





# Experiment 10 Michelson Interferometer



יבקי

Laserzentrum FH Münster University of Appl. Sciences Dept. Engineering Physics - Prof. Dr. Klaus Dickmann -





1	FUN	DAMENTALS	3
	1.1	Characteristics of light	3
	1.2	Superimposition and Phase $\delta$	5
	1.3	Interferometer	6
	<b>1.4</b> 1.4.1 1.4.2	Real interferometer Contrast or visibility Light sources and coherence	<b>8</b> 8 8
	1.5	Coherence	9
	1.6	Thermal light	9
	1.7	Cold light	9
	1.8	Wave fronts	13
	1.9	Gaussian beams	14
	<b>1.10</b> 1.10.2 1.10.2 1.10.3 1.10.4 1.10.4	Technical interferometers1Beam splitting prism2Triple reflector3Homodyne interferometer4Detection optics5Heterodyne interferometer	<b>15</b> 16 17 17 18 19
2	EXP	ERIMENTS	21
	<b>2.1</b> 2.1.1 2.1.2 2.1.3 2.1.4 2.1.5	Michelson InterferometerAdjustment of the Michelson Interferometer Step 1Alignment Step 2Step 3, alignment of the mirrorsInterference patternsMeasuring the contrast function V(ΔL)	<b>21</b> 23 23 24 25 26
	<b>2.2</b> 2.2.1 2.2.2 2.2.3 2.2.4 2.2.5 2.2.6	Technical interferometer Set-up of Technical Interferometer step 1 Set-up step 2 Set-up step 3 Set-up step 4 Signal detection Length measurement	<b>27</b> 29 30 31 31 32
	2.3	PC-Card	32

#### 2.3 PC-Card

# 1 Fundamentals

### 1.1 Characteristics of light

Light, the giver of life, has always held a great fascination for human beings. It is therefore no coincidence that people have been trying to find out what light actually is for a very long time. We can see it, feel its warmth on our skin, but we cannot touch it. The ancient Greek philosophers thought light was an extremely fine kind of dust, originating in a source and covering the bodies it reached. They were convinced that light was made up of particles. As humankind progressed and we began to understand waves and radiation, it was proved that light did not, in fact, consist of particles but that it is an electromagnetic radiation with the same characteristics as radio waves. The only difference is in the wavelength. We now know that the characteristics of light are revealed to the observer depending on how he sets up his experiment. If the experimenter sets up a demonstration apparatus for particles, he will be able to determine the characteristics of light particles. If the apparatus is one used to show the characteristics of wavelengths, he will see light as a wave. The question we would like to be answered is: What is light in actual fact? The duality of light could only be understood using modern quantum mechanics. Heisenberg showed, with his famous "uncertainty relation", that strictly speaking, it is not possible to determine the place x and the impulse p of any given occurrence at the same time

$$\Delta \mathbf{x} \cdot \Delta \mathbf{p}_{\mathbf{x}} \ge \frac{1}{2}\hbar$$

If, for example, the experimenter chooses a set up to examine particle characteristics, he will have chosen a very small uncertainty of the impulse px. The uncertainty x will therefore have to be very large and no information will be given on the course of the occurrence. Uncertainties are not given by the measuring apparatus, but are of a basic nature. This means that light always has the particular quality the experimenter wants to measure. We can find out about any characteristic of light as soon as we think of it. Fortunately the results are the same, whether we work with particles or wavelengths, thanks to Einstein and his famous formula:

$$\mathbf{E} = \mathbf{m} \cdot \mathbf{c}^2 = \hbar \cdot \boldsymbol{\omega}$$

This equation states that the product of the mass m of a particle with the square of its speed c corresponds to its energy E. It also corresponds to the product of Planck's constant  $h = \hbar \cdot 2\pi$  and its radian frequency  $\omega = 2\pi \cdot \nu$ . In this case v represents the frequency of luminous radiation.

In our further observations of the fundamentals of the Michelson interferometer, we will use the wave representation and describe light as electromagnetic radiation. All types of this radiation, whether in the form of radio waves, X-ray waves or light waves consist of a combination of an

electrical field E and a magnetic field H. Both fields are bound together and are indivisible. Maxwell formulated this observation in one of his four equations, which describe electromagnetic fields

$$\nabla \times \vec{H} \approx \frac{\partial \vec{E}}{\partial t}$$

According to this equation, every temporal change in an electrical field is connected to a magnetic field (Fig.1).



Fig. 1: Light as electromagnetic radiation

Due to the symmetry of this equation, a physical condition can be sufficiently described using either the electrical or the magnetic field. A description using the electrical field is preferred since the corresponding magnetic field can then be obtained by temporal derivation. In the experiments (as presented here) where light is used as electromagnetic radiation, it is advantageous to calculate only the electrical fields since the light intensity is:

$$\mathbf{I} = \frac{\mathbf{c} \cdot \mathbf{\varepsilon}}{4\pi} \cdot \left| \vec{\mathbf{E}} \right|^2.$$

This is also the measurable size as perceived by the eye or by a detector. In this case, the speed of light is c in the respective medium and  $\varepsilon$  is the corresponding dielectrically constant. Since we are comparing intensities in the same medium, it is sufficient to use



### Fig. 2: In this experiment we need only observe the electrical field strength E

The experimental findings agree with the theory of electromagnetic radiation if a harmonic periodic function becomes temporally dependent on the field strength of light. In its simplest form this is a sine or cosine function. An amplitude  $E_o$  and a wavelength  $\lambda$  should be used in the definition of this kind of function. Let us begin with the equation:

$$\mathbf{E}_{\mathbf{X}} = \mathbf{E}_0 \cdot \sin\left(\frac{2 \cdot \pi}{\lambda} \cdot \mathbf{x}\right), \qquad (1.1.1)$$

which we will elaborate and explain further.



Fig.3: Amplitude and wavelength

In the above figure the light wave no longer oscillates in the Z-direction as in Fig.2 but at a certain angle to the Z- or Y-axis. The X-axis has been chosen as the direction of propagation of the wave. We still require information on the direction in which the electrical field strength  $E_x$  oscillates to complete the description of the wave. Strictly speaking, the field  $E_x$  oscillates vertically to the direction of propagation X. However, we have to give information regarding the Z-and Y-axis. This leads to the term 'Polarisation' and Direction of Polarisation. In Figs. 1 and 2 we used linearly polarised light with a polarisation direction in Z and in Fig.3 we used a different direction. We will now introduce the polarisation vector P, which is defined in the following Fig.4. We have to look into the light wave in the direction of the X-axis for this purpose.



### Fig.4: Definition of the polarisation vector

We can observe a wave expanding in the X-direction and oscillating at the electrical field amplitude  $E_o$  below an angle  $\alpha$  to the Y-axis. The amplitude  $E_o$  is separated into its components, which oscillate in the Z- or Y-direction. We now write  $\hat{E}_0$  instead of  $E_o$  to indicate that the amplitude  $E_o$  is now made up of individual components.

$$\hat{\mathbf{E}}_{0}^{2} = (\mathbf{E}_{0}^{Y})^{2} + (\mathbf{E}_{0}^{Z})^{2} = (\mathbf{E}_{0}^{Y} \cdot \hat{\mathbf{e}}_{Y} + \mathbf{E}_{0}^{Z} \cdot \hat{\mathbf{e}}_{Z})^{2}$$

In this case  $\hat{e}_{Z} = (0,1), \hat{e}_{Y} = (1,0)$  is the unit vector in the Z- or Y-direction on the ZY-plane. Characteristically the unit vectors yield  $|\hat{e}_{Z}| = 1$  and the scalar product  $\hat{e}_{Z} \cdot \hat{e}_{Y} = 0$ . The equation (1.1.1) can now be generalised to:

$$\mathbf{E}_{\mathbf{X}}(\mathbf{Y}, \mathbf{Z}) = \left(\mathbf{E}_{0}^{\mathbf{Y}} \cdot \hat{\mathbf{e}}_{\mathbf{Y}} + \mathbf{E}_{0}^{\mathbf{Z}} \cdot \hat{\mathbf{e}}_{\mathbf{Z}}\right) \cdot \sin\left(\frac{2\pi}{\lambda} \cdot \mathbf{x}\right)$$

At this point we come across a fundamental principle in classic wave theory, i.e. the principle of superimposition. A big word for the simple statement:

Every wave can be represented as the sum of individual waves.

