Université de Neuchâtel

Institut de Microtechnique

Signal Processing For Optical Phase Detection

Thèse

présentée à la Faculté des sciences
pour obtenir le grade de docteur ès sciences

par

Eric Zimmermann

Neuchâtel, juin 1997

IMPRIMATUR POUR LA THÈSE

Signal Processing for Optical Phase Detection

de M. Eric Zimmermann

UNIVERSITÉ DE NEUCHÂTEL

FACULTÉ DES SCIENCES

La Faculté des sciences de l'Université de Neuchâtel sur le rapport des membres du jury,

MM. R. Dändliker (directeur de thèse),
F. Pellandini, P. de Groot (Middlefield USA),
J.-F. Willemin (Greifensee)

autorise l'impression de la présente thèse.

Neuchâtel, le 19 décembre 1997

Le doyen:

F. Stoeckli.

F. Stoeckli

La vraie force de l'intelligence n'est pas de comprendre les choses difficiles, mais de les dépouiller de ce qui les empêche d'être simples.

Pierre Billion,

"L'Ultime Alliance"

Abstract

In the first part, a novel full-dynamic phase-demodulator for heterodyne interferometric vibration measurements is described. Contrary to a conventional phase-demodulator, which has a limited dynamic range of ± 90°, the proposed phase-demodulator has an unlimited dynamic range. This allows to recover phase information for displacements of the surface under test which are much greater than the wavelength of light. An analog version with a cut-off frequency of 6.7 kHz and a digital version of the phase- demodulator have been developed. The digital version of the full-dynamic phase-demodulator has an extended frequency response up to one half of the heterodyne frequency. For a carrier-to-noise ratio of 70 dB within 3 kHz bandwidth, we measured a phase noise of less than 3 μrad/√Hz, for the analog version, and 5 μrad/√Hz, for the digital version. The performance of both versions of the phase-demodulator are analyzed and several examples illustrate how well our vibration measurement system can be used to study complex movements of objects.

In the second part, an original signal processing method to determine the center position of white-light fringe signals is presented. This is achieved in two steps. First, the center of gravity of the signal power is calculated to better than half a fringe period. Second, the phase of the fringe signal is determined to get the exact center position of the white-light fringes. For the implementation of this method, we introduce an optimum sampling concept, which is based on synchronous sampling with four samples per fringe period. The theoretical predictions for systematic and statistical errors are compared with computer simulations. Experimental results show, that the proposed signal processing is simple to operate, fast, accurate and extremely noise resistant. Finally, we apply this signal processing scheme to absolute force measurements using the induced birefringence in a optical monomode fiber. Here, the polarization interferometer is composed of a Wollaston prism to compensate the phase retardation induced by the force.

Table of contents

1. Introduction

In the preamble to this introduction, we would like to cite a part of the preface to the proceedings of a SPIE conference on Laser Interferometry [1.1].

> "The continued drive for higher performance and greater reliability of devices, structures and processes places stringent demands on the methods used in their development and operation. Optical metrology is a major and inseparable part of these methods. Interferometry is an essential field of optical metrology and relates more and more to industrial applications."

Before interferometry has become an important technique to provide high resolution and high accuracy measurement systems, both for scientific and industrial applications, an impressive amount of research had been carried out by many famous scientists. Indeed, the history of optical interferometry extends over more than three hundred years since the first description by Boyle and Hooke in the second half of the 17th century of what is actually known as "Newton rings" and the first development of wave theory by Huyghens in 1690. Therefore, since three centuries, optical interferometry is intimately linked to history of science and it has played a crucial role in the development of several important chapters in science and engineering.

Until the early 1960's, the light source of choice for interferometry was the mercury arc lamp followed by a filter to select the green line and a pinhole to increase the spatial coherence. If this source was very helpful for laboratory experiments, its poor spatial and temporal coherence, and in addition, the very low intensity available, limited the number of industrial applications. But in 1960, with the advent of lasers, an intense light source with a high degree of spatial and temporal coherence was, for the first time, available. Since that time, optical interferometry has been applied to an incredibly large range of applications, as can be seen from the large number of books, scientific papers, patents, etc., devoted to this subject.

One main issue for laser interferometry is related to the electronic processing and the analysis of the interference signal. Indeed, the measurement accuracy depends strongly on the signal processing unit (hardware and/or software) employed to

recover the phase information from the interference signal. The initial motivation behind the work presented in this thesis emerges directly from the application of interferometric methods in two practical applications, where the signal processing for optical phase detection plays the leading role to reach the required specifications of the whole measurement system.

The first part of this thesis deals with the application of heterodyne interferometry to highly accurate vibration measurement. Indeed, the interference of two waves of slightly different frequencies produces a beat signal. Since measurements of either the frequency or the phase of this beat signal can be made with very high precision, such heterodyne technique allows to achieve fringe interpolation to better than one thousandth of a fringe. The work developed in this thesis follows directly earlier studies carried out several years ago by J.-F. Willemin [1.2 and 1.3]. In this previous development, the interference signal was demodulated using a frequency-demodulator to recover the modulation which is proportional to the instantaneous velocity of the vibrating object. Unfortunately, with a FM-demodulator the displacement sensitivity decreases with frequency. Although tolerable for frequency vibrations greater than a few hundreds of Hz, this limitation is a severe drawback for studying vibrating objects at low frequencies. Since many drawbacks of the measurement system based on FM-demodulation are given by the demodulator itself, it was decided to develop a novel demodulator based on phase-demodulation rather than frequency-demodulation. The concept and circuit design of the phase-demodulator presented in this thesis has been undertaken by J.-F. Willemin and R. Dändliker. One main advantage of the proposed full-dynamic phase-demodulator is its unlimited dynamic range. This allows to recover phase information for displacements of the surface under test which are much greater than the wavelength of light. Moreover, a digital approach to phase-demodulation was developed to increase the throughput speed up to one half the heterodyne frequency. Two nearly identical phase-demodulators were built. One of them was transferred to USA and is actually tested by S. M. Khanna at the Columbia University, New York in a heterodyne interferometer for accurate measurements of the mechanical vibration of individual cells in the organ of Corti (inner ear) [1.4].

The main drawback of classical laser interferometry is the incremental manner of measuring, resulting from the periodic nature of the interference signal. Various techniques have been developed to overcome this severe limitation and to obtain

absolute (non-incremental) measurements with interferometric accuracy over long distances. One of these techniques is white-light interferometry, where a broadband source is used instead of a laser. The second part of this thesis is devoted to the description and analysis of a novel type of signal processing, particularly adequate for electronically scanned white-light interferometry. In this method, the interferogram is projected onto a photodiode array and is scanned in space rather than in time. The absence of moving parts in the interferometer increases the mechanical stability of the system and reduces its complexity. In the framework of a research project with the company Hilti AG (Schaan, Liechtenstein), a polarizing interferometer for absolute force measurements using the induced birefringence in an optical fiber has been developed. For this application, we introduced an elegant solution for the receiving interferometer, using a Wollaston prism to compensate the phase retardation which is induced by the force. Two patent applications were submitted in the framework of this project [1.5, 1.6].

Chapter 2 presents a general overview of high resolution interferometric techniques. We specially emphasize the influence of the light source on the interference signal detected at the interferometer output. The possibility to extend the range of non-ambiguity for absolute measurements using either multiple-wavelength or white-light interferometry is exposed.

In chapter 3, the full-dynamic phase-demodulator for heterodyne interferometric vibration measurements is described. The signal processing program employed to recover off-line the phase information is also presented in this chapter. After a complete description and analysis of the demodulator, several examples are shown to illustrate how well our vibration measurement system can be used to study complex movements of objects.

Chapter 4 is dedicated to the study of the signal processing for white-light interferometry. The behavior of this signal processing in presence of noise is analyzed and solutions to improve its performances are proposed. Finally, a white-light interferometer is used for absolute measurements of force using the induced birefringence in a monomode optical fiber.

The research projects were partially supported by the Commission pour l'Encouragement de la Recherche Scientifique (CERS Project 1905.1) and by the Swiss Priority Program "OPTIQUE".

2. High resolution interferometric techniques

An extraordinary variety of measurements can be made using optical interferometers [2.1, 2.2, 2.3]. Indeed, interferometry is a fast and highly accurate measurement technique, which has been used in many scientific and industrial metrology problems. Actually, interferometric sensors continue to be the subject of an intense research and development effort. New applications have been grown up by the use of optical fibers to create fiber versions of classical interferometer configurations, such as two-beam Michelson, Mach-Zehnder, Fizeau, Fabry-Pérot and Sagnac interferometers [2.4]. Moreover, using single frequency laser diodes leads to the development of compact and accurate sensors [2.5, 2.6]. Yet another very interesting improvement for interferometry that begins to appear is the new development of smart detectors for optical metrology with pre-processing electronics on chip [2.7, 2.8]. Today, examples of current applications for optical interferometric systems are highly accurate measurements of length or displacement [2.9], absolute distance [2.10, 2.11, 2.12, 2.13] and vibration [2.14], optical-surface testing [2.15], surface profilometry [2.16, 2.17], measurements of temperature, strain and pressure [2.18, 2.19], velocity measurements of fluids [2.20].

The aim of this chapter is to provide an overview of interferometric techniques. The first two paragraphs will shortly summarize the basic properties of electromagnetic waves and cover such topics as interference between two monochromatic light beams, the interference of partially coherent light and the concept of optical coherence. In the next paragraphs, we will describe the influence of the spectral characteristics on the interference signal, this for different types of light sources, namely: single frequency or highly coherent source, multiple-wavelength source and broadband or white-light source. We will also see that two major and fundamental problems associated with the used of highly coherent source in interferometry are their short unambigous operating range and the inability to identify the interference order when the source is switched on and off. The use of multiple-wavelength and white-light sources can help to overcome these drawbacks, and as a result, to extend both the range and utility of such sensors. In this part, we will specially emphasize the influence of the source on the interference signal observed at the interferometer output.

2.1 Electromagnetic waves

To simplify the purpose, we only consider the propagation of an electromagnetic wave in a linear, homogeneous and isotropic medium. In this situation, the electric field **E** and the magnetic field **H** of the electromagnetic wave lie in a plane normal to the direction of propagation [2.21]. Moreover, **E** and **H** are perpendicular and the magnitudes of the electric and magnetic fields are related by $|\mathbf{H}| = \sqrt{\varepsilon\varepsilon_0/\mu\mu_0}\,|\mathbf{E}|$, where ε_0 is the electric permitivity of vacuum, ε is the relative permitivity of the medium, μ_0 is the magnetic permeability of vacuum and μ is the relative permeability of the medium [2.21]. Using a scalar representation, the electric field E(r, t) of a monochromatic electromagnetic wave can be written as

$$E(\mathbf{r},t) = \frac{1}{\sqrt{2}}\left(\frac{\varepsilon\varepsilon_0}{\mu\mu_0}\right)^{1/4} \mathrm{Re}[U(\mathbf{r},t)], \tag{2.1}$$

where **r** is the position vector and t is the time. The complex wavefunction U(**r**, t) can be express in the form

$$U(\mathbf{r},t) = V(\mathbf{r})\exp(i2\pi\nu t), \tag{2.2}$$

where ν is the frequency of the electromagnetic wave. The frequency ν is related to the wavelength of the wave by $\nu\cdot\lambda = c_0/n$, where $c_0 = 1/\sqrt{\varepsilon_0\mu_0}$ is the velocity of the light in free space and $n = \sqrt{\varepsilon\mu}$ is the refractive index of the medium. The time-independent factor V(**r**) is called the complex amplitude and can be reformulated as

$$V(\mathbf{r}) = V_0(\mathbf{r})\exp[i\varphi(\mathbf{r})], \tag{2.3}$$

where $V_0(\mathbf{r})$ is the amplitude and φ is the spatial dependent phase of the wave. The optical intensity I(**r**, t), defined as the optical power per unit area, is proportional to the average of the squared wavefunction [2.22], namely

$$I(\mathbf{r},t) = \left\langle |U(\mathbf{r},t)|^2 \right\rangle, \tag{2.4}$$

where <·> denotes averaging over a time interval that is much longer than an optical period 1/ν. Substituting Eq. (2.2) into Eq. (2.4), it can be shown that the

optical intensity of a monochromatic wave is the absolute square of its complex amplitude, namely

$$I(\mathbf{r}) = |V(\mathbf{r})|^2. \tag{2.5}$$

It should be noted that the intensity of a monochromatic wave does not change with time.

