) the ability of a Kepler telescope to image fine details can be expressed by the aperture size:

$$d = 2{,}44\lambda\frac{f_1}{b}$$

$$d' = 2{,}44\lambda\frac{f_2}{b} \tag{61}$$

A Kepler telescope is the ideal optical system for imaging the wavefront under test onto the detector in wavefront measuring interferometers. It is also a system, where no additional spherical wavefront terms are introduced when a plane wave at the front focal plane of the first system is transformed through the instrument to the back focal plane of the second system.

4.2 Coupled ray-path

A coupled ray-path is often found in optical instruments, which contain their own light source. In order to let the light flux pass through the optical instrument undisturbed, it is necessary to image the light source, in some cases more than once. The main purpose of the instrument might be imaging an object. For instance, the purpose of a microscope is the magnification of the image of a specimen. The principle of the optical layout with "coupled rays" is to position the stops for the definition of the object size and the cone of rays in such a way, that they do not influence each other. This is achieved, if the field stops for imaging the object are at the same time the aperture stops for imaging the light source and the images of the light source are located at the aperture stops.

When the image of a light source is smaller than the aperture stop for imaging an object, i.e. the pupil is only partially "filled" by the image of the light source, "partial coherent imaging" occurs. Coherent imaging, partial coherent imaging and incoherent imaging for instance produce different intensity distributions in the image plane, if a knife edge is the object.

4.3 Difference of coherent/incoherent optical imaging

Generally, the purpose of optical imaging is to transfer the given intensity distribution in the object plane to a similar distribution in the image plane. The transfer is accomplished by lenses, mirrors, prisms, etc.

In most imaging cases of daily life, such as with spectacles or photographic lenses, so-called incoherent imaging occurs. Here only "intensities" are important; a radiation density distribution in the object plane is transmitted into the image plane.

In the terms of the scalar theory of wave-optics, which then can also be applied to the case of coherent imaging, the following explains the relationships:

The light disturbance in the object may be $a(x)$ (for reasons of simplicity written only in one dimension) and the disturbance in the image may be $a'(x)$, where x is the reduced space-coordinate and therefore identical in object and image. $a(x)$ and $a'(x)$ both have amplitude and phase and therefore are written more precisely as:

$$a(x) = A(x)\mathrm{e}^{\mathrm{i}\varphi(x)} \quad \text{and} \quad a'(x) = A'(x)\mathrm{e}^{\mathrm{i}\varphi'(x)} \tag{62}$$

The irradiance or intensity distribution is proportional to $a(x)\ a^*(x)$ and the irradiation density distribution is proportional to $a'(x)\ a'^*(x)$, where * denotes the conjugate complex quantity. Obviously radiation distribution and irradiation distribution are only dependent on the amplitude distributions $A(x)$ and $A'(x)$; the phase distributions are not concerned, whichever values they may take. Therefore it is not important how the transformation of $\varphi(x)$ to $\varphi'(x)$ took place in incoherent imaging.

In the case of coherent imaging the transformation of the light disturbance $a(x)$ in the object to the light disturbance $a'(x)$ in the image should be considered. Here, amplitude *and* phase are important. In particular, it is the aim to transform the phase $\varphi(x)$ in the object to the phase $\varphi'(x)$ in the image in such a way, that they are