In our example we had separated the wave as shown in Fig.4 into two individual waves, i.e. one that oscillates in the Z-direction and another in the Y-direction. We could just as well say that our wave was formed by the superimposition of these two individual waves. The word *interference* can also be used to mean superimposition. In this context our wave was formed by the interference of two individual waves. This is the basis for the functioning of the Michelson interferometer. An introduction to this interferometer now follows. For the time being, let us return to the polarisation vector.

The polarisation vector P is also a unit vector, which always points in the direction of the oscillation of the electrical field  $E_{\rm x}$ 

$$\hat{\mathbf{P}} = \frac{\hat{\mathbf{E}}_{0}}{\left|\hat{\mathbf{E}}_{0}\right|} = \frac{\mathbf{E}_{0}^{Y}}{\mathbf{E}_{0}} \cdot \hat{\mathbf{e}}_{Y} + \frac{\mathbf{E}_{0}^{Z}}{\mathbf{E}_{0}} \cdot \hat{\mathbf{e}}_{Z}$$

or as is written for vectors

Ê

$$\hat{\mathbf{P}} = \left(\frac{\mathbf{E}_0^{\mathrm{Y}}}{\mathbf{E}_0}, \frac{\mathbf{E}_0^{\mathrm{Z}}}{\mathbf{E}_0}\right).$$

The polarisation vector for a polarisation in the Z-direction  $(0^{\circ})$  would then be, for example:

$$\hat{P} = (0, 1)$$

for a polarisation direction of 45° it would be:

$$\hat{\mathbf{P}} = \frac{1}{\sqrt{2}} (1, 1).$$

The equation of the wave with any given polarisation direction will thus be

$$\hat{\mathbf{E}}_{\mathbf{X}}(\mathbf{Y}, \mathbf{Z}) = \hat{\mathbf{P}} \cdot \mathbf{E}_{0} \cdot \sin\left(\frac{2\pi}{\lambda} \cdot \mathbf{x}\right), \text{ or}$$

$$\mathbf{x}(\mathbf{Y}, \mathbf{Z}) = \left(\mathbf{E}_{\mathbf{Y}}, \mathbf{E}_{\mathbf{Z}}\right) \cdot \sin\left(\mathbf{k} \cdot \mathbf{x}\right). \quad (1.1.2)$$

We have introduced the wave number **k** in the above equation

$$k = \frac{2\pi}{\lambda}.$$

The wave number k has the length dimension<sup>-1</sup> and was originally introduced by spectroscopists because it was a size that could be measured immediately with their equipment. We are using this size because it simplifies the written work.

Till now we have only described the wave as a function of the location. This would be sufficient in order to understand the classical Michelson interferometer, but not for technical interferometers. We carry out the following hypothetical experiment to introduce the "time coordinate":



Fig. 5: Hypothetical experiment for the introduction of dependency on time

The talented physicist Walter S. \* (\*Names have been changed) is working on new experiments with electromagnetic waves in his laboratory. His colleague Gerd N. who is jealous of his rival's success sees that the door to Walter S.' laboratory has been left open a crack and uses the opportunity to find out what his colleague is working on. In spite of his nervousness, Gerd N. forces himself to make painstakingly accurate notes of his observations. He measures the time with his Swiss stop watch, a present from his father, who was also a physicist and notes the respective intensities of what he sees through the crack in the door. He rushes into his modest study, red in the face, and writes his observations into his laboratory records. Here we find the following entries:

".....I stood at location X and looked through the crack in the door. I observed periodic oscillations in field strength, which fluctuated between a maximum and a minimum. I began measuring at the time t=0, when the field strength was at its minimum. At the time t<sub>1</sub> I calculated the maximum field strength. The differences in time between the extreme values stayed constant." A graph of his measurement values follows:



Fig. 6: Gerd N's measurement curve

Gerd N. further states that:"...... the time that passed between two maxima as the duration period  $\tau$ . I have observed n of such maxima within one second. Obviously the field

strength has a frequency of  $v = \frac{n}{\tau}$  and follows a periodic

function  $E = E_0 \cdot \sin(2\pi \cdot \nu \cdot t)$  although this function begins with positive values. The measured values only correspond to the observations if a constant is added to the argument of the sine

$$\mathbf{E} = \mathbf{E}_0 \cdot \sin\left(2\pi \cdot \mathbf{v} \cdot (\mathbf{t} - \mathbf{t}_0)\right)$$

Some weeks later both colleagues meet at a specialists' conference. As is often the case, the evening session of the conference took place in a suitable atmosphere, where the participants committed themselves to the team spirit over a glass of wine and agreed on all other things as well. Walter S. spoke openly about how he had managed to formulate the position of the course of a wave and wrote his formula on the beer mat, commonly used in this area

$$\mathbf{E} = \mathbf{E}_0 \cdot \sin\left(\frac{2\pi}{\lambda} \cdot (\mathbf{x} - \mathbf{x}_0)\right)$$

How the evening finally ended is left to your own imagination. What is important is that both experimentalists measured the same field strength, one with a stopwatch in his hand, the other with a scale. Therefore

$$\frac{\mathrm{E}}{\mathrm{E}_{0}} = \sin\left(2\pi \cdot \nu \cdot (t-t_{0})\right) = \sin\left(\frac{2\pi}{\lambda} \cdot (x-x_{0})\right).$$

This is the same as

or

$$\mathbf{v}\cdot\boldsymbol{\lambda} = \frac{\left(\mathbf{x} - \mathbf{x}_{0}\right)}{\left(\mathbf{t} - \mathbf{t}_{0}\right)} = \mathbf{c}.$$

 $\mathbf{v} \cdot (\mathbf{t} - \mathbf{t}_0) = \frac{1}{\lambda} \cdot (\mathbf{x} - \mathbf{x}_0)$ 

This hypothetical experiment has shown us that we can describe the wave by its temporal course on the one hand, and by the position of the course of a wave on the other. We have also found out the importance of the relationship of the

speed c of a wave to its frequency and wavelength  $v = \frac{c}{\lambda}$ .

If we write the connection with  $\omega = 2\pi v$  as a rotary frequency we get:

$$\omega = \frac{2\pi}{\lambda} \cdot \mathbf{c} = \mathbf{k} \cdot \mathbf{c} \, .$$

Let us now return to the generalised formula for the temporal and spatial course of the field strength of a light wave. Since the sine is a periodic function we can include the temporal and spatial dependency into the argument. We would then get

$$\hat{\mathrm{E}}_{\mathrm{X}}(\mathrm{Y},\mathrm{Z}) = \left(\mathrm{E}_{\mathrm{Y},\mathrm{E}_{\mathrm{Z}}}\right) \cdot \sin\left(\mathrm{k}\cdot\left(\mathrm{x}-\mathrm{x}_{0}\right) + \omega\cdot\left(\mathrm{t}-\mathrm{t}_{0}\right)\right)$$

If we now make the constants  $kx_0$  and  $\omega t_0$  into one constant  $\delta$  we obtain the general formula

$$\hat{\mathbf{E}}_{\mathbf{X}}(\mathbf{Y}, \mathbf{Z}) = \left(\mathbf{E}_{\mathbf{Y},} \mathbf{E}_{\mathbf{Z}}\right) \cdot \sin\left(\mathbf{k} \cdot \mathbf{x} + \boldsymbol{\omega} \cdot \mathbf{t} + \boldsymbol{\delta}\right) \quad \textbf{(1.1.3)}$$

### 1.2 Superimposition and Phase $\delta$

 $\delta$  is also described as a phase. Since this term is often inconvenient we would like to examine it more closely. If we put x=0 and t=0 into (1.1.3) the field strength will have a value of  $\hat{E}_x(Y,Z) = (E_{Y,}E_Z) \cdot \sin(\delta)$  and thus defines an initial value for the amplitude. This value is or will be determined according to the physical situation.



Fig.7: Definition of phase δ

Obviously phase  $\delta$  contains information about the relationship between two or more waves. Let us presume that the waves originate in a light source and phase  $\delta$  contains information on how the wave was formed. Light waves are formed by emission processes. There is an emission procedure for every photon or light wave. Such processes are always taking place when the light source is continuously illuminated. The emission procedures are distributed statistically according to the type of light source. Thus, phase  $\delta$  is also distributed statistically. If the emission procedures are coupled to each other, as is the case with lasers, and all photons or waves have the same frequency or wavelength (they are monochromatic) the light is then described as coherent (holding itself together). If, however, phase  $\delta$  is randomly distributed, then this light is incoherent. This is the case with thermal light sources, e.g. light bulbs. To judge the coherence of a light source, the characteristics of the emitted waves and/or photons would have to be classified. The waves (photons) are first sorted out according to their frequencies (wavelengths) and then according to their phases. If we form small "containers" in our minds with the labels:

$$k = \frac{2\pi}{\lambda} = numerical value and \delta = numerical value$$

and if we now sort out the photons in these containers and then count the photons per container, we could obtain a statement on the coherence. This kind of container is also called a phase cell. If all photons were in one container or phase cell, the light would be completely coherent.