The simplest solutions of the Helmholtz equation in a homogeneous medium are the plane and the spherical waves [2.22]. In general, the complex amplitude of a plane wave can be expressed as

$$V(\mathbf{r}) = V_0 \exp[-in\mathbf{k}\cdot\mathbf{r}], \tag{2.6}$$

where n is the refractive index of the medium and $\mathbf{k} = (k_x, k_y, k_z)$ is called the wavevector. The magnitude of the wavevector is equal to

$$k = \frac{2\pi}{\lambda} = \frac{2\pi\nu}{c_0}, \tag{2.7}$$

where c_0 is the velocity of the light in free space. Using Eqs. (2.5) and (2.6), the intensity of a plane wave is simply given by $I = V_0^2$. One important practical case of Eq. (2.6) is a plane wave propagating in the z-direction. In this situation the complex amplitude is expressed as

$$V(z) = V_0 \exp(-inkz) = V_0 \exp\left(-in\frac{2\pi}{\lambda}z\right). \tag{2.8}$$

2.2 Interference phenomena

From the linearity of the wave equation and the Helmholtz equation, it follows that when two or more electromagnetic waves are present simultaneously in a same point of space, the total wavefunction is the sum of the individual wavefunctions. As we will see in the next paragraphs, this principle of superposition is not applicable to the optical intensity. Indeed, the total intensity is not necessarily the sum of the individual intensity. This situation is a common characteristic to all interference phenomena.

2.2.1 Interference between two monochromatic waves

We consider two monochromatic waves $V_1(\mathbf{r})$ and $V_2(\mathbf{r})$ of same frequency ν. The result of the superposition is a monochromatic wave of frequency ν with a complex amplitude $V(\mathbf{r}) = V_1(\mathbf{r}) + V_2(\mathbf{r})$. Following Eq. (2.5) the total intensity is given by

$$I = |V_1 + V_2|^2 = |V_1|^2 + |V_2|^2 + 2\mathrm{Re}\left[V_1^* V_2\right]. \tag{2.9}$$

Substituting the complex amplitude of each wave into Eq. (2.9), namely

$$V_1 = \sqrt{I_1}\exp[i\varphi_1] \quad \text{and} \quad V_2 = \sqrt{I_2}\exp[i\varphi_2], \tag{2.10}$$

where I_1, I_2 and φ_1, φ_2 are the intensity and the phase of the individual light beams, respectively, we obtain

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\varphi), \tag{2.11}$$

with $\varphi = \varphi_2 - \varphi_1$.

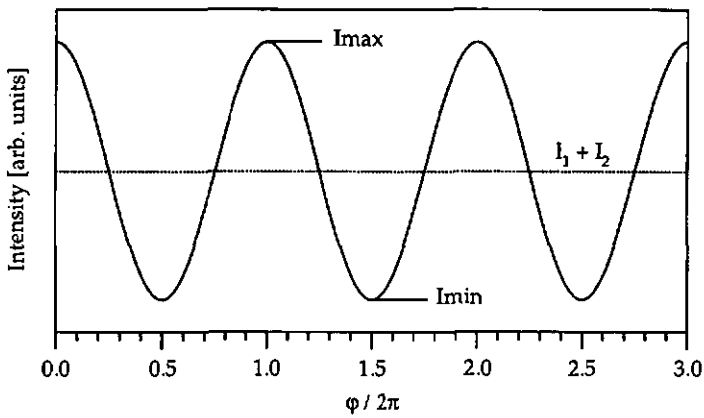


Fig. 2.1 Interference signal due to the superposition of two monochromatic waves.

This relation is of primordial importance. It shows that the total intensity due to the superposition of two monochromatic waves is equal to the sum of individual intensities plus a term which is proportional to the cosine of the phase difference

between the two beams φ as drawn in Fig. 2.1. The last term in Eq. (2.11) is attributed to the interference between the two waves. The visibility Γ of the interference fringes is defined by the relation

$$\Gamma = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \tag{2.12}$$

where $0 \le \Gamma \le 1$. In the present situation, from Eq. (2.11) and (2.12), we obtained

$$\Gamma = \frac{2\sqrt{I_1 I_2}}{I_1 + I_2}, \tag{2.13}$$

which is equal to $\Gamma = 1$ only when $I_1 = I_2$.

2.2.2 Interference of two partially coherent waves

If we consider the superposition between a partially coherent wave $A_1 \cdot U(t)$ and a portion of itself but delayed by the time τ, $A_2 \cdot U(t+\tau)$, where A_1 and A_2 are constants, a separate fringes system is produced for each wavelength and the interference signal can be interpreted as a weighted superposition of interferograms produced by each of the monochromatic component of the beam. Mathematically, the average intensity of their sum is

$$I = \left\langle \left| A_1 U(t) + A_2 U(t+\tau) \right|^2 \right\rangle = I_0 \left\{ 1 + 2 \frac{A_1 A_2}{A_1 + A_2} \mathrm{Re}\left[g(\tau) \right] \right\}, \tag{2.14}$$

where $I_0 = \left\langle |U(t)|^2 \right\rangle = \left\langle U^*(t) \cdot U(t) \right\rangle$ and $g(\tau)$ is the complex degree of temporal coherence [2.23]. The complex degree of temporal coherence is a measure of the degree to which $U(t)$ and $U(t+\tau)$ are correlated and is given by

$$g(\tau) = \frac{\left\langle U^*(t) \cdot U(t+\tau) \right\rangle}{\left\langle U^*(t) \cdot U(t) \right\rangle}. \tag{2.15}$$

The magnitude of $g(\tau)$ cannot exceed unity, $0 \le |g(\tau)| \le 1$. When $|g(\tau)|$ decreases monotonically with τ, the value at which it drops to a defined value, in general 1/2 or 1/e, is known as coherence time τ_c. The coherence length l_c of the source is defined as the distance covered by the light during the time τ_c in free space, that is

$l_c = c_0 \cdot \tau_c$. Using the Wiener-Khinchin theorem [2.23], it is easy to relate the complex degree of temporal coherence $g(\tau)$ to the power spectral density of the source $S(\nu)$ by

$$g(\tau) = \frac{\int_0^\infty d\nu S(\nu) \exp[i2\pi\nu\tau]}{\int_0^\infty d\nu S(\nu)}. \tag{2.16}$$

Thus the power spectral density of the source and the complex degree of temporal coherence form a Fourier transform pair. Following Eq. (2.14), a typical white-light fringe signal is depicted in Fig. 2.2, where we assumed a Gaussian shape for the power spectral density and $A_1 = A_2 = 1$.

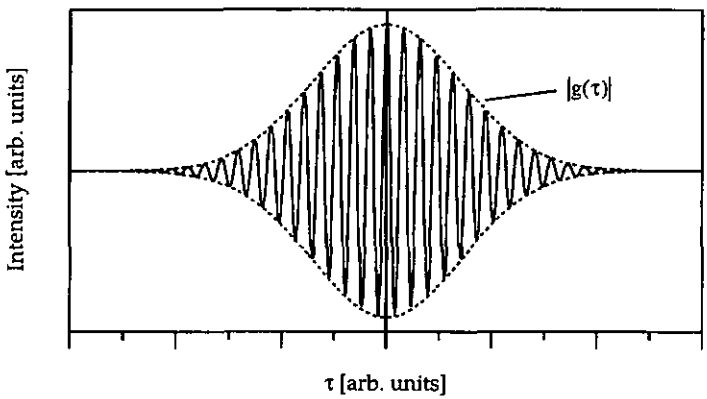


Fig. 2.2 Typical white-light interference signal.

2.3 Interferometric techniques

As already mentioned, interferometric measurements are made, in most cases, by interfering two optical beams so that the phase difference φ between the two beams is proportional to the quantity being measured. This class of optical phase-sensitive instruments uses the wavelength of the light source as the length standard. In fact, the problem behind precise phase measurement seems clear: following Eq. (2.11) we detect the optical intensity I (interference signal) and we want to determine the phase difference φ. Although this problem appears simple,

many difficulties must be overcome to get a reliable and outstanding interferometric measurement system.

There are two basic tasks in building a interferometric sensor: (a) design the transducer so that the measured phase is proportional to the desired physical parameter and does not depend on other physical quantities, and (b) devise an appropriate mean for processing the measured data from the transducer. The art of the engineer consists of imaging a measurement system which fulfills the requirements, keeping in mind that the optics and mechanical components must be small, simple and rugged, and to embody whatever complexity is needed to process the interference signal.

2.3.1 One-wavelength interferometry

Let us consider the classical Michelson interferometer, illuminated by a highly coherent source, as sketched in Fig. 2.3.

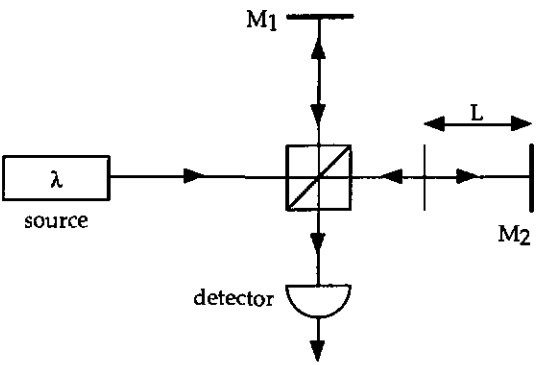


Fig. 2.3 Michelson interferometer.

From Eq. (2.11), the interference signal is given by

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\varphi) \tag{2.17}$$

where I_1 and I_2 are the intensity of the two individual interfering beams. The interference phase φ depends directly on the path difference L and can be expressed by

$$\varphi = n\frac{2\pi}{\lambda}2L, \tag{2.18}$$

where n is the index of refraction of the medium. It can be seen from Eqs. (2.17) and (2.18), that the interference signal is a periodic function of the distance L with a period equal to $\lambda/2$ (n = 1). This establishment leads to an important drawback of one-wavelength interferometers. The non-ambiguity range is limited to a phase difference of $\varphi = \pi$, or in other words, these measurement systems have a position ambiguity of $m\cdot\lambda/4$, where m is an integer. If the mirror displacement is greater than the non-ambiguity range, the distance can be determined by counting the number of periods. Moreover, when an optical system is used to produce two interference signals which are in quadrature, an accurate estimate of the fractional fringe order can be made [2.24]. This is why these kind of interferometric sensors, which use monochromatic sources, are widely employed as displacement sensors (length measurement), for vibration measurements and also in the field of interferometric profilometry to inspect sample surface for quality control purposes.

As we will see later (chapter 3), heterodyne techniques are very useful for the detection at very low-level of light intensity. In one method, a frequency difference between the two interfering waves is introduced by a pair of acousto-optic modulators operating at slightly different frequencies. For this situation, the two interfering beams can be expressed by their complex wavefunctions

$$U_1 = V_1 \exp[i2\pi\nu_1 t] = V_{10}\exp[i(2\pi\nu_1 t + \varphi_1)], \tag{2.19}$$

and

$$U_2 = V_2 \exp[i2\pi\nu_2 t] = V_{20}\exp[i(2\pi\nu_2 t + \varphi_2)]. \tag{2.20}$$

Substituting Eqs. (2.19) and (2.20) into Eq. (2.4), the interference signal becomes

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos[2\pi(\nu_2 - \nu_1)t + \varphi] \tag{2.21}$$

where $I_1 = V_{10}^2$, $I_2 = V_{20}^2$ and $\varphi = \varphi_2 - \varphi_1$. We note that in heterodyne interferometry the phase difference φ is equal to the phase of a beat signal [see Eq. (2.21)], contrary to homodyne techniques where the phase difference φ is

translated into a sinusoidal variation of the intensity [see Eq. (2.17)]. This implicates several advantages of heterodyne techniques. Indeed, the problem of the phase interpolation can be solved easier, and the sign of φ can also be measured. This means that using heterodyne interferometry, the direction of the movement can be determined. Moreover, the interference term is an ac component and it is thus well separated from the other dc terms in Eq. (2.21) and from low-frequency noise sources.

We would like to remark that the amplitude and the phase of an optical wave reflected from a rough surface have statistical behavior (speckle effect). This is why interferometry using monochromatic source is not suitable when operating with non-cooperative targets. This represents a serious drawback for the application of such measurement systems in an industrial environment.