In the example according to Fig.7 the wave 2 has a phase of  $\delta$  as opposed to the wave 1, in other words, the waves have a phase difference of  $\delta$ , presuming that we have produced two such waves (this is exactly what the Michelson interferometer does). We expect a third wave through the principle of superimposition, which is formed by the superimposition or interference of the two basic waves. We will find out how this wave looks by simply adding both basic waves:

Wave 1  $\hat{E}_1(Y, Z) = (E_{Y,}E_Z) \cdot \sin(k \cdot x + \omega \cdot t + \delta)$ 

Wave 2 
$$\hat{E}_{2}(Y, Z) = (E_{Y,}E_{Z}) \cdot \sin(k \cdot x + \omega \cdot t)$$

Wave 3 
$$\hat{E}_{3}(Y,Z) = \hat{E}_{1}(Y,Z) + \hat{E}_{2}(Y,Z)$$

It is now easy to imagine that a large number of waves with different frequencies  $\omega$  or wavelengths  $\lambda$  and phases  $\delta$  result in such a mixture and that it makes little sense to carry out superimposition or interference experiments with this light. Therefore light sources which emit light within a narrow emission spectrum with a phase as constant as possible are selected. Lasers are an example of such light sources. But when Michelson carried out his experiments around 1870 he could not use lasers. He used the red emission line of a cadmium lamp whose emission bandwidth showed a coherence length of only 20 cm. This means that when, for example, waves at the position x=0 were superimposed with those at the position x=20, there was no readable interference any more. We will come back to the important term "coherence length" later on and discuss it in more detail.

### 1.3 Interferometer

An apparatus that produces this physical condition is shown in the following diagram.





We beam light into the entrance of the interferometer from some light source. The light is split into two bundles on a beam-splitting plate S. One bundle hits the mirror M1 and the other the mirror M2. The bundles will reflect back in themselves at these mirrors and reunite at the beam-splitting plate S. The respective bundles are split into two further bundles due to the characteristics of the beam-splitting plate and one bundle is led in the direction of exit 1, the other in the direction of exit 2 in the process. Exit 2 of the Michelson interferometer points in the direction of the light source, so this exit is practically of no use to us to set up photodetectors or imaging screens. This is why only exit 1 will be mentioned. However, exit 2 must also be taken into consideration for the energy balance. Whatever technical model of an interferometer is chosen, it can be represented easily in an optically circuit diagram.



Fig.9:"Optical circuit diagram" of an interferometer

The light, which is put into the entrance of the interferometer, is split into two bundles. How this happens technically is not important for the time being. This kind of element can generally be called a directional coupler. Bundle 1, or, in the simplest case the wave 1 runs through the path L1 and the other wave 2 runs through the path L2. Both waves are brought together in a mixer. This mixer has two exits. In the Michelson interferometer directional couplers and mixers are one and the same element. We are only interested in exit 1 for the time being. As we will see later on, exit 2 is symmetrical to exit 1. We now just have to calculate wave 3 at exit 1 which is formed out of the superimposition of wave 1 with wave 2 which have travelled along a path from L1 or L2. Without jeopardizing the general validity of the solution, we can assume that the electric field strength only oscillates in the Y-direction. As already defined in the beginning the waves propagate in the X-direction. Although the direction of the bundle of rays can point in other directions after separation, they should have the same direction of propagation, at least in the mixer, if they are to interfere at all.

Wave 1 
$$E_1 = E_{01} \cdot \sin(k \cdot L_1 + \omega \cdot t)$$
  
Wave 2  $E_2 = E_{02} \cdot \sin(k \cdot L_2 + \omega \cdot t)$   
Wave 3  $E_3 = E_1 + E_2$ 

Since both waves  $E_1$  and  $E_2$  are formed when the entering wave  $E_0$  is split and the splitter should separate them symmetrically, both partial waves do not have any phase shift  $\delta$  with regard to each other.

$$\mathbf{E}_{3} = \mathbf{E}_{01} \cdot \sin\left(\mathbf{k} \cdot \mathbf{L}_{1} + \boldsymbol{\omega} \cdot \mathbf{t}\right) + \mathbf{E}_{02} \cdot \sin\left(\mathbf{k} \cdot \mathbf{L}_{2} + \boldsymbol{\omega} \cdot \mathbf{t}\right)$$

A screen or photodetector is installed at exit 1. The human eye and the photodetector are not in a position to register electric field intensities, but can only register the light intensity I which is connected to the field strength:

$$I = E^{2}$$

$$I = (E_{01} \cdot \sin(k \cdot L_{1} + \omega \cdot t) + E_{02} \cdot \sin(k \cdot L_{2} + \omega \cdot t))^{2}$$

$$I = E_{01}^{2} \cdot \sin^{2}(kL_{1} + \omega t)$$

$$+ 2E_{01}E_{02} \sin(kL_{1} + \omega t)\sin(kL_{2} + \omega t)$$

$$+ E_{02}^{2} \cdot \sin^{2}(kL_{2} + \omega t)$$

To simplify the mixed term we use the relation:

$$2 \cdot \sin \alpha \cdot \sin \beta = \cos (\alpha - \beta) + \cos (\alpha + \beta)$$

and obtain:

$$I = E_{01}^{2} \cdot \sin^{2} (kL_{1} + \omega t)$$
  
+  $E_{01}E_{02} \cos(k(L_{1} - L_{2}))$   
+  $E_{01}E_{02} \cos(k(L_{1} + L_{2}) + 2\omega t)$   
+  $E_{02}^{2} \cdot \sin^{2} (kL_{2} + \omega t)$ 

The expression for light intensity, which is perceived either by a detector or by our own eyes, consists of four terms. Only the second term is not dependent on the time t. All other terms oscillate with the frequency  $\omega$ . We use

$$\omega = 2\pi\nu = 2\pi\frac{c}{\lambda}$$

to determine this  $\omega$ . The frequency of light is v and has a wavelength  $\lambda$  and a speed c. In the later experiments we will select the light emitted by a Helium-Neon laser. The wavelength of this light is 633mm. Using this value and the speed of light  $c = 3 \cdot 10^8$  m/s the frequency v is calculated as:

$$v = \frac{3 \cdot 10^8}{633 \cdot 10^{-9}} = 4, 7 \cdot 10^{14} \,\mathrm{Hz}.$$

The sine of the first and last term oscillates at this frequency and the third even oscillates at double this frequency. Neither the eye nor any photodetector is capable of following this extremely high frequency. The fastest photodetectors nowadays can follow frequencies of up to approx.  $2 \cdot 10^9$  Hz. This is why a detector, and even more so, our eyes can only perceive average values. The sin<sup>2</sup> terms oscillate between 0 and 1; their temporal average value is therefore1/2. The cosine term oscillates between -1 and +1, the average value is zero. The intensity *I* would therefore be:

$$I = \frac{1}{2}E_{01}^{2} + \frac{1}{2}E_{02}^{2} + E_{01}E_{02}\cos(k \cdot \Delta L)$$
$$\Delta L = L_{1} - L_{2}.$$

Obviously *I* is maximum if the cosine is one. This is always the case when its argument is zero or an integral multiple of  $2\pi$ . *I* is minimum just at the time when the cosine is -1

$$I_{max} = \frac{1}{2}E_{01}^{2} + \frac{1}{2}E_{02}^{2} + E_{01}E_{02} = \frac{1}{2}(E_{01} + E_{02})^{2}$$
$$I_{min} = \frac{1}{2}E_{01}^{2} + \frac{1}{2}E_{02}^{2} - E_{01}E_{02} = \frac{1}{2}(E_{01} - E_{02})^{2}$$

Let us remind ourselves that

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

and that it is constant at a stable wavelength. The light intensity at exit 1 is therefore obviously only dependent on the path difference L1-L2. If both paths having the same length, both partial waves interfere constructively and the

light intensity observed is maximum. If the path difference is just  $\lambda/2$  then:

$$\mathbf{k} \cdot \Delta \mathbf{L} = \frac{2\pi}{\lambda} \cdot \frac{\lambda}{2} = \pi$$

The cosine is then -1 and the light intensity *I* at the exit becomes minimum. Let us divide the initial intensity into two partial ones of equal size, i.e.  $E_{01}$  and  $E_{02}$ . In this case even the light intensity is zero. Here, both partial waves interfere destructively. Keeping in mind that the wavelength for our experiment is 633 nm and that it leads to a shift from one wave to another by only  $\lambda/2=316.5$  nm = 0.000000316 mm (!) from a light to a dark transition at exit 1, this type of interferometer is a highly precise apparatus for measuring length.