2.3.2 Multiple-wavelength interferometry

Multiple-wavelength interferometry is, as classical interferometry, a coherent method, but it offers great flexibility in sensitivity by an appropriate choice of the different wavelengths. Indeed, as we will see soon, the use of two different wavelengths, λ_1 and λ_2, permits to generate a synthetic wavelength $\Lambda = \lambda_1 \cdot \lambda_2 / |\lambda_1 - \lambda_2| = c_0/\Delta\nu$, where $\Delta\nu = \nu_2 - \nu_1$, much longer than the individual optical wavelengths [2.25, 2.26, 2.27]. This method thus makes it possible to increase the range of non-ambiguity for interferometry and to reduce the sensitivity of the measurement. Moreover, this technique is also applicable to rough surfaces [2.17, 2.28].

Two-wavelength heterodyne interferometry was reported by Fercher et al. [2.28]. By simultaneous phase measurements at both wavelengths, the interference phase at the synthetic wavelength Λ can be directly determined. Figure 2.4 shows the optical setup of a Michelson two-wavelength interferometer. At the output of the interferometer, the two wavelengths are first optically separated (grating, prism) and then individually detected. For each wavelength λ_i, i = 1 and 2, an interference signal of the form

$$I_i = I_{0i}[1 + m_i \cos\varphi(\lambda_i)] \tag{2.22}$$

is detected, where m_i is the relative interference amplitude. Because the two wavelengths must be optically separated before detection, the technique can be

used only for relatively large wavelength differences and thus small synthetic wavelengths. Moreover, interferometric stability at the optical wavelengths is required.

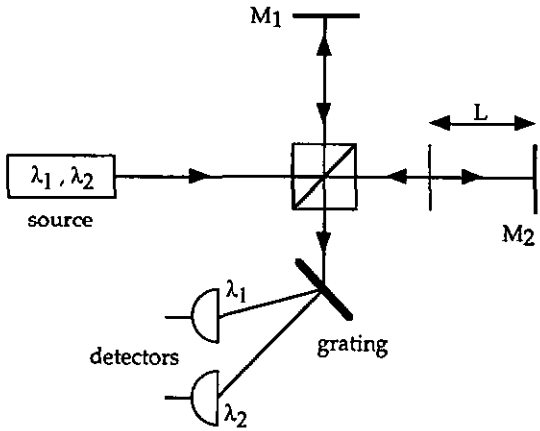


Fig. 2.4 Two-wavelength Michelson interferometer: the two wavelength are optically separated.

After processing, the phase $\varphi(\lambda_i)$ modulo 2π is obtained. Then the phase difference $\Delta\varphi = \varphi(\lambda_2) - \varphi(\lambda_1)$ is calculated, namely

$$\Delta\varphi = \varphi(\lambda_2) - \varphi(\lambda_1) = \frac{2\pi}{\Lambda} 2L. \tag{2.23}$$

The remarkable fact is that $\Delta\varphi$ is now only sensitive to the synthetic wavelength Λ given by

$$\Lambda = \frac{\lambda_1\lambda_2}{|\lambda_1 - \lambda_2|} = \frac{c_0}{\nu_2 - \nu_1}. \tag{2.24}$$

Modern laser sources provide great variety of optical frequencies or wavelengths to get interesting synthetic wavelengths. Thus, multiple-wavelength interferometry is a promising technique for absolute distance measurement [2.29, 2.30].

As shown in Fig. 2.5, an alternative method can be used. A two-wavelength source emitting at two different optical frequencies ν_1 and ν_2, corresponding to the wavelengths λ_1 and λ_2 are used to illuminate the interferometer. The light at the output of the interferometer is directed toward a photodiode. The interference signal is now given by

$$I(L) = A_0 + A_1 \cos[\varphi(\lambda_1)] + A_2 \cos[\varphi(\lambda_1)], \tag{2.25}$$

where A_0, A_1 and A_2 are constants. This signal corresponds simply to the incoherent superposition of the two interference signals for the wavelengths λ_1 and λ_2, with the corresponding interference phase $\varphi(\lambda_1)$ and $\varphi(\lambda_2)$. As previously, for an interferometric path difference L, the phases $\varphi(\lambda_1)$ and $\varphi(\lambda_2)$ are given by $\varphi(\lambda_1) = 4\pi L/\lambda_1$ and $\varphi(\lambda_2) = 4\pi L/\lambda_2$.

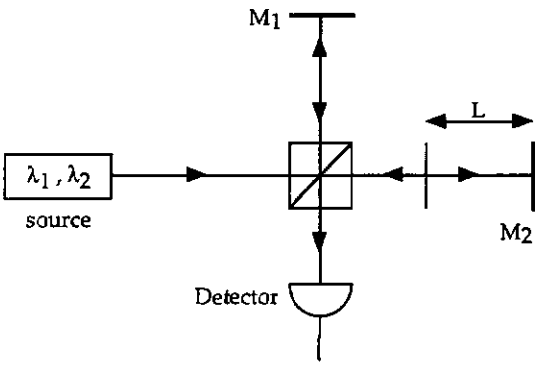


Fig. 2.5 Two-wavelength Michelson interferometer

The signal described by Eq. (2.25) is depicted in Fig. 2.6. The interference fringe function, that is the envelope of this interference signal (modulation power), is then obtained by detecting the electrical power of the ac part of this signal, which is

$$P_{ac} = \left(A_1^2 + A_2^2\right) \cdot \left\{1 + \frac{2 \cdot A_1 \cdot A_2}{A_1^2 + A_2^2} \cdot \cos[\varphi(\lambda_2) - \varphi(\lambda_1)]\right\}. \tag{2.26}$$

The phase of this signal $\Delta\varphi$ is given by

$$\Delta\varphi(L) = \varphi(\lambda_2) - \varphi(\lambda_1) = 2\pi\left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right)\cdot 2L = \frac{2\pi}{\Lambda}2L, \tag{2.27}$$

and is again only sensitive to the synthetic wavelength Λ.

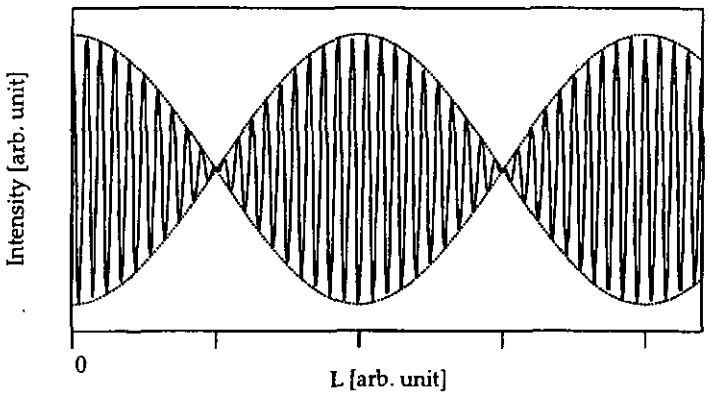


Fig. 2.6 Interference signal resulting from the superposition of two beams of slightly different frequencies.

This phase $\Delta\varphi$ can be now determined by techniques similar to those used for phase interpolation in classical interferometry (one-wavelength). For example, super-heterodyne detection [2.31, 2.32] and heterodyne detection [2.33] enables high-resolution at arbitrary synthetic wavelength without the need for interferometric stability or optical separation of the wavelengths λ_1 and λ_2 before detection.

2.3.3 White-light interferometry

An obvious extension of multiple-wavelength interferometry presented in the previous paragraph, is white-light interferometry or interferometry using low-coherence sources [2.34]. Indeed, over the last few years white-light interferometry, which uses broad-band light sources with short coherence lengths, has become an important technique to overcome the phase ambiguity from which suffer the interferometers using monochromatic sources [2.35, 2.36]. A great advantage of white-light interferometry is the possibility to make absolute measurements rather than incremental measurements. Moreover, the

mechanical stability of the measurement system can be increased by working with white-light channeled spectrum interferometry [2.13, 2.37] or when the interferogram is scanned in space rather than in time, and is projected onto a linear photodiode array (electronically scanned interferometer) [2.18, 2.38]. White-light interferometers have been successfully developed, among others, as absolute displacement sensors [2.11, 2.13], as strain or force sensors [2.19, 2.38], and for profilometry [2.16, 2.39].

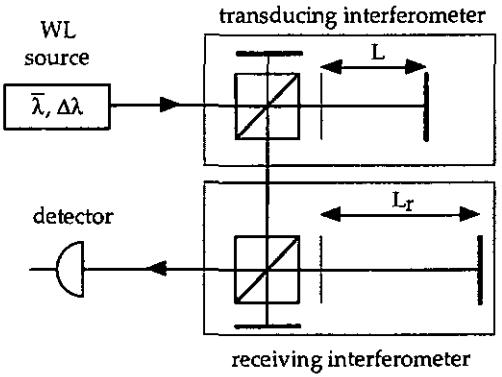


Fig. 2.7 White-light interferometer, mechanically scanned in the temporal regime.

Although different types of configuration exist, the basic arrangement is the same, and one commonly employed is built from two interferometers in tandem, as schematically drawn in Fig. 2.7. Here, both interferometers are two-beam Michelson interferometers. The first one acts as the transducing interferometer, in which the distance L is much greater than the coherence length l_c of the light source. The aim of this interferometer is to convert the measurand into a change of the optical path difference. In the receiving interferometer, the large optical path difference induced by the transducing interferometer is compensated by the optical path difference of the receiving interferometer.

A typical white-light interference signal is shown in Fig. 2.8. A first set of interference fringes is observed when the two arms of the receiving interferometer have equal lengths, that is when $L_r = 0$, which determined the internal absolute zero position. A second set of fringes is also observed when the receiving interferometer compensates the optical path difference induced by the

transducing interferometer, that is when $L_r = L$. The envelope of each set of fringes is given by the modulus of the complex degree of temporal coherence, which is proportional the Fourier transform of the power spectrum density of the source (see § 2.2.2). Thus, by measuring the distance between the maxima of both sets of fringes, an absolute value for L can be determined. This involves the accurate determination of the center position of both sets of fringes. A novel and simple signal processing technique will be presented in chapter 4.

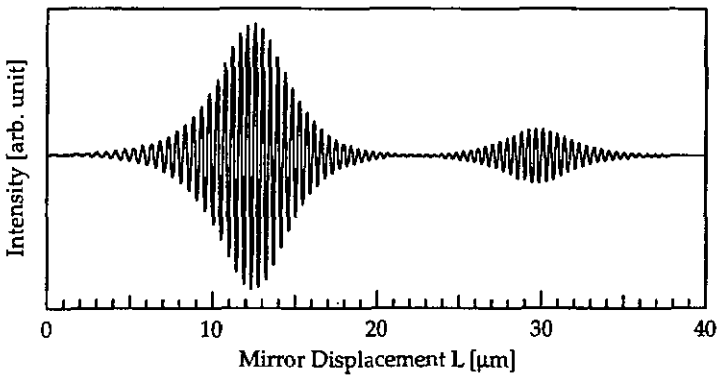


Fig. 2.7 Measured white-light interference signal using an optical setup similar to that sketched in Fig. 2.6.

Several practical sensor configurations can be conceived, either for the transducing or the receiving interferometer [2.35]. But we would like to mentioned two major improvements with respect to the basic configuration described above. Electronically scanned white-light interferometry has been used since several years in the domain of optical fiber sensors. Young's interferometers [2.18], as well as Michelson's interferometers with one mirror tilted [2.11], or Wollaston prisms as phase compensator [2.38] were applied as receiving interferometer. In these examples, the interferogram is scanned electronically using a linear photodiode array in the spatial regime. It should be noted that if mechanical scanning suffers from limited speed, spatial electronic scanning is limited to small operating range. To avoid these drawbacks, white-light channeled spectrum (or dispersive) interferometers has be proposed as optical profilometer [2.37] and for absolute distance measurement [2.38]. Here, the output light of the transducing interferometer passes through a grating and the diffracted light is then focused on a linear photodiode array to scan the interference signal.

3. Full-dynamic phase-demodulator for heterodyne interferometer

Most parts of this chapter have been published as regular papers in international journals or presented at conferences [3.1, 3.2].

Since the development of the first lasers in the early 1960's, laser interferometry has become an important technique to provide high resolution and high accuracy measurement systems, at once for scientific applications but also in the industrial framework. One crucial issue for laser interferometry is related to the electronic treatment and the analysis of the interference signal. Indeed, the measurement accuracy depends strongly on the signal processing unit (hardware and/or software) employed to recover the phase information from the interference signal. Moreover, this signal processing unit will depend itself on the interferometric technique used. Thus, a compromise between the complexity of the system (optical setup, signal processing unit, etc.) and the measurement accuracy has to be found to finally reach the required specifications of the measurement system.

The interferometric measurement techniques can be classified in two broad categories: the static and dynamic methods [3.3]. The static techniques, also called homodyne, work with only one optical frequency for the interfering beams. The dynamic techniques are based on the principle of shifting the phase between the interfering waves. In the phase-shifting (quasi-heterodyne) technique, the optical path difference between the interfering beams is varied either stepwise (phase-stepping method) or linearly with time (integrating-bucket method) [3.4]. A great advantage of the quasi-heterodyne methods is that a detector array can be used to make measurements simultaneously at a very large number of points covering the interference pattern. In heterodyne interferometry, a frequency offset between the two interfering beams is introduced with the help, for example, of acousto-optic modulators. Thus, in heterodyne interferometry, the optical phase difference to be measured corresponds to the phase of the beat signal at the heterodyne frequency (see chapter 3.1). We would like to mention that until recently, heterodyne interferometry was not compatible with standard CCD array detectors. Indeed, this technique requires synchronous detection of the heterodyne frequency. To solve this problem, novel smart CCD arrays where each

pixel acts as an independent synchronous detector were recently developed [3.5]. This lock-in CCD image sensor has been successfully implemented in a multiple-wavelength interferometer for applications with non-cooperative targets [3.6]

Among the different techniques used to measure vibrations in the inner ear, previous studies [3.7, 3.8] show, that an optical heterodyne measurement system is most adequate. But the measurement of vibrations in the inner ear imposes specific constraints to the interferometer. Indeed, the objects whose vibrations are to be measured are very small (5 µm to 15 µm) and the vibration amplitudes are very low (tens of pm). In addition, the linearity of the interferometer must be high in order to measure the nonlinear behavior of the mechanical structure of the inner ear. To fulfill these requirements, a heterodyne interferometer measurement system was specially designed [3.7, 3.8]. In this previous development the output signal of the detector is demodulated using a frequency-demodulator to recover the modulation, which is proportional to the instantaneous velocity of the vibrating object. Unfortunately, with a FM-demodulator, the displacement sensitivity decreases with frequency. Although tolerable for vibration frequencies greater than a few hundreds of Hz, this limitation is a severe drawback for studying vibrating objects at low frequencies. Moreover, for nearly transparent cells the amount of back-scattered light becomes very small [3.9], resulting in a poor carrier-to-noise ratio (CNR) of the detected interference signal. And when the CNR decreases, that is when the noise at the demodulator input increases, the FM-demodulator ceases to work properly. This phenomena is the well known "threshold effect". It can be shown that when the CNR falls below 10 dB for the full receiver bandwidth, the signal-to-noise ratio at the demodulator output shows a rapid deterioration [3.10]. This threshold determines the minimum signal quality (CNR for a fixed bandwidth) for which the FM-demodulator works properly. Since many drawbacks of the measurement system based on FM-demodulation are given by the demodulator itself, we decided to develop a novel demodulator based on phase-demodulation rather than frequency-demodulation.

In this chapter, a novel full-dynamic phase-demodulator for heterodyne interferometric vibration measurements is described. Contrary to a conventional phase-demodulator, which has a limited dynamic range of ± 90°, the proposed phase-demodulator has an unlimited dynamic range. This allows to recover phase information for displacements of the surface under test which are much greater than the wavelength of light.

After a brief description of the principle of operation of the heterodyne interferometer, the noise equivalent phase for frequency-demodulation and phase-demodulation will be determined and compared. Then, we will see how heterodyne detection offers an elegant way to increase the CNR up to a point where the shot-noise dominates the thermal noise of the electronics. Next, a complete description of the full-dynamic phase-demodulator will be presented. Specially, we will see that by adopting a digital approach to phase-demodulation, a high throughput speed up to one half of the heterodyne (carrier) frequency can be obtained. The signal processing employed to recover the phase will also be introduced in this paragraph. Finally, after a complete noise characterization of the demodulator, several examples will illustrate how well our vibration measurement system can be used to study complex movements of objects.

3.1 Heterodyne interferometer

3.1.1 Principle of operation and optical setup

The principle of the heterodyne interferometer employed to investigate our vibrations measuring system is depicted in Fig. 3.1.

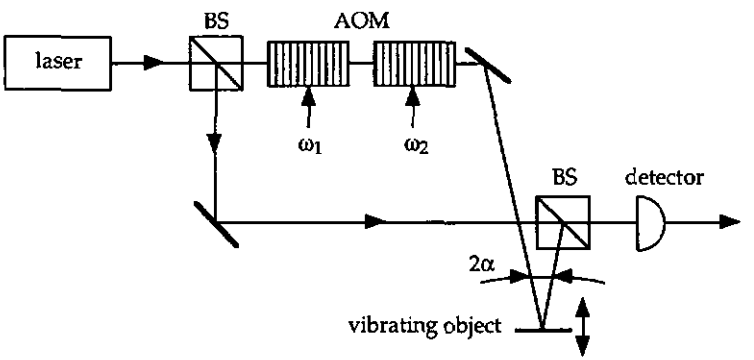


Fig. 3.1 Schematic of a heterodyne interferometer. BS: beamsplitter, AOM: acousto-optic modulator.

It is quite simple in concept, a HeNe laser emitting at a wavelength of $\lambda = 633$ nm is used as a coherent light source. Utilizing two acousto-optic modulators operating at $\omega_1/2\pi = 40$ MHz and $\omega_2/2\pi = 40.455$ MHz, a heterodyne frequency

shift of Δf_h = 455 kHz between the two interfering optical frequencies is obtained. It is important to note that by this way, the heterodyne frequency is only presents optically, which is very significant for highly sensitive system to avoid electrical interferences at the heterodyne frequency. Moreover, the heterodyne frequency can be easily modified by changing the operating frequency of the acousto-optic modulators.

The interference signal is detected using a simple photodiode followed by a transimpedance amplifier (see § 3.1.2). The detector bandwidth has to be large enough to resolve the heterodyne frequency. From Eq. (2.21), the detector output voltage is then given by

$$V(t) = a + b\cos[\Delta\omega t + \phi(t)], \tag{3.1}$$

where a is the dc and b the ac amplitude of the interference signal, and $\Delta\omega = \omega_2 - \omega_1 = 2\pi\Delta f_h$. The interference phase $\phi(t)$ of such a beat signal depends on the path difference between the two interfering beams and thus on the position of the vibrating object. For a periodic movement, the interference phase $\phi(t)$ in Eq. (3.1) is of the form

$$\phi(t) = \beta u\cos(\Omega t + \Phi) + \phi_c, \tag{3.2}$$

where u is the amplitude, $\Omega/2\pi$ is the frequency and Φ is the phase of the vibration. ϕ_c is a constant phase given by the geometry of the experimental setup. The interferometric sensitivity β is given by

$$\beta = 2\left(\frac{2\pi}{\lambda}\right)\cos\alpha, \tag{3.3}$$

where 2α is the angle between the illumination and observation directions as shown in Fig. 3.1. At normal incidence and for λ = 633 nm, we find an interferometric sensitivity of $\beta \cong 20\ \mu m^{-1}$. Substituting Eq. (3.2) into (3.1), the interference signal is then written as

$$V(t) = a + b\cos[\Delta\omega t + \beta u\cos(\Omega t + \Phi) + \phi_c]. \tag{3.4}$$

Equation (3.4) shows that the detector output is a phase- or frequency-modulated signal with a carrier frequency $\Delta\omega$ and a modulation frequency Ω. The frequency spectrum of this signal is a well-known discrete Bessel spectrum [3.11]. For small vibration amplitudes, this spectrum consists mainly of three lines: the carrier and two sidebands at $\pm\,\Omega$, as shown is Fig. 3.2.

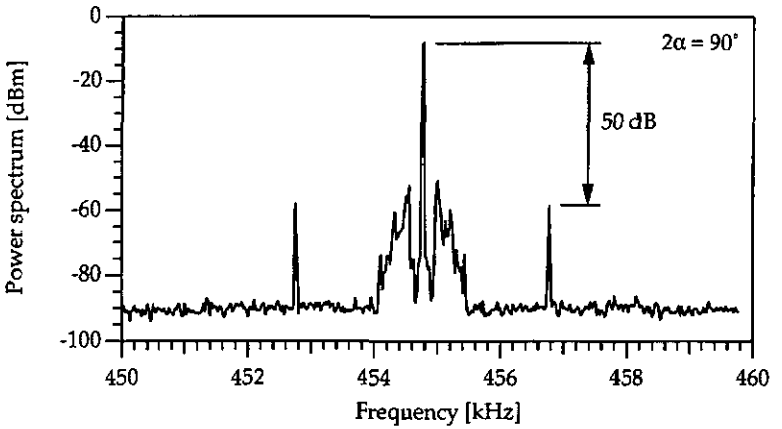


Fig. 3.2 Measured spectrum corresponding to Eq. (3.4) for a small vibration amplitude of approximately 0.45 nm (β = 14 μm^{-1}) at 2 kHz (resolution bandwidth = 30 Hz).

Thus, the vibration amplitude u can easily be deduced from the power ratio of the carrier P_0 and the first sideband P_1 with the relation [3.7]

$$\sqrt{\frac{P_0}{P_1}} = \frac{J_0(\beta u)}{J_1(\beta u)} \cong \frac{2}{\beta u}, \tag{3.5}$$

where J_0 and J_1 are the Bessel functions of the first kind of order 1 and 2, respectively. Of course, this method is only valid for sinusoidal vibrations. For example, the spectrum of Fig. 3.2 corresponds to a vibration amplitudes of approximately 0.45 nm at a frequency of 2 kHz. You can also observed the low frequency noise (between 100 Hz and 700 Hz) due to external perturbations (mechanical noise). The noise floor is at –70 dBm for a CNR = 62 dB (measured within 30 Hz bandwidth, converted to 3 kHz bandwidth), equivalent to an amplitude of 0.1 nm.

By comparison to homodyne interferometry, heterodyne interferometry has several advantages [3.3, 3.7]. First, the linearity of heterodyne technique is not limited to small vibration amplitudes. Indeed, the phase variations are not converted to intensity variations [see Eq. (2.11)] but to phase variations of the beat signal [see Eq. (3.1)]. Second, the familiar quadrature condition that has to be maintained for homodyne techniques is not necessary with heterodyning, because the sensitivity is independent of the phase difference between the two interfering beams. Therefore, drifts taking the interferometer away from the quadrature condition do not cause signal fading. Moreover, the interference term is an ac component and it is thus well separated from the other dc terms of Eq. (3.1) and from low-frequency noise sources.

From Eqs. (3.1) and (3.2), the frequency of the interference signal can be calculated by

$$f(t) = \frac{\Delta\omega}{2\pi} - \beta u \frac{\Omega}{2\pi}\sin(\Omega \cdot t + \psi) = \Delta f_h - \frac{\beta}{2\pi} v(t), \tag{3.6}$$

where $\Delta f_h = \Delta\omega/2\pi$ and $v(t)$ is the velocity of the surface, given by

$$v(t) = u\Omega\sin(\Omega \cdot t + \psi). \tag{3.7}$$

From these equations it is obvious that if we consider as modulating signal the position of the surface, the interference signal given by Eq. (3.1) must be seen as a phase modulated signal, but if we consider as modulating signal the velocity of the surface, the interference signal must be seen as a frequency modulated signal. We emphasize again that for a FM-demodulator the displacement sensitivity decreases with vibration frequency Ω, as explicitly given by Eq. (3.6).