#### Fig.10:Interferogram

#### 1.4 Real interferometer

In the preceding chapter, we assumed that we had ideal light sources and optical components at our disposal. Practical interferometry is, in fact, characterised by these components. Michelson (1871) did not construct his interferometer to examine the interference of light, but to prove the 'world ether' theory. It was therefore an optical instrument that underwent constant development. Since the interference fringes had to be judged with the naked eye, it was the aim of this further development to maximise the contrast in light and dark shades.

### 1.4.1 Contrast or visibility

Michelson defined as a scale for the measurement of contrast:

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

Historically, the word visibility or clarity was also chosen for contrast.



#### **Fig.11: Signal at exit 1 at** $I_{01}$ = 0.2 *Io* **and** $I_{02}$ = 0.8 *Io*

If, for example, the splitting relationship of the beam splitter is not exactly 1 then the interferogram will change as shown in Fig.11.

Using

$$I_{max} = \frac{1}{2}I_{01} + \frac{1}{2}I_{02} + \sqrt{I_{01} \cdot I_{02}}$$

and

$$\mathbf{I}_{\min} = \frac{1}{2} \mathbf{I}_{01} + \frac{1}{2} \mathbf{I}_{02} - \sqrt{\mathbf{I}_{01} \cdot \mathbf{I}_{02}}$$

we obtain

$$V = \frac{2 \cdot \sqrt{I_{01} \cdot I_{02}}}{I_{01} + I_{02}} = \frac{2 \cdot \sqrt{I_{01} \cdot I_{02}}}{I_0}$$

for the contrast. According to Fig.10 the contrast is V=1 and in Fig.11 V=0.8. A contrast reduction also occurs if, for example the adjustment is not optimal, i.e. the overlap of both the interfering rays is not 100 %. In this case





#### 1.4.2 Light sources and coherence

Till now we have tacitly assumed that the light source only has a sharp frequency  $\omega$ . This is never the case in practice. We would therefore like to examine the influence of the spectral emission of the light source on interference formation. Radiation that gets its energy from the warmth of a radiating body is known as thermal radiation and is sent out by glowing solids (metal, coal etc.) or gases at high pressures. But radiation can also be produced without extracting energy from the warmth of a medium, e.g. by introducing electrical energy (discharge in gas). In this type of radiation

production, the temperature of the radiator (except in loss mechanisms) does not change. The light produced in this way is known as "cold light" or also luminous radiation. The type of emission depends on the excited electronic or vibronic states of the atoms and molecules concerned. Light sources that emit light based on these mechanisms are an important component of modern light technology. Whereas light bulbs are thermal radiators, fluorescent lamps, gas discharge lamps, monitors as well as lasers are sources of cold light.

### 1.5 Coherence

In classic optics, light was described as coherent if it could produce interferences. The term coherence was elaborated by the photon statistics introduced by quantum optics. It says, that light is coherent when photons originate in the same phase cell, that is; they have the same frequency and phase within Heisenberg's uncertainty principle. Monochromatic light is only coherent if the partial waves also have the same phases.



# Fig.13: Light, whose photons originate from the smallest possible phase cell according to Heisenberg's uncertainty principle, is coherent.

Classic optics states that light is coherent when it shows signs of interference. The narrower the emission line width  $\Delta \omega$  is the higher will be the contrast observed in this process. There is obviously a connection between the contrast or visibility V of the interferogram and the line width of the light source used.



# Fig.14: Light with a narrow band emission with emission bandwidth $\Delta \omega$ .

If we follow the views of classic optics in the next chapters, this is because we will not touch the limits of Heisenberg's uncertainty principle by far, in spite of using a laser as the light source for the interferometer. The emission bandwidth and the emitted line width of the technical HeNe laser used is very big compared to the theoretical natural bandwidth of a line.

### 1.6 Thermal light

We know from our daily observations that a hot body transmits light radiation. The higher the temperature of the body, the whiter the light seems.



### Fig.15: The spectral distribution of a black body radiator (thermal radiator, Planck's radiation law) for various temperatures

It can thus be concluded that this kind of radiator has a very large emission bandwidth  $\Delta \omega$ . Therefore thermal light sources are not suitable for interferometry.

### 1.7 Cold light

Another type of light production is the light emission of atoms and molecules, which show a clear spectral structure

characteristic of certain atoms, and molecules, contrary to the continuous radiation of the temperature radiators. Apart from their characteristics as a light source, the main interest in spectra was to use them to find out more about the structure of atoms and molecules.



#### Fig.16: Bohr's idea of the atom

Einstein elaborated Bohr's' theories on the atom model to the extent that light emitted or absorbed from the atoms can only have energies of  $E_2 \cdot E_1 = hv$  (Fig.16). His work and measurements proved that both in a resonant cavity and in the case of atomic emission discrete energies must be assumed. Einstein started off by trying to find one single description for these two types of light sources. He found the solution to this problem in 1917 when he derived Planck's hypothesis once more in his own way. Photons or a radiation field are produced through the transition from  $2 \rightarrow 1$ . With regard to what we have already considered, the frequency of the radiation produced in this way should actually be determined by  $E_2$ - $E_1 = h v_0$ . However, keeping in tune with Heisenberg's uncertainty principle

$$\Delta \mathbf{x} \cdot \Delta \mathbf{p}_{\mathbf{x}} \ge \frac{1}{2}\hbar \text{ or } \Delta \mathbf{E} \cdot \Delta \mathbf{t} \ge \frac{1}{2}\hbar$$

which states that every emission process is affected by an uncertainty and the emission line therefore always has a certain width  $\Delta \omega$ .



### Fig. 17: Natural line width

$$\rho(\omega) = \frac{1}{\left(\omega - \omega_0\right)^2 + \left(1/2 \cdot \tau_s\right)^2}$$

This type of curve represents a Lorentz curve. Here,  $\omega_0$  is the resonant frequency and

$$\tau_{s} = \frac{1}{A_{21}}$$

is the average life of state 2. The half-width value  $\Delta \omega$  (or also FWHM, Full Width at Half Maximum) of the curve according to Fig.17 is calculated by inserting the value for  $\rho(\omega) = 1/2$ . We find that

$$\delta(\omega)_{\text{not}} = A_{21} \tag{1.7.1}$$

is the natural line width of a transition, defined by the Einstein coefficient  $A_{21}$ , which has a particular value for every transition. The results gained can also be interpreted to mean that state 2 does not have any sharply defined energy but an expansion with a half-width value of  $\Delta E = 2 \pi h A_{21}$ . There would therefore be some uncertainty in this state. Quantum mechanics has shown that this effect is of fundamental importance. It has been called the Heisenberg's uncertainty relation after its inventor. In normal optical transitions, the value of  $\tau_s$  is between 10<sup>-8</sup> and 10<sup>-9</sup> sec. This lifetime, determined by the spontaneous transitions alone is a decisive factor for the so-called natural half-width value of a spectral line. Regarding descriptions, it must be stressed at this point that we must differentiate between the width of a state and the width of a line as well as between the terms state and line. Atomic states always exist, so no statement will be made on whether the state is occupied or empty. A line can only be formed if, for example, an emission occurs by the transition from the state from to  $2\rightarrow 1$ . The line is a word used by spectroscopists. They may produce, for example, photographic plates using their spectral apparatus, on which fluorescent light is shown divided up into its wavelengths. They use slits in the optical radiation path for an easy evaluation of the spectra. Fig 18 shows this kind of line spectrum.



### Fig.18: Photograph of the emission of a light source with the corresponding energy level showing a socalled line spectrum.

The spectrum according to Fig.18 can be seen from the line widths as well as from the emission wavelengths. Please note, however, that the measurement apparatus makes the line widths appear larger than they actually are.

If the spectral distribution of a line radiator were carried into that of the thermal radiator, the result would be a narrower line than the one in Fig.15.

Let us return to the influence of the spectral bandwidth  $\Delta \omega$  on the contrast of our interferometer. To make it easier we have selected the intensities as

$$I_{01} = I_{02} = \frac{1}{2} \cdot I_0$$

and obtain for the intensity of the interfering waves:

$$\mathbf{I} = \mathbf{I}_0 \left\{ 1 + \cos\left(\mathbf{k} \cdot \Delta \mathbf{L}\right) \right\}.$$

However, now the intensity  $I_{\theta}$  is a function of  $\omega$ :  $I_{0} \rightarrow \hat{I}(\omega) = \hat{I} \cdot \rho(\omega) \cdot \partial \omega$ 

So, we must integrate over all  $\omega$ :

$$\mathbf{I} = \int_{-\infty}^{+\infty} \left\{ \hat{\mathbf{I}} \cdot \boldsymbol{\rho}(\boldsymbol{\omega}) + \hat{\mathbf{I}} \cdot \boldsymbol{\rho}(\boldsymbol{\omega}) \cdot \cos(\boldsymbol{\omega} / \mathbf{c} \cdot \Delta \mathbf{L}) \right\} \cdot \partial \boldsymbol{\omega}$$

and obtain for the interferogram

$$\mathbf{I} = \mathbf{I}_{0} + \int_{-\infty}^{\infty} \hat{\mathbf{I}} \cdot \boldsymbol{\rho}(\boldsymbol{\omega}) \cdot \cos(\boldsymbol{\omega} / \mathbf{c} \cdot \Delta \mathbf{L}) \cdot \partial \boldsymbol{\omega} \quad (1.7.2)$$

We can obtain the contrast or the visibility of the interference from the extreme values of (1.7.2):

$$V(\Delta L) = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

Thus the contrast function V, which is still dependent on  $\Delta L$  after the integration, is the envelope of the interferogram I( $\Delta L$ ). According to the form the emission line takes, or how the function  $\rho(\omega)$  is made up, we can obtain the contrast of the function of the path difference  $\Delta L$  and the corresponding contrast function V( $\Delta L$ ) by using the above integration. The emission line of the laser, which we will use in the later experiment, can be described using a Gaussian function with a bandwidth  $\Delta \omega$ . The integral according to (1.7.2) can be calculated in this way. Counting over is carried out in [1] and the contrast function is shown in the following diagram.



Fig.19: Contrast as a function of the path difference of a light source with an emission bandwidth of 5 or 10 MHz (Laser).

The diagram on the right hand side of Fig.19 shows the contrast function. It shows that the contrast in a Gaussian distribution curve gradually drops as the path difference between both interfering waves increases. Theoretically, if the path difference is infinitely long, the contrast will eventually be zero. However this is not a practical answer. Although technical interferometers are being used for a long time there is no information on a marginal contrast, which merits an evaluation. We have therefore taken the liberty of defining the marginal contrast ourselves. We would say that if the contrast V has dropped to

$$\frac{1}{e^2} \cdot V_{max} \approx 0.14$$

the practical coherence length  $L_e$  has been reached. Since these associations are important, we should now examine the practical conditions in more detail in anticipation of the experiments that are still to come. In the course of our studies, we had taken a partial ray from a ray of light at a location situated at any given distance from a light source. We will give the coordinate  $z_1$  to this location. Another partial ray is taken from a second location with the coordinate  $z_2$ . Both the partial rays are joined together and the intensity of the superimposed ray is measured.

### Light source



#### Fig.20: Definition of coherence time

The photons taken from location  $z_2$  have to travel a longer path than the photons at location  $z_1$ . Since all photons fly at the speed of light c, the photons at location  $z_2$  must have been "born" earlier than those at  $z_1$ . The time difference between both photon generations is obviously

$$\Delta t = \frac{\Delta L}{c}$$

If we choose a path difference  $\Delta L$ , that is large enough to cause a drop in contrast to the value we defined, then the corresponding running time  $\Delta t$  will be known as the coherence time. We have introduced this term for didactic reasons. It is not required in our further observations. The construction according to Fig.20 has a considerable disadvantage in that it cannot be carried out practically. The simplest basic form of a regulatory device for interference is the Michelson interferometer. Although the physical situation we have just discussed is the same, the path difference must be observed more closely.



### Fig.21: Calculating the path difference △L using the Michelson interferometer

The beam splitter S splits the incoming beam. This is the origin of our system of coordinates. The mirror M1 should always be static, whereas the mirror M2 should be movable. Thus the two radiation paths form two arms of the interferometer. The one with the static mirror is known as the reference arm and the one with the movable mirror is the index arm. The partial beam in the reference arm travels a path of 2  $L_1$  before it returns to the reunification point at the beam splitter S. The path  $2L_2$  is crossed in the reference arm. The path difference is therefore:

$$\Delta L = 2 \cdot (L_2 - L_1)$$

We assume that the thickness of the beam splitter plate is zero. If  $L_2 = L_1$ , the path difference is zero. This can be adjusted for any number of intervals between the mirrors. In this state, the Michelson interferometer is symmetrical and even white light produces an interference pattern. This type of interferometer is sometimes also known as a white light interferometer. To find out the coherence length of the light source, the path difference is increased by shifting the mirror M2. There should be no maladjustment while doing this, since this will also reduce the contrast. If we measure this contrast as a function of the path difference, we get the contrast function. The point, at which the contrast is below the value already given by us, is the point where the coherence length has been reached. The following interference patterns would probably be seen on a screen.



# Fig.22: Decreasing contrast with an increase in path difference.

Michelson carried out his experiments with the red line of a cadmium lamp that had a coherence length of only 20 cm. Since the index arm is crossed twice he had a measurement area of only 10 cm at his disposal. By measuring the contrast function, it is possible to find out the line width of the light source with (1.7.2). Can you imagine that Michelson carried this out with the green Hg line and saw 540,000 wavelengths in path difference with his naked eye, Perot and Fabry saw 790,000 and Lummer and Gehrcke as much as 2,600,000! Try and picture this process: The shifting of the measurement reflector, observing the light and dark band with the naked eye and counting to 2,600,000 at the same time. Till now we have examined the situation where the light source has different emission bandwidths but only one individual emission line. If we use a laser as a light source, the emission can consist of several so called 'longitudinal oscillation modes', as long as we do not take specific measures to produce only one. In the following experiment a laser is used which emits only two modes. Each of the two modes has a narrow bandwidth (approx. 2 MHz) but both modes have a difference in frequency of 900 MHz.



# Fig.23: Emission spectrum of the two-mode HeNe laser used

The expected interferogram will consist of a superimposition of the waves with both the frequencies  $\omega$  and  $\delta\omega$ .

$$I = I_0 \cdot \left\{ 1 + \frac{1}{2} \cos\left(\frac{\omega}{c} \cdot \Delta L\right) + \frac{1}{2} \cos\left(\frac{\omega + \delta \omega}{c} \cdot \Delta L\right) \right\}$$

We can form a picture of the contrast function without having to make an explicit evaluation of the integral according to (1.7.2) by discussing the interferogram for some values of  $\Delta L$ , using

$$\cos\alpha + \cos\beta = 2 \cdot \cos\left(\frac{\alpha + \beta}{2}\right) * \cos\left(\frac{\alpha - \beta}{2}\right)$$
$$I = I_0 \left\{ 1 + \cos\left(\frac{2\omega + \delta\omega}{2c} \cdot \Delta L\right) \cdot \cos\left(\frac{\delta\omega}{2c} \cdot \Delta L\right) \right\}$$

Since  $\omega >> \delta \omega$ 

$$\mathbf{I} = \mathbf{I}_0 \left\{ 1 + \cos\left(\frac{\omega}{c} \cdot \Delta L\right) \cdot \cos\left(\frac{\delta \omega}{2c} \cdot \Delta L\right) \right\}.$$

Since  $\delta \omega$  is not dependent on  $\omega$ , this term is constant as far as the integration of all wavelengths is concerned and the contrast function of the single mode emission  $V_1(\Delta L)$  becomes:

$$V_{2}(\Delta L) = V_{1}(\Delta L) \cdot \cos\left(\frac{\delta \omega}{2c} \cdot \Delta L\right)$$

The contrast function resulting from the two-mode emission is reproduced by the single mode emission, but the periodicity  $\delta\omega/2$  superimposes with the cosine. The contrast is therefore zero for

or

$$\frac{\delta\omega}{2\cdot \mathbf{c}}\cdot\Delta\mathbf{L} = (2\mathbf{n}+1)\cdot\frac{\pi}{2}, \, \mathbf{n} = 0, 1, 2 \dots$$

$$\Delta \mathbf{L} = (2\mathbf{n}+1) \cdot \frac{\mathbf{c}}{\delta \omega}.$$

The intervals between the zero points is exactly

$$\Delta L = \frac{c}{2 \cdot \delta v}$$

Let us remind ourselves of the interval in frequencies of longitudinal modes of a laser resonator:

# **Michelson Interferometer**

$$\delta v = \frac{c}{2 \cdot L}$$

In this case L is the interval between the laser mirrors, so the interval between the zero points of the contrast corresponds exactly to the resonator length of the two-mode laser used.