The experimental setup utilized to analyze our full-dynamic phase-demodulator is depicted in Fig. 3.3. In this setup, the vibrating object is a mirror mounted on a piezoelectric translator (PZT). Contrary to FM-demodulation, phase-demodulation needs a stable reference signal. It is also created optically, using the beam-splitters BS_3 to BS_5 shown in Fig. 3.3. Another way to produce the reference signal would be to mix electronically the two driving signals of the acousto-optic modulators at ω_1 and ω_2 and to extract the reference signal at the heterodyne frequency with a low-pass filter. But, as we will see in paragraph 3.4, using a reference signal created optically instead of electronically increases the accuracy of

the measurement. Indeed, the optical setup sketched in Fig. 3.3 must be considered as two nearly identical interferometers, with a long common optical path. This leads to a very good "common mode (path) rejection", to employ electronic terminology.

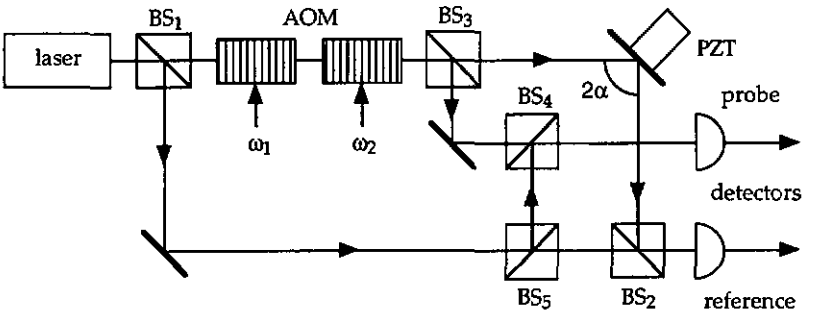


Fig. 3.3 Experimental setup for testing the phase-demodulator.

3.1.2 Transimpedance amplifier

At the interferometer output, the light is detected with a standard pin photodiode (BPW 34) followed by a transimpedance amplifier (high speed operational amplifier Burr-Brown OPA 602, transition frequency = 6.5 MHz). Indeed, the transimpedance amplifier is usually the amplifier of choice for the photodetection circuit. It is sketched in Fig. 3.4 and has the function to convert the photocurrent generated by the photodiode to a more practical output voltage.

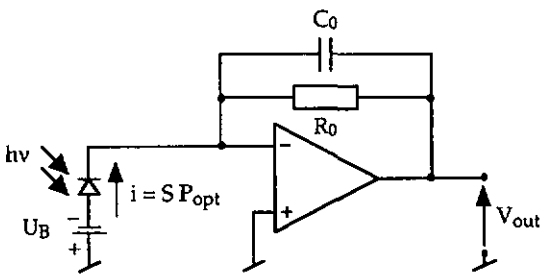


Fig. 3.4 Detection circuit consisting of a standard pin photodiode followed by a transimpedance amplifier.

If the optical power P_{opt} is modulated at a frequency $\omega/2\pi$, the output voltage $V_{out}(\omega)$ of the transimpedance amplifier can be expressed by

$$V_{out}(\omega) = \Re(\omega) \cdot S \cdot P_{opt}(\omega) \tag{3.8}$$

where $\Re(\omega)$ is the frequency response of the detection circuit and S is the spectral sensitivity [A/W] of the photodiode. Taking into account the photodiode capacitance C_d and the amplifier input capacitance C_a, the frequency response $\Re(\omega)$ may be written as

$$\Re(\omega) = -\frac{R_0}{1 + j\,2\,\xi \frac{\omega}{\omega_0} - \left(\frac{\omega}{\omega_0}\right)^2} = |\Re(\omega)| \exp[i\chi(\omega)], \tag{3.9}$$

where we use the definition

$$\omega_0 = \sqrt{\omega_T \frac{\omega_d\,\omega_a}{\omega_d + \omega_a}}, \tag{3.10}$$

with

$$\omega_a = \frac{1}{R_0 C_0} \quad \text{and} \quad \omega_d = \frac{1}{R_0 C}. \tag{3.11}$$

In Eqs. (3.9) to (3.11), $\omega_0/2\pi$ is the undamped resonance frequency or eigen frequency of the system, $\omega_T/2\pi$ is the transition frequency of the amplifier and $C = C_d + C_a$. The damping factor ξ is given by

$$\xi = \frac{\omega_0}{2\omega_T} + \frac{\omega_0}{2\omega_a}. \tag{3.12}$$

Equation (3.9) corresponds to the frequency response of a 2nd-order system. In Fig. (3.5), the amplitude and phase of the normalized frequency response $\Re(\omega)/R_0$ is plotted as a function of the normalized frequency ω/ω_0 for various damping factors ξ. From this figure we can see that the low frequency gain is $-R_0$, as obviously expected from the schematic of Fig. 3.4.

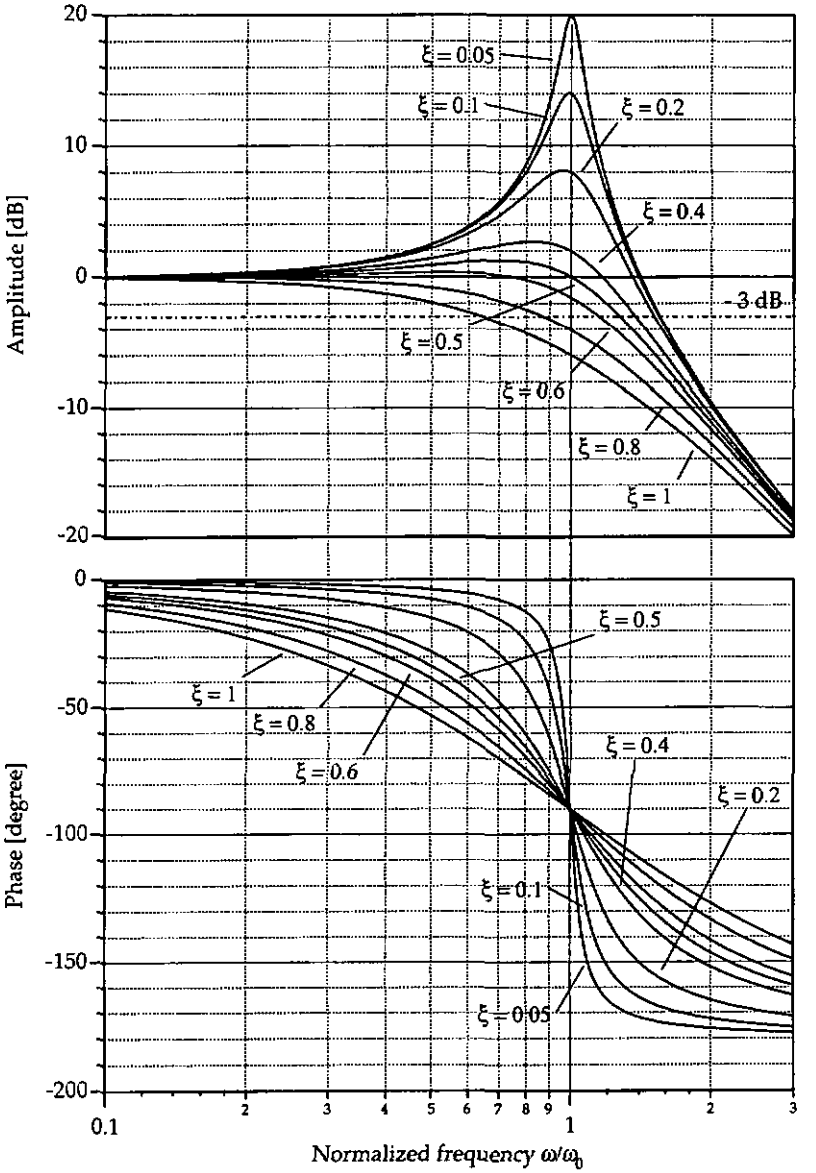


Fig. 3.5 Normalized frequency response of a 2nd-order system.

The 3-dB bandwidth B of such a 2nd-oder system is given by

$$B = \frac{\omega_0}{2\pi}\sqrt{1-2\xi^2+\sqrt{\left(1-2\xi^2\right)^2+1}}. \tag{3.13}$$

From Eqs. (3.12) and (3.13), we can see that if C_d is reduced by increasing the reverse voltage U_B applied to the photodiode, the damping factor ξ and the bandwidth B are increased.

When the amplitude response $|\Re(\omega)|$ of the transimpedance amplifier is not constant with frequency inside the frequency band of interest around the carrier frequency (455 kHz), the frequency components of the interference signal are detected with different gain. This effect is called amplitude distortion and results in an amplitude modulation of the detected signal superposed to the frequency modulation. But, because the input stage of the phase demodulator is composed of limiter amplifiers which clip the amplitude of the input signal (see § 3.3.1.1), the amplitude distortion does not significantly influence the performances of the phase demodulator.

The second form of distortion arises when the phase of the frequency response $\chi(\omega)$ is not linear with frequency (see Fig. 3.5). Indeed, when the interference signal is composed of a set of components, each of them is subject to a different delay when passing through the transimpedance amplifier. Thus, the output voltage has a different waveform from the optical input signal. This form of distortion is called phase distortion and influences directly the output signals of the phase demodulator.

The group delay $\tau_g(\omega)$ is proportional to the slope of the phase of the frequency response, namely

$$\tau_g(\omega) = -\frac{d}{d\omega}\left[\chi(\omega)\right]. \tag{3.14}$$

In Fig. 3.6, the group delay τ_g of the normalized frequency response $\Re(\omega)/R_0$ is plotted versus the normalized frequency ω/ω_0 for various damping factors ξ. In fact, the phase distortion results from the variations of the group delay inside the frequency band of interest. In consequence, the transimpedance amplifiers were designed with a damping factors of approximately $\xi = 0.8$. Indeed, the group delay

fluctuations are minimized with this damping factor, resulting in a small amount of phase distortion at the transimpedance outputs.

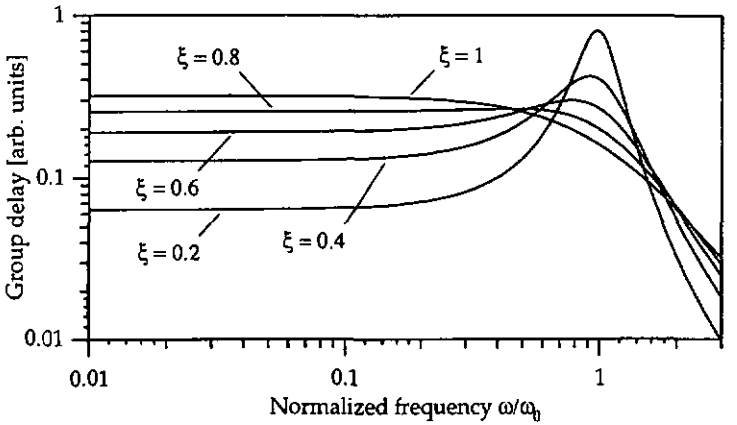


Fig. 3.6 Group delay of the transimpedance amplifier.

Substituting Eqs. (3.10) and (3.11) into Eq. (3.12), the feedback capacitance C_0 as a function of the damping factor ξ is expressed by

$$C_0 = \frac{2\xi^2 - 1 + 2\xi\sqrt{\xi^2 - 1 + \frac{\omega_T}{\omega_d}}}{\omega_T R_0}. \tag{3.15}$$

Using Eq. (3.15), we can calculate the feedback capacitance C_0 to obtain any damping factor ξ. This allows to adapt the frequency response $\Re(\omega)$ of the optoelectronic detection circuit to minimize the phase distortion.

3.2 Noise consideration and fundamental limits

3.2.1 Noise equivalent phase for PM- and FM-demodulation

The basic concept of the delay-line frequency demodulator used in previous experiments [3.7, 3.8] is shown in Fig. 3.7. It consists essentially of a limiter amplifier, a delay line to get a 90° or 270° phase shift of the carrier frequency, a mixer, and a low-pass filter to eliminate the carrier frequency after mixing. The

demodulator output is then proportional to the phase difference $\Delta\phi = \phi(t+\tau) - \phi(t)$, where τ is the delay, at least as long as $|\Delta\phi| < 90°$. This is a good approximation for the frequency deviation $d\phi/dt \cong \Delta\phi/\tau$.

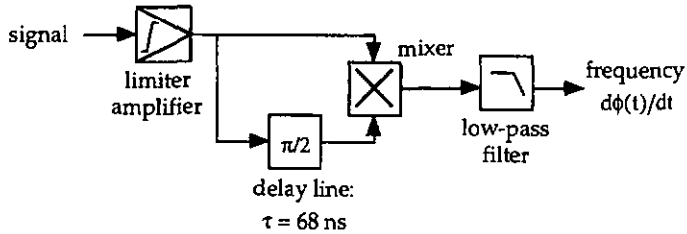


Fig. 3.7 Delay-line frequency-demodulator (Revox B760) [3.7, 3.8].

As shown in Fig. 3.8, the basic concept of a phase-demodulator looks very similar to the delay-line frequency-demodulator, except for the fact that now the phase of the signal is compared with the constant phase of a reference signal at the carrier frequency. Note that contrary to a FM-demodulator, the phase-demodulation requires a stable reference signal.

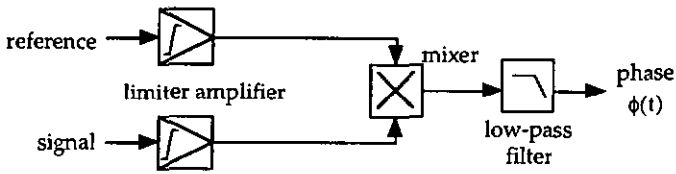


Fig. 3.8 Phase-demodulator.

Before going into the details of the noise equivalent phase for frequency- and phase-demodulators, we would like to specify that two distinct situations must be considered. For high values of the carrier-to-noise ratio (CNR) the internal demodulator noise will dominate; the output noise level is then independent of the CNR. For low values of the CNR the output noise is essentially due to the applied input noise. Therefore, the output noise decreases as the CNR increases.

To determine the noise equivalent phase $\delta\varphi^{PM}_{noise}$ caused by the noise of the input signal, we consider the effect of a noisy component n(t) on a harmonic signal

$s(t) = \sqrt{2}S_0 \sin(2\pi ft)$. The phasor representation of such a noisy signal is depicted in Fig. 3.9.

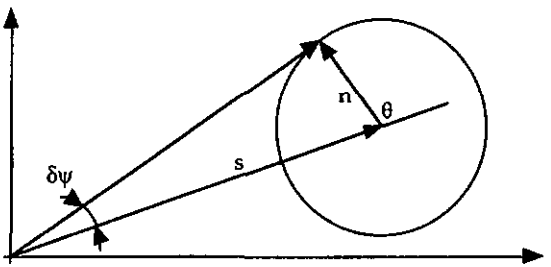


Fig. 3.9 Phasor model of a noisy harmonic signal.

In this situation, the CNR is given by

$$CNR = \frac{\langle s^2(t) \rangle}{\langle n^2(t) \rangle} = \frac{S_0^2}{\langle n^2(t) \rangle}. \tag{3.16}$$

The noise component n(t) is assumed to be random, following a Gaussian distribution, with $\langle n(t) \rangle = 0$ and $\langle n^2(t) \rangle = N^2$. The phase θ is uniformly distributed between 0 and 2π. Assuming that $s(t) >> n(t)$, the resulting phase error $\delta\psi(t)$ becomes

$$\delta\psi(t) = \frac{n(t)\sin\theta(t)}{S_0}. \tag{3.17}$$

Finally, and assuming that the CNR of the reference is much higher than the CNR of the signal, we find for the noise equivalent phase $\delta\varphi_{noise}^{PM}$ in the case of phase-demodulation

$$\delta\varphi_{noise}^{PM} = \sqrt{\langle \delta\psi^2(t) \rangle} = \frac{1}{\sqrt{2 \cdot CNR}}. \tag{3.18}$$

Equation (3.18) relates directly the phase error to the CNR at the PM-demodulator input. Using Eq. (3.2) and the condition $\langle \phi^2(t) \rangle = \langle \delta\psi^2(t) \rangle$, the corresponding noise equivalent vibration amplitude is found to be

$$u_{noise}^{PM} = \frac{\sqrt{2}\,\delta\varphi_{noise}^{PM}}{\beta} = \frac{1}{\beta\sqrt{CNR}}. \tag{3.19}$$

For a FM-demodulator, the noise equivalent amplitude for a detection limited by the CNR is given by

$$u_{noise}^{FM} = \frac{\sqrt{2}}{\beta\sqrt{CNR}}. \tag{3.20}$$

Compared with Eq. (11) in Ref. [3.7], we have taken into account that the phase is only affected by the quadrature noise component of the input signal [3.10] The remaining factor of $\sqrt{2}$ between Eqs. (3.19) and (3.20) is due to the fact, that in the frequency-demodulator two phase values of the noisy signal are compared, whereas in the phase-demodulator the reference signal is considered noise-free.

On the other hand, for high CNR the noise equivalent amplitude depends on the demodulator noise itself. Thus, from Eq. (3.2), we get for the phase-demodulator a noise equivalent amplitude

$$u_{noise}^{PM} = \frac{\sqrt{2}\,\delta\varphi_{noise}}{\beta}, \tag{3.21}$$

where $\delta\varphi_{noise}$ is the internal noise equivalent phase. From Eq. (3.2) the frequency deviation (Doppler shift) is given by

$$\frac{d\phi(t)}{dt} = -\beta u\Omega\sin(\Omega t + \Phi), \tag{3.22}$$

and from Eq. (3.22), we get for the frequency-demodulator a noise equivalent amplitude

$$u_{noise}^{FM} = \frac{\sqrt{2}\,\delta\varphi_{noise}}{\beta\Omega\tau} = \frac{\sqrt{2}\,2\pi\delta\nu_{noise}}{\beta\Omega}, \tag{3.23}$$

where τ is the delay of the delay-line frequency demodulator (τ = 68 ns for the Revox B760) and $\delta\nu_{noise}$ is the internal frequency noise of the demodulator. It can be assumed that both $\delta\varphi_{noise}$ or $\delta\nu_{noise}$ are independent of the vibration

frequency Ω. Therefore u_{noise}^{PM} is essentially independent of Ω for phase-demodulation, whereas u_{noise}^{FM} increases with decreasing vibration frequency.

Figure 3.10 shows the noise equivalent amplitude versus vibration frequency for frequency- and phase-demodulation assuming the same internal demodulator noise $\delta\varphi_{noise} = 0.1\mu rad/\sqrt{Hz}$. These value being those measured with the Revox B760 tuner used in previous experiments [3.7, 3.8] It is clear from this figure that for low frequencies the phase-demodulator is superior to the frequency-demodulator. For CNR = 40 dB (3 kHz bandwidth) the cross-over point is at a vibration frequency of approximately 1 kHz.

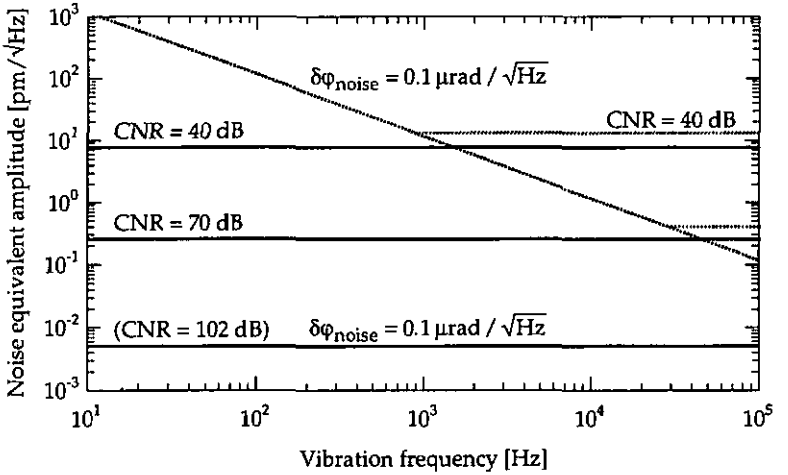


Fig. 3.10 Noise equivalent amplitude versus vibration frequency for frequency-demodulation (dotted line) and for phase-demodulation (plain line). The CNR is given for 3 kHz bandwidth and the interferometric sensitivity is $\beta = 20\ \mu m^{-1}$.

3.2.2 Shot-noise-limited heterodyne detection

Heterodyne detection offers an elegant way to increase the CNR up to a point where the shot-noise dominates the thermal noise of the electronics. In our analysis we considered two sources of noise, namely the shot-noise due to the quantum nature of light and the Johnson- or thermal-noise.

Considering the detector circuit depicted in Fig. 3.11, the electric power corresponding to the shot-noise at the detector output is given, for frequencies within the detection bandwidth B, by [3.12]

$$P_{sn} = \frac{(R_0 \cdot i_{sn})^2}{R} = 2eB\left(S\overline{P}_0 + i_d\right)\frac{R_0^2}{R}, \tag{3.24}$$

where e is the electron charge, S is the spectral sensitivity of the photo-detector, i_d the dark-current of the photodiode and $\overline{P}_0 = P_{RB} + P_{OB}$ is the sum of the light powers in the reference beam P_{RB} and object beam P_{OB} that reach the detector.

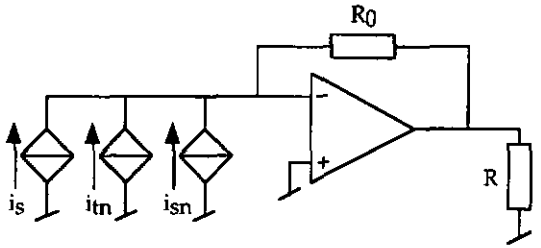


Fig. 3.11 Detector circuit including the different sources of noise. i_s: photo-current (signal), i_{tn}: current due to the thermal-noise and i_{sn}: current due to the shot-noise.

The contribution of the thermal-noise at the output of the transimpedance amplifier can be expressed as [3.12]

$$P_{tn} = \frac{(R_0 \cdot i_{tn})^2}{R} = 4kTB\frac{R_0}{R}, \tag{3.25}$$

where k is the Boltzmann constant and T the temperature. The electrical power of the modulated signal is

$$P_{ac} = 2m^2S^2P_{OB}P_{RB}\frac{R_0^2}{R}, \tag{3.26}$$

where m is the relative interference amplitude ($m \leq 1$). This factor measures the degree of temporal and spatial coherence between the two interfering beams [see Eq. (2.14)]. From Eqs (3.24) to (3.26), we get the carrier-to-noise ratio

$$CNR = \frac{P_{ac}}{P_{sn} + P_{tn}} = \frac{2m^2 S^2 P_{RB} P_{OB}}{2eB\left(S\bar{P}_0 + i_d\right) + \frac{4kTB}{R_0}}. \tag{3.27}$$

Neglecting the dark-current and assuming that m = 1, the CNR gets

$$CNR = \frac{2P_{RB}P_{OB}}{\frac{2eB}{S}\left(P_{RB} + P_{OB}\right) + \frac{4kTB}{SR_0}} \tag{3.28}$$

As shown in Fig. 3.12, it is always possible to increase the power in the reference beam P_{RB} until the shot-noise dominates the Johnson noise.

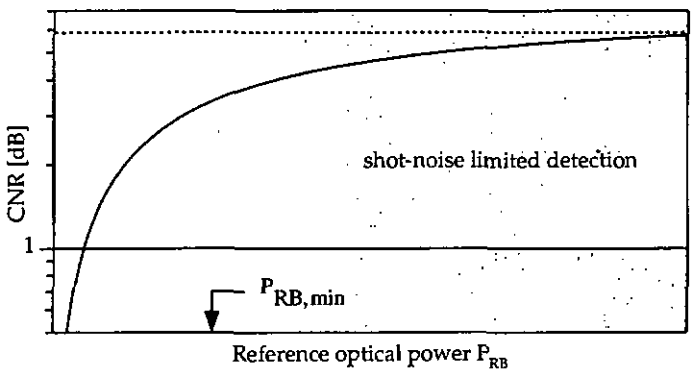


Fig. 3.12 Carrier-to-noise ratio versus optical power in the reference beam for heterodyne detection.

Using Eqs. (3.24) and (3.25) and assuming that $P_{RB} >> P_{OB}$, the minimum optical power in the reference beam to assure shot-noise limited detection is given by [3.7]

$$P_{RB,min} = \frac{2kT}{eR_0 S}. \tag{3.29}$$

The maximum CNR value that can be obtain for a given light power in the object beam is then

$$CNR_{max} = \frac{S}{eB} P_{OB}. \tag{3.30}$$

As an example, which corresponds to our laboratory setup, we assume a feedback resistance R_0 = 100 kΩ, a spectral sensitivity S = 0.6 A/W, a bandwidth B = 500 kHz and a temperature T = 300 K. With these values, we find a minimum optical power for the reference beam $P_{RB,min}$ = 0.9 µW and a maximum carrier-to-noise ratio CNR_{max} = 80 dB (3 kHz bandwidth) for P_{OB} = 100 nW.