# Fig.10: Contrast as a function of the path difference of a single mode and a two-mode laser

Although the radiation is coherent, there is a disappearing contrast. In this case, it is not due to a lack of coherence, but to the phenomenon of optical interference. If the laser emission is put onto a photodetector at a suitably fast speed, we will have direct proof of optical interference. Optical interference can be demonstrated using an oscilloscope or a spectrum analyser. However, this would mean that when a two-mode laser is used in interferometry the available measurement area of the index arm would be limited to L/2. Although the contrast does increase when this value is exceeded there is no information available during the zero crossings.

### 1.8 Wave fronts

The planes at which light has the same phases in a spatially wide radiation emission are known as wave fronts.



### Fig.24: Plain wave fronts of a parallel light bundle

If the light spreads itself in a parallel fashion, the planes of identical phases will be plane and therefore the wave fronts are called plane.



### Fig.25: Curved wave fronts of a divergent light bundle

However, if the radiation expansion is divergent, the wave fronts will be curved. If the radiation spreads in a solid angle with the same divergence in all directions, the wave fronts look like the surface of calotte shells. If two parallel light bundles are superimposed with plane wave fronts, we will see only brightness or darkness on a screen according to the relationship of the phases of the waves. If both bundles are attracted to each other there will be a striped pattern.



# Fig.26: Two light bundles attracted to each other with plane wave fronts

If one plane of a wave front, which represents the maximum of the field strength, cuts the plane of the minimum of a second wave front, then these are the locations for destructive interference. Since the wave fronts in the example according to Fig.26 represent planes these intersecting points are on a straight line.

# **Michelson Interferometer**



### Fig.27: Superimposition of two divergent light bundles

If, however, two light bundles with a slight divergence are superimposed, the intersecting points of the field strength minima or maxima lie on circles in space. A ring structure is formed.

### 1.9 Gaussian beams

In reality, there are no actual parallel light bundles and even wave fronts also only exist at a particular point. The cause for the failure of geometric optics lies in the fact that they were established at a time when people did not yet know that light is an electromagnetic wave and that its behaviour could therefore be described using Maxwell's equations. We will again use the wave equation already known to us for this purpose

$$\Delta \vec{\mathrm{E}} - \frac{\mathrm{n}^2}{\mathrm{c}^2} \cdot \frac{\partial^2 \vec{\mathrm{E}}}{\partial \mathrm{t}^2} = 0$$

If light were limitless, it would spread itself in all directions in space as a spherical wave.

$$\vec{E} = \vec{E}(r)$$
 with  $r^2 = x^2 + y^2 + z^2$ 

If, however we are interested in the technically important case of the spherical wave spreading in direction z in a small solid angle then the solution is an equation for the electrical field

$$\vec{E} = \vec{E}(r, z)$$
 with  $r^2 = x^2 + y^2 + z^2$ 

The solution to the wave equation is found in fields, which show a Gaussian shaped intensity distribution over the radiation cross-section and are therefore called Gaussian beams. According to the existing marginal conditions, Gaussian beams can be found in various modes. Such beams, especially the Gaussian basic mode (TEM<sub>00</sub>) are preferably produced by lasers. However, the light coming from any light source can be seen as a superimposition of many such Gaussian modes. But the intensity of a pure mode is very small compared to the total intensity of a light source. The situation is different with lasers where the whole light intensity can be produced in the basic mode alone. This, as well as the monochromatic nature of laser radiation, is the main difference compared to conventional light sources. A Gaussian beam always has a beam waist. The beam radius w (w=waist) is a result of the solution of the wave equation

$$\mathbf{w}(\mathbf{z}) = \mathbf{w}_0 \cdot \sqrt{1 + \left(\frac{\mathbf{z}}{\mathbf{z}_R}\right)^2}$$

 $w_0$  is the smallest beam radius of the beam waist and  $z_r$  is the Rayleigh length.

Fig.29 shows the passage of the beam diameter when it is dependent on the length z.

$$z_{\rm R} = w_0^2 \frac{\pi}{\lambda}$$



# Fig.28: Beam diameter of a Gaussian beam in the basic mode TEMoo as a function of the location z.

The beam spreads itself in the z-direction. At the point when  $z=z_0$  the beam's radius is at its smallest. As the interval increases, the beam radius grows linearly. Since Gaussian beams are spherical waves, there is a radius of curvature of the wave front for every point z. The radius of curvature R can be calculated using the following equation:

$$R(z) = z + \frac{z_r^2}{z}$$

This situation is shown in Fig.29.



# Fig.29: The radius of curvature of a wave front as a function of the distance from the beam waist at z=0

The radius of curvature has a minimum at  $z = z_r$  and R increases by 1/z compared to z = 0. The radius of curvature is infinite at z=0. The wave front is even at this point. The ra-

# **Michelson Interferometer**

dius increases again linearly above the Rayleigh length  $z_r$ . This statement is of fundamental importance. It states that a parallel beam only exists at one point of a light wave, that is, in its focal point.

In the region for

$$-z_r \le z \le z_r$$

a beam can be seen as approximately parallel or collinear. In Fig.30 the Rayleigh area is drawn in as well as the divergence  $\Theta$  in the far field region, i.e.  $z >> z_0$ .

The graphic representations suggest that one of the significant characteristics of laser beams, i.e. their lack of divergence, cannot be shown in this way.



Fig.30: Rayleigh range  $z_R$  and Divergence  $\Theta$  in the far field region, i.e.  $z > z_R$ )

This is because the relationship of the beam diameter to z in the graphs has not been standardised. If we take a HeNe laser (632 nm) with a beam radius of  $w_o = 1$ mm at the exit of the laser as an example, we get:

$$2 \cdot z_{R} = 2w_{0}^{2} \frac{\pi}{\lambda} = 2 \cdot 10^{-6} \frac{3.14}{623 \cdot 10^{-9}} = 9,9 \text{ m}$$

for the Rayleigh range  $2z_r$ . A moderate picture would therefore show only a thin line in the drawing.

### 1.10 Technical interferometers

In modern technical interferometers used in the calibration of CNC machines, the main light sources used are HeNe lasers. Modern CNC machines are positioned precisely at 1  $\mu$ m. Since this precision is dependent on the condition of the machines, a periodic calibration of the machines must follow. The measurement apparatus, the laser interferometer is supposed to be more precise than the test specimen according to Bernd's Law, i.e. it has a precision of 0.1  $\mu$ m. As we have already said previously and will be studying in more depth, it is possible to carry out resolutions into the nm area with laser interferometers. However, we have yet to establish the precision. We must first observe the errors for this purpose. If we want to reach the path S of the movable reflector M<sub>2</sub> through the number of light/dark transitions N, we must use the following equation

$$S = N \cdot \frac{\lambda}{\epsilon}$$

In this equation,  $\lambda$  is the wavelength of the light source used and  $\varepsilon$  is an integer scaling factor which had a value of 2 in the interferometer we have talked about till now. We must also keep in mind that the value of the wavelength  $\lambda$  depends on the refractive index n of the medium in which the wave expands.

$$\mathbf{S} = \mathbf{N} \cdot \frac{\boldsymbol{\lambda}_0}{n} \cdot \frac{1}{\varepsilon}$$

 $\lambda_0$  is the wavelength in vacuum. If we use the total differential to determine the relative error of S, we will obtain:

$$\left|\frac{\partial S}{S}\right| = \left|\frac{\partial N}{N}\right| + \left|\frac{\partial \lambda_0}{\lambda_0}\right| + \left|\frac{\partial n}{n}\right| + \left|\frac{\partial \varepsilon}{\varepsilon}\right|$$

In a machine working range S of 1m the typical relative mechanical error of the machine will be  $1\mu m/1m=10^{-6}$ . This means that the relative error of the laser interferometer for determining S should be smaller than  $10^{-7}$ 

$$\left|\frac{\partial S}{S}\right| \le 10^{-7}!$$

The sum of the absolute values of individual error therefore has to be smaller than this value

$$\left|\frac{\partial N}{N}\right| + \left|\frac{\partial \lambda_0}{\lambda}\right| + \left|\frac{\partial n}{n}\right| + \left|\frac{\partial \varepsilon}{\varepsilon}\right| \le 10^{-7}$$

Let us first consider the term

$$\frac{\partial N}{N}$$

The absolute error  $\partial N$  for determining a change from a dark to light fringe is 1/100 of one counted fringe. There are 3,164,557 fringe counts N for 1m resulting in

$$\left|\frac{\partial N}{N}\right| = \frac{0.01}{3.164.557} = 3 \cdot 10^{-9}$$

This is sufficiently precise. The errors formed through the variation in wavelength are:

$$\left|\frac{\partial \lambda_0}{\lambda}\right| = \frac{\partial \lambda}{0,632 \cdot 10^{-6}}$$

This leads to the demand that the absolute precision  $\delta\lambda_0$  of the vacuum wavelength of the laser is smaller than  $10^{-13}$  m or  $10^{-7}$  µm. The wavelength of a commonly used HeNe laser can fluctuate in a range of 1.5  $10^{-6}$ . The laser will then have to be additionally stabilised for a measurement area of 1m. The term

∂n
n

is caused by fluctuations in the refractive index caused by the surrounding air. The respective values of n are dependent on the air pressure, air temperature, the relative humidity and foreign gas in the surrounding air. The biggest influence on errors is caused by the variation in air pressure and in temperature. These values must be applied to reduce these errors and be taken into consideration in a correction calculation using the Edlen formula. The air pressure should be measured with an accuracy of 0.1 Torr and the temperature of  $0.1^{\circ}$  C, so as not to exceed the permitted errors. The

final source for an error is the scaling factor  $\varepsilon$  that is, however, a flawless constant in the apparatus. Modern HeNe laser with stabilised frequencies reach a precision of  $10^{-9}$ . Apart from the type of stabilisation, all lasers have one disadvantage: Light cannot be reflected back into the laser. This is because the returning photons have another phase as those responsible for the stimulated emission in the laser resonator. Photons coming from outside would stimulate an emission which is not in phase with the basic oscillation of the resonator and can therefore stimulate undesired modes, which have a frequency other than the desired one. Because of this reason, the Michelson interferometer had to be modified so that no direct back coupling of light in the resonator would be possible.



# Fig.31: Commonly used HeNe laser tube with mirrors soldered on to it with a distance of L.

The resonator length is approx. 130mm for a single mode tube and approx. 180mm for a two-mode tube.



Merging point Beam splitter

### Fig.32: Setup of a technical Laser Interferometer.

Instead of the beam splitter plate in the Michelson construction, a beam splitting prism is now used and instead of the plane mirrors, triple reflectors are used. This arrangement ensures that no radiation can return to the laser, contrary to the Michelson interferometer. The surfaces of optical components are also given anti-reflection layers to avoid return reflections from their surfaces. Polarising beam splitting prisms are usually used. The light in the reference arm is then vertically or orthogonal polarised with regard to the index arm. Before we discuss signal formation we will explain the structure and function of the optical components.

### 1.10.1 Beam splitting prism

The term beam splitting prism is used in common jargon. Actually a plate could also have been used instead of the prism. However, as shown in the following diagram, the plate has some disadvantages.



Fig.33: Arrangement with a beam splitting plate

On one hand, there is a misalignment of beam for the sensitive index arm, which depends on the angle adjustment of the plate. On the other hand, the index beam crosses the plate twice. If the temperature of the plate changes, then its refractive index also changes considerably. This causes a phase shift that is interpreted as a movement of the triple reflector. The structure is not thermally compensated for.





In this arrangement, the reference beam crosses the paths C+ triple reflector +D after the splitting point in the glass and up to the reunification point. The index beam crosses the paths E + triple reflector + A. These paths have the same length. A temperature change in the optical components therefore does not lead to a phase shift. The arrangement is compensated for thermally with regard to the optical components. Polarising beam splitting prisms are mainly used in technical laser interferometers that measure length. Contrary to neutral beam splitting prisms, the light is not split according to intensity, but is divided into orthogonal polarisation states.



### Fig. 35: Polarising beam splitting prisms

Vertically polarised light (S) is reflected on the inner layer of the prism, parallel light (P) is transmitted. If the polarisation direction of the light bends towards  $45^{\circ}$ , the beam splitting prism divides this light into its orthogonal components.

### 1.10.2 Triple reflector

If you cut a corner off a quadratic cube and put mirrors on the inner surfaces, you will get the important optical components of a triple reflector. Every light beam, which is led into the corner of this cube under any given angle, goes through three different reflections (A,B,C) and is finally reflected back into the same direction. There is, however, an offset between the incoming and outgoing beam, which is dependent on the place in the triple reflector onto which it falls.



### Fig.36: Triple reflector

A corner can also be cut out of a cube made totally of glass. The reflection is then total reflection. This type of triple reflector is used most frequently. The advantage is, that the reflection surfaces are protected from dust and dirt.

### 1.10.3 Homodyne interferometer

The term homodyne interferometer was conceived at a time when heterodyne interferometers were in predominant use. The first commercial laser interferometer was patented and brought into the market about 25 years ago by Hewlett Packard. Other manufacturers because of the patent situation produced Homodyne interferometers. Today both types of systems are in usage. The light beam in the reference arm has a very different frequency to the light beam in the index arm. In this way the possibility of an alternating current coupled with gating is possible. The fluctuations in contrast are then of no significance. A detailed discussion will follow in the relevant chapter. The homodyne technique will be used for this experiment. This means that there is no frequency difference between the light in the index and reference arms. The structure is identical to Fig.32. However, special measures must be taken in detection so that the changing contrast does not influence consistent counting of the interference fringes. The following signal sequences are produced and evaluated for this purpose.



# Fig.37: Signal sequences required for safe counting of interference fringes.

The curves correspond to the equations

 $\mathbf{A} \qquad \mathbf{I} = \frac{1}{2} \mathbf{E}_{01}^2 + \frac{1}{2} \mathbf{E}_{02}^2 + \mathbf{E}_{01} \mathbf{E}_{02} \cos\left(\mathbf{k} \cdot \Delta \mathbf{L}\right)$  $\mathbf{R} \qquad \mathbf{I} = \frac{1}{2} \mathbf{E}_{01}^2 + \frac{1}{2} \mathbf{E}_{02}^2 + \mathbf{E}_{01} \mathbf{E}_{02} \cos\left(\mathbf{k} \cdot \Delta \mathbf{L}\right)$ 

$$\mathbf{B} = \frac{1}{2} \mathbf{E}_{01}^2 + \frac{1}{2} \mathbf{E}_{02}^2 - \mathbf{E}_{01} \mathbf{E}_{02} \cos(\mathbf{k} \cdot \Delta \mathbf{L})$$
$$\mathbf{D} = \mathbf{I} = \frac{1}{2} \mathbf{E}_{01}^2 + \frac{1}{2} \mathbf{E}_{02}^2 + \mathbf{E}_{01} \mathbf{E}_{02} \sin(\mathbf{k} \cdot \Delta \mathbf{L})$$

$$\mathbf{D} \qquad \mathbf{I} = \frac{1}{2} \mathbf{E}_{01} + \frac{1}{2} \mathbf{E}_{02} + \mathbf{E}_{01} \mathbf{E}_{02} \sin(\mathbf{k} \cdot \Delta \mathbf{L})$$
$$\mathbf{E} \qquad \mathbf{I} = \frac{1}{2} \mathbf{E}_{01}^2 + \frac{1}{2} \mathbf{E}_{02}^2 - \mathbf{E}_{01} \mathbf{E}_{02} \sin(\mathbf{k} \cdot \Delta \mathbf{L})$$

Curve A is the reference curve. Curve B is formed out of A through a phase shift of 180° and curve D is formed when curve A undergoes a phase shift of 90°. Curve E is formed when curve D goes through a phase shift of 180°. If we subtract curve B from curve A, the base is eliminated. The same is done with curves D and E. We can spare ourselves the subtraction work if we put the signals A and B, as well as E and D on a TTL-comparator. This type of comparator

$$Output = \begin{cases} High & \text{if } A > B \\ Low & \text{if } A < B \end{cases}$$

has the following output characteristic:

Signals C and F are produced in this way. Strictly speaking only counting the signal C is necessary, but we would then have no information on the direction in which the measurement reflector moves. The signal whose phase has been shifted by 90° has been produced for this purpose alone. The recognition of the direction happens in such a way that a toggle is dependent on the entrance at which a decreasing side is first detected. The exit signal from the toggle controls the following counter in the up or down mode. Let us make a timed observation of the signals C and F to understand this better. If the measurement reflector moves out of a condition at z=0 (Fig.37) e.g. to the right, then the intensity at detector A increases and decreases at B. A is larger than B; the exit of the comparator is high. After some time B>A and the corresponding comparator produces a decreas-

ing side. The toggle sets the counter to up. If we were to go to the left from z=0 the first decreasing side would come from F and the toggle would be set back to enable the counter to switch to the Down mode. The necessary signals cannot be produced electronically. They therefore have to be provided by the interferometer.