Table 3.1 summarizes the different results obtained in this paragraph for frequency- and phase-demodulation. Specially, the advantage and drawback of both methods are highlighted.

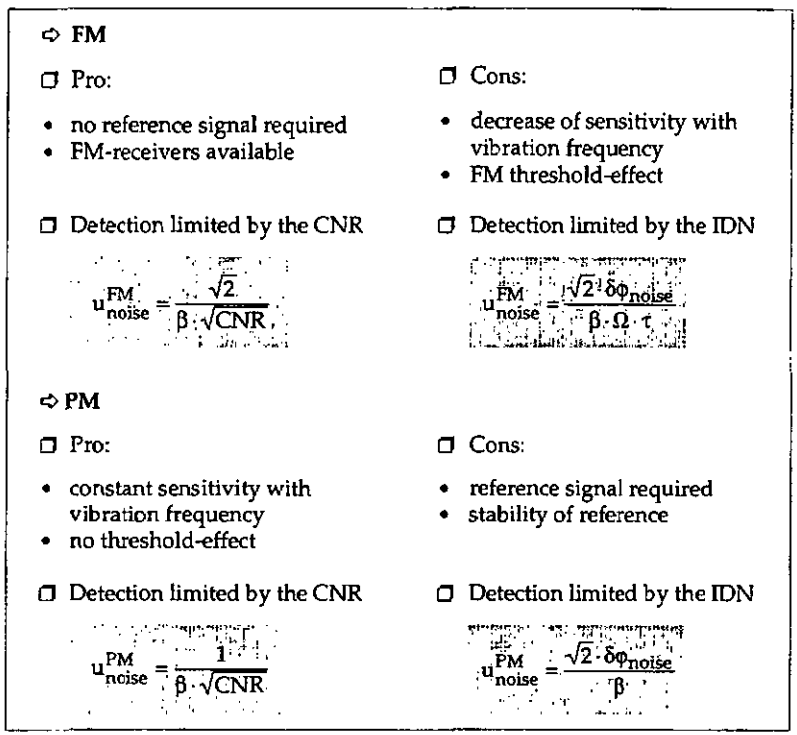


Table 3.1 Comparison between frequency- and phase-demodulation (IDN: internal demodulator noise).

3.3 Full-dynamic phase-demodulator

3.3.1 Analog full-dynamic phase-demodulator

The concept of the full-dynamic phase-demodulator is schematically presented in Fig. 3.13. The phase of the modulated signal (probe) is compared with the constant phase of the reference signal at the heterodyne frequency (455 kHz). We utilized the optical setup shown in Fig. 3.3 to generate optically these two signals. Typical interference signals are shown in Fig. 3.14. We would like to emphasize again that we have to measure the phase difference between these two signals, which is proportional to the position of the vibrating object.

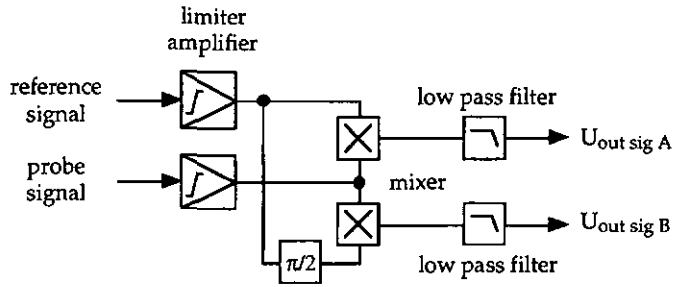


Fig. 3.13 Full-dynamic phase-demodulator: analog version.

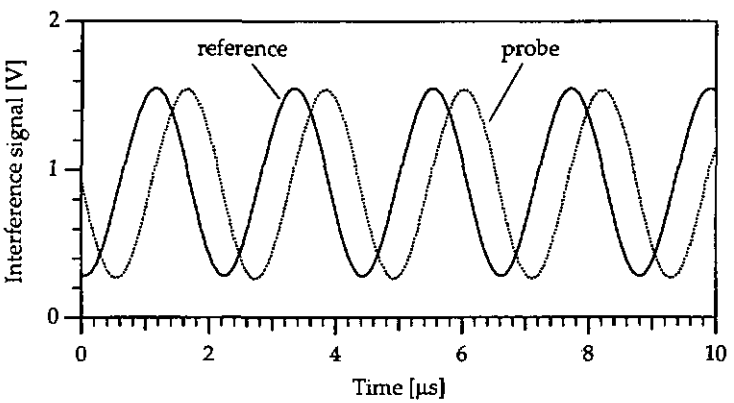


Fig. 3.14 Measured interference signals at the heterodyne frequency (455 kHz). The phase difference between these two signals is proportional to the position of the vibrating object.

The analog version of the phase-demodulator consists essentially of a limiter amplifier, a mixer and an active low-pass filter (cut-off frequency = 6.9 kHz) to eliminate the carrier frequency after mixing. The amplitude of the output signal is thus proportional to the phase difference between the modulated signal and the reference signal. Contrary to a classical phase-demodulator, which has a limited dynamic range of ± 90°, this phase-demodulator has a unlimited dynamic range. Indeed, from two signals in quadrature ($U_{out\ sig\ A}$ and $U_{out\ sig\ B}$), it is possible to recover the phase of the interference signal using a simple algorithm (see § 3.3.3). To obtain these two signals in quadrature, the reference signal undergoes first a 90° phase-shift and is then mixed with the modulated signal. The circuit diagrams of the full-dynamic phase-demodulator (FDPD) are given in Annex A. Here we will go through this FDPD step by step to understand how it works.

3.3.1.1 Input stage

The reference and probe channels are identical, as shown in Fig. 3.16. The input stage is composed of three non-inverting amplifier stages built from wide bandwidth operational amplifiers (Burr Brown OPA 654). These amplifiers are ac-coupled and Germanium diodes (General Instruments AA143) are used to clip the signal to ± 0.35 V. Finally, the signals are applied to high speed differential comparators (National Semiconductor LM360) with complementary TTL outputs.

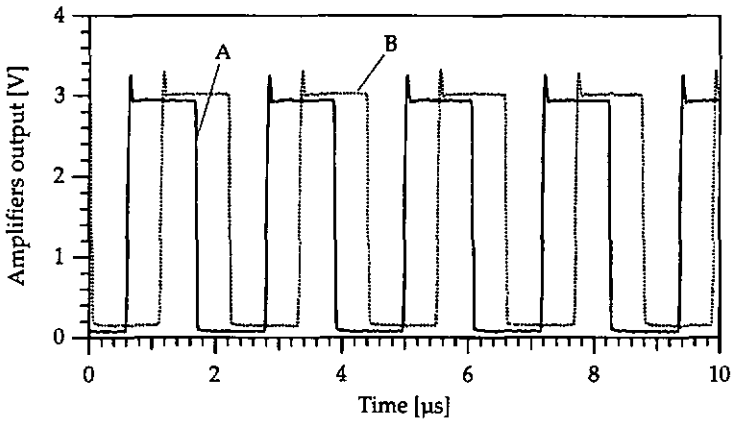


Fig. 3.15 Output signals of the input stage. A: reference signal and B: probe signal.

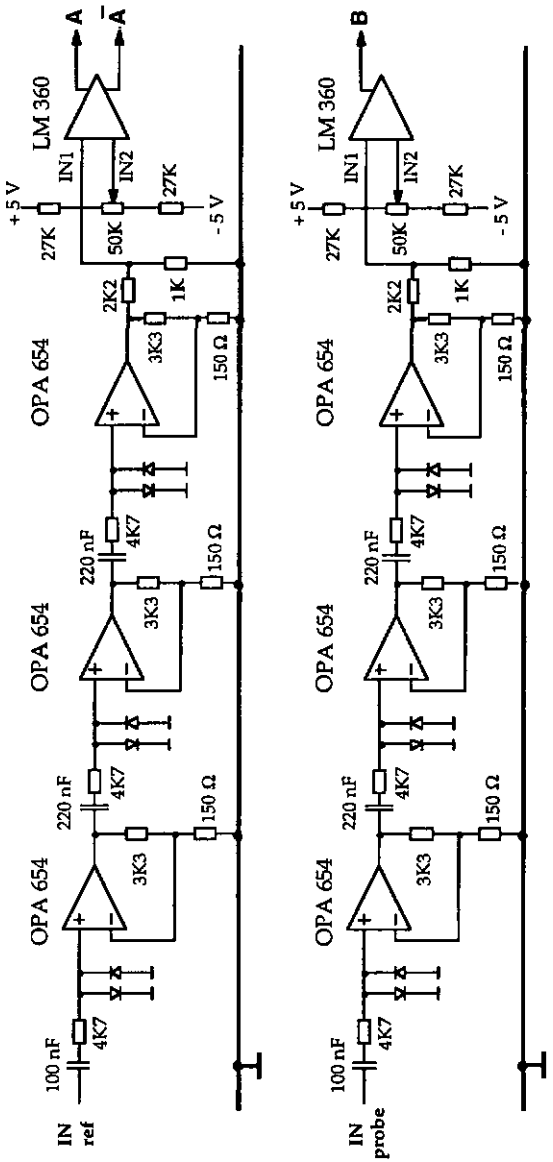


Fig. 3.16 Input stage of the full-dynamic phase-demodulator.

Thus, two TTL signals (0–5 V) are obtained at the output of comparators, as represented in Fig. 3.15. One acts as the reference signal (A) and the second acts as the probe signal (B). Special care was taken during the design of the input stage to reduce at maximum the delay difference between the two channels. Indeed, this delay difference will be interpreted as a phase difference between the reference and probe signals, resulting, at the end, in a measurement error. The input sensitivity is about 5 mV and the dynamic range is more than 80 dB.

3.3.1.2 Phase-shifter module and mixer

As already mentioned, we need two signals in quadrature (90° phase-shift) to recover the phase of the interference signal. To obtain these two signals, the reference and the probe signals are divided by two, using JK flip-flop (74HCT73) as shown in Fig. 3.17. The complementary signal of the reference signal ($\overline{A}$ in Fig. 3.17) is also divided by two.

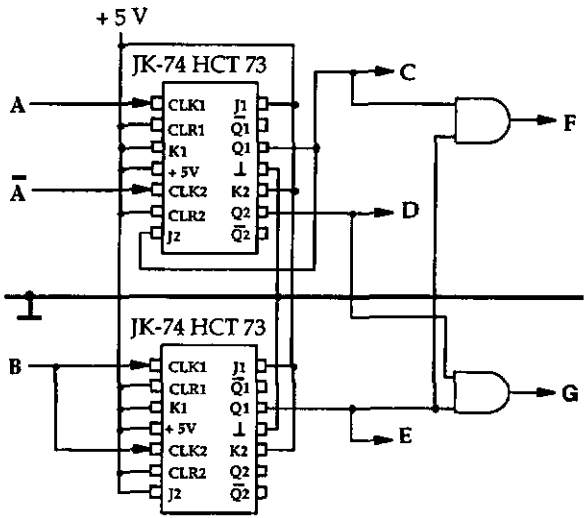


Fig. 3.17 Phase-shifter module (90°) and mixers.

Thus, three signals at half the heterodyne frequency (227.5 kHz) are obtained. The first one is the reference signal (C), the second one is the probe signal (E) and the third one is the reference signal shifted by 90° (D), as plotted in Fig. 3.18. Indeed, the 180° phase-shift between A and $\overline{A}$ (at the heterodyne frequency) corresponds now and after the division by two, to a 90° phase-shift between C and D (at half

the heterodyne frequency). At present, we would like to draw you attention to the fact that an input phase difference ϕ between A and B will now correspond to a phase difference of $\phi/2$ between C and E. Then, these three signals C, D and E are applied to AND-gates (74HCT04) according to Fig. 3.17. As shown in Fig. 3.19, the AND-gates operate as mixers and produce two pulse trains at 227.5 kHz. Thus, the phase difference between the reference and the probe signal (modulating signal) is converted to the duration of the individual pulses. This form of modulation is called pulse-width modulation or pulse-length modulation [3.13]. The modulating signal may vary the time of occurrence of the leading edge (as for the signal F), the trailing edge (as for the signal G) or both edges of the pulse.