**T1** 

### 1.10.4 Detection optics



### Fig. 1: Detection optics

By using a polarising beam splitter, the beams in the reference arm are polarised vertically to those in the index arm. It is well known that such polarising beams do not produce any interference pattern. They are sent through a small  $\lambda/4$ plate (L4) to "persuade" them to make an interference pattern. This type of small plate consists of a birefringent crystal, e.g. natural quartz. This kind of crystal is not isotropic optically. Due to the natural crystal growth it has two particular directions optically speaking that are caused by the asymmetry of the crystal cells. The inner structure of such a crystal can be compared to a forest of arranged antennae. An equal number of antennae receive both vertically and horizontally polarised radiation. These antennae are, in fact, dipoles that release the received radiation after a certain period of time. These dipoles take up a light wave entering the crystal one after the other and release it again. The two types of dipole also cause two different speeds of light in the crystal.



### Fig.39: Birefringent crystal as a certain number of orthogonal dipoles

If light with a polarisation of  $45^{\circ}$  is beamed into the "antennae forest", the light intensity will be equally spread on the dipoles since the light with a polarisation of  $45^{\circ}$  can be interpreted as the superimposition of two orthogonal components.



Fig. 40: Effect of a quarter-wave plate

At the exit of the crystal there is a phase shift between the S and P components, which is composed by the different speeds of the components in the crystal. If we select the number of antennae or the thickness of the crystal in such a way that this phase delay is just 90 or  $\lambda$  /4, then we obtain two merging waves which are vertically polarised towards one another and also have a phase shift of 90°. If we add up these field strength components vectorally, the resulting field strength vector will rotate with the frequency of the light beam on the axis of the direction of expansion. The light is now circularly polarised. Depending on whether the quarter-wave plate is reached under a beaming in angle of 45° or 135°, the polarised light will be right or left-handed circular. The quarter-wave plate sometimes has a mark on the edge that shows the direction "fast optical axis". This axis is also described as the direction of the regular refractive index and the one at 90°, which is turned at a certain angle to it, is called the "slow axis, also known as irregular refractive index.



Fig.41: Circular polarised light

Since we are sending two waves that are vertically polarised to each other from the interferometer into the crystal, we will obtain one left-handed and one right-handed circular wave behind the crystal. This can be shown more simply in the diagram of a circle.



Fig.42: Orthogonal circularly polarised waves and their resultants.

The above diagram shows a time exposure of the waves behind the quarter-wave plate. An indicator rotates at the frequency  $\omega_{index}$ ; another one rotates in the opposite direction at a frequency  $\omega_{ref}$ . If both frequencies are the same, the resulting indicator will be symmetrical to both the other ones and will only change its number. These changes happen so fast that a photodetector cannot follow them. If there is a difference in frequency between the beams coming from the index and reference arms, the resulting indicator starts to rotate with the frequency of the difference. The direction of rotation depends on the signs indicating the difference. If we allow this light to pass through a polariser, a photodetector will detect a signal which has been temporally changed and which coincides exactly with the frequency difference. But what is the reason for this frequency difference between the reference and index beam? As long as the measurement reflector is not used, both frequencies are the same. If, however, the measurement reflector moves, the light frequency changes due to the Doppler effect. This happens according to:

$$\omega_{\rm MeB} = \omega_0 \cdot \left(1 \pm \frac{2 \cdot v}{c}\right) \quad \omega_{\rm Ref} = \omega_0$$

In this case  $\omega_0$  is the frequency of the laser, v is the speed of the measurement reflector and c the speed of light. Factor 2 of the speed moves, because the beam length is changed by 4 cm, when, for example, 2cm is shifted due to the double passage of the light in the index arm. The difference in frequency is therefore:

$$\Delta \omega = \omega_{\text{Ref}} - \omega_{\text{MeB}} = \mp \omega_0 \frac{2 \cdot v}{c}$$

We will not use the sign  $\pm$  just now. The following are also used:

or

$$dz = \frac{c}{2\omega_0} \cdot \Delta \omega \cdot dt$$
$$\frac{dz}{dz} = \frac{c}{2\omega_0} \cdot \int_{t_1}^{t_2} \Delta \omega \cdot dt = \frac{\lambda_0}{2} \cdot \int_{t_1}^{t_2} \Delta \nu \cdot dt$$

 $\Delta \omega \cdot \frac{c}{2 \cdot \omega_0} = \frac{dz}{dt}.$ 

If we therefore take into consideration that  $\Delta v$  is the number  $\Delta N$  of the rotations of the resulting indicator per time unit counted by a counter, then

$$\Delta v = \frac{\Delta N}{dt}$$
$$s = z_2 - z_1 = \frac{\lambda_0}{2} \cdot \int_{t_1}^{t_2} \frac{\Delta N}{dt} \cdot dt = \frac{\lambda_0}{2} \cdot N$$

The result is therefore the same as in the simple Michelson interferometer. The reason why this method is used is because it is now possible to split up the resulting beam and to send it through differently oriented polariser. Let us return to Fig.38 where we can see in the optical structure that both the orthogonal circularly polarised waves meet after the quarter-wave plate (L4) on the neutral beam splitter (P1) and are split into two channels at that point. One part of the radiation reaches the polarising beam splitting prism P2. This fulfils two tasks. On one hand it splits the radiation and on the other hand it is effective as a polarising analyser. The phases of the signals A and B are then shifted towards each other by 180°, as desired. The polarising beam splitter (P3) is arranged at a 45° angle in the second channel, so this channel has a phase shift of 90° compared to the other channel. The phases of the signals C and D are again shifted towards each other by 180°. In this way we obtain the necessary phase shifts to compensate the varying contrast on one hand, and to obtain signals for the forward/reverse recognition on the other hand. The difference between the Michelson interferometer and this method is, that we do not obtain the path information from a contrasting picture dependent on location, but from the fluctuation or mixture of two optical frequencies. Since we use only one basic frequency for the mixing process, i.e. the frequency of light, this process is also called the homodyne process.

### 1.10.5 Heterodyne interferometer

An optical intermediate frequency is used in this process. It is obtained mainly from the frequency interval between the modes of a laser.



### Fig.43: Heterodyne interferometer

This frequency difference is approx. 2MHz in the case of older equipment. It is produced when a single mode HeNe laser tube is surrounded by a longitudinal magnetic field. The mode splits into two circularly polarised modes due to the Zeeman effect. The size of the slit and the frequency difference depends on the magnetic field strength. After both circular modes have crossed a quarter-wave plate, the modes that are now orthogonal are sent into the interferometer, which would otherwise have the same optical structure.

A polarisation filter is placed in front of detector B, which detects the frequency

 $\omega_1\, \lrcorner\, \omega_2$ 

Detector A sees the frequency

$$\omega_1 - \omega_2 + \delta \omega$$

In this case  $\delta \omega$  is the change in frequency that is produced by the Doppler effect or by the movement of the measurement reflector. Both signals are electronically mixed, subtracted from each other and carried out using a phase discriminator, which carries out forward-reverse recognition. Since both detector A and detector B only "see" alternating stress signals, fluctuations in the contrast are already eliminated. Optical detection is obviously easier but the electronic preparation would be more expensive. Nowadays both interferometers have an equal place in the market and both perform with the high precision required.