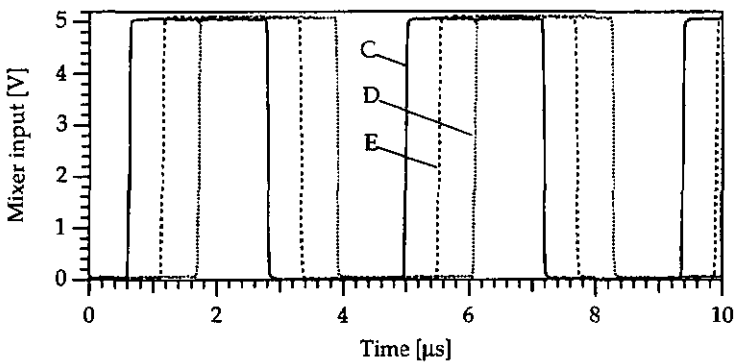


Fig. 3.18 Input signals applied to the mixers (AND-gates). C: reference signal, D: reference signal shifted by 90° and E: probe signal.

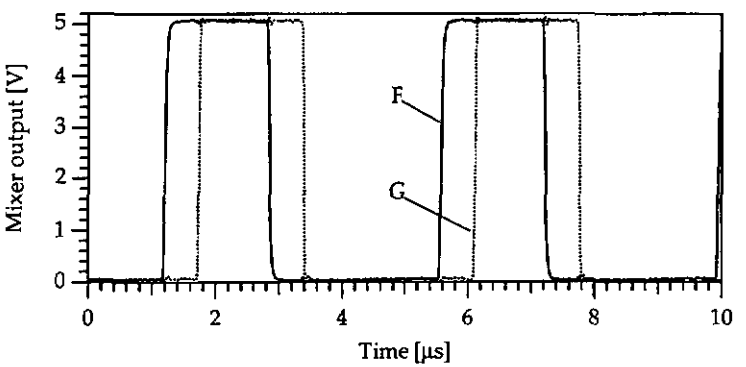


Fig. 3.19 Output signals of the mixers. These pulse trains correspond to a phase to pulse-width conversion.

The spectral analysis of pulse-width modulation (PWM) is very complicated and a complete description can be found elsewhere [3.14]. We will present only a qualitative description of the spectrum of PWM signals.

If T_s is the period of the PWM signal (T_s = 4.4 μs for f_s = 227.5 kHz) and assuming a pure sinusoidal modulating wave of small amplitude [βu < 1 in Eq. (3.2)] at frequency f_m, we find that the spectrum of this PWM signal consists of the following components, as sketched in Fig 3.20: (A) a dc component equal to the average value of the pulses. (B) Sinusoidal components of frequency equal to n/T_s, where n= 1, 2, 3.... (C) A sinusoidal component of frequency equal to f_m and in phase with the modulated signal. (D) Sinusoidal components of frequencies $(n/T_s) \pm p \cdot f_m$, where n, p = I, 2, 3....The frequency components (A) and (B) are supplied by the unmodulated pulse train, which may be regarded as the carrier of the PWM signal.

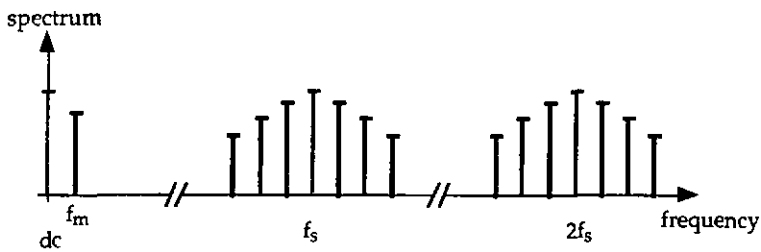


Fig. 3.20 Spectrum of pulse-width modulation signal.

One important observation is, that when natural sampling is used to generate a PWM wave, there are no harmonics of the modulating signal frequency at $2f_m$, $3f_m$, etc. [3.15]. Thus, the modulating wave can be recovered by passing the PWM signals F and G through low-pass filters. It would be believed, that the recovery of PWM signals is distortionless because no harmonics of the modulating signal frequency are generated. However, this is not true, because some cross-modulation products can fall within the filter bandwidth. The dominant ones are the harmonics of the first lower sideband, at frequencies $(1/T_s) - p \cdot f_m$, where p is sufficiently large. Their relative amplitudes, as a fraction of the signal amplitude at f_m, are given by [3.15]

$$d_p(\Delta\phi) = \frac{J_p(\Delta\phi/2)}{\Delta\phi/2}, \tag{3.31}$$

where J_p is the Bessel function of the first kind of order p and $\Delta\phi$ is the excursion of the phase difference between the reference and the probe channels. In Fig. 3.21, the relative amplitude of the 1st, 3rd and 5th harmonic is plotted as a function of $\Delta\phi$.

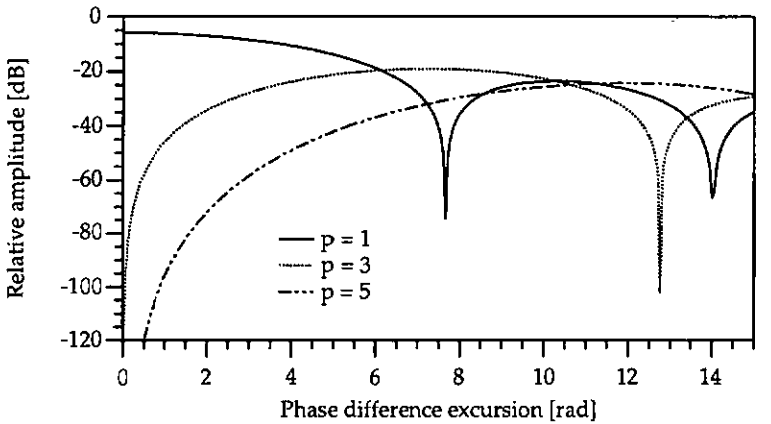


Fig. 3.21 Relative amplitude of the 1st, 3rd and 5th harmonic of the first lower sideband.

Taking the derivative of Eq. (3.31), the maximum relative amplitudes $d_p|_{max}$ are given by the solutions of the equation

$$(p+1)\,J_p\left(\frac{\Delta\phi_{max}}{2}\right) - \Delta\phi_{max}\,J_{p-1}\left(\frac{\Delta\phi_{max}}{2}\right) = 0. \qquad (3.32)$$

The numerical solutions $\Delta\phi_{max}$ of this equation and the maximum relative amplitude $d_p|_{max}$ are represented, for the 50 first harmonics, in Fig. 3.22.

Therefore, as far as the carrier frequency ($1/T_s$) is much higher than the cut-off frequency of the low-pass filter, the amplitudes of the harmonics of the first lower sideband are very small. This allows to reconstruct the modulating wave with a very small level of phase distortion.

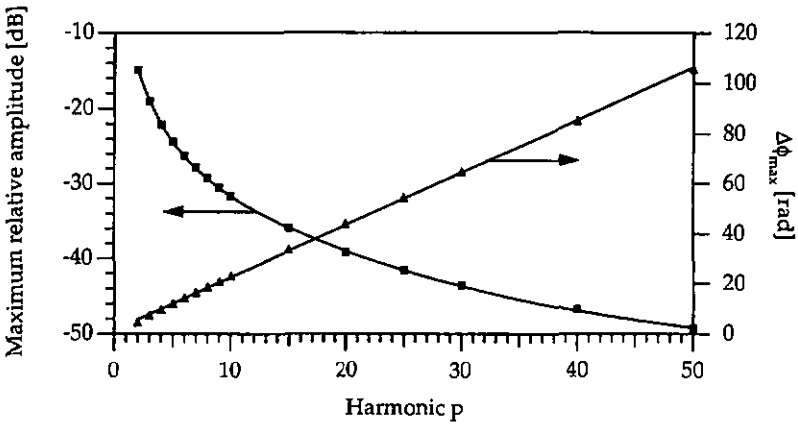


Fig. 3.22 Numerical solutions of Eq. (3.36) for the 50 first harmonics.

3.3.1.3 Active low-pass filter

We have designed an 2nd-order active low-pass filter to recover the modulating wave or, in others words, to remove the carrier frequency. The circuit diagram of this filter is shown in Fig. 3.23.

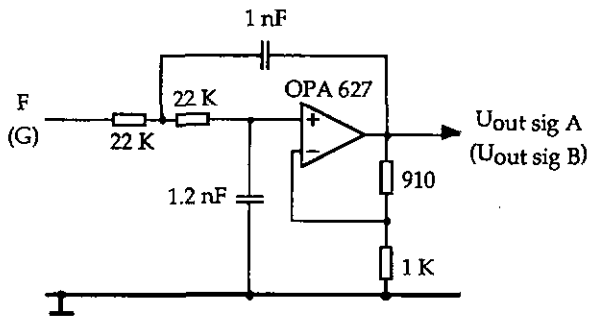


Fig. 3.23 Active low-pass filter (cut-off frequency = 6.9 kHz), used to eliminate the carrier frequency after mixing.

The frequency response of this 2nd-order system is given by [3.16]

$$T(\omega) = \frac{G_0}{1 + j2\xi\frac{\omega}{\omega_0} - \left(\frac{\omega}{\omega_0}\right)^2}, \qquad (3.33)$$

where the undamped resonance frequency ω_0 and the low frequency gain G_0 are

$$\omega_0 = \frac{1}{\sqrt{R_1 C_1 R_2 C_2}} \tag{3.34}$$

and

$$G_0 = 1 + \frac{R_4}{R_3}. \tag{3.35}$$

We get for the damping factor

$$\xi = \frac{1}{2}\left[\sqrt{\frac{R_1 C_2}{R_1 C_1}} + \sqrt{\frac{R_1 C_2}{R_2 C_1}} + (1 - G_0)\sqrt{\frac{R_1 C_1}{R_2 C_2}}\right]. \tag{3.36}$$

The bandwidth of this filter is given by Eq. (3.13). Substituting the values for the circuit shown in Fig. 3.23 into Eqs. (3.34) to (3.36), we obtain a dc-gain of $G_0 \cong 1.91$, a damping factor of $\xi \cong 0.68$, and a cut-off frequency of $f_c = B \cong 6.9$ kHz.

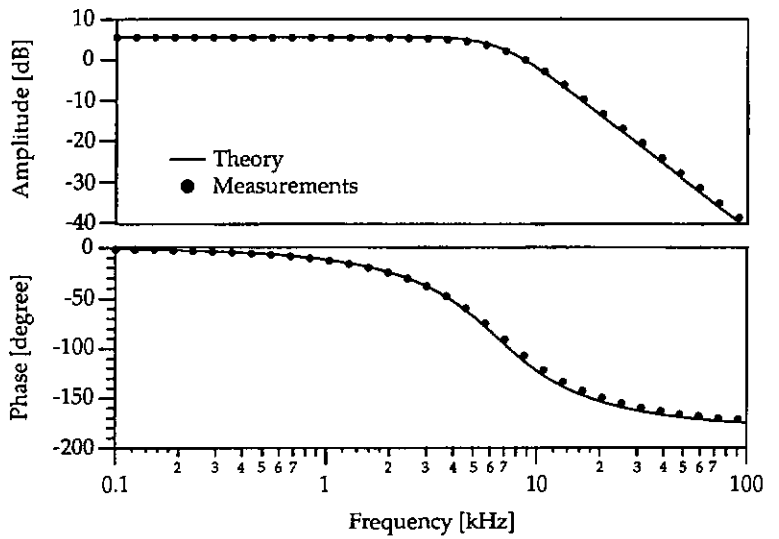


Fig. 3.24 Frequency response of the active low-pass filter.

The amplitude and phase of the frequency response of low-pass filters shown in Fig. 3.24 was measured with a network analyzer. There is a good agreement between the theoretical prediction (plain line) and measured values (dots).

The amplitude of the output signal, after the low-pass filter, is thus proportional to the phase difference between the modulated signal and the reference signal. From the two signals in quadrature shown in Fig. 3.25, it is possible to recover the phase of the interference signal using a simple algorithm, which is described in paragraph 3.3.3. The frequency response of this full-range phase-demodulator is limited by the cut-off frequency of the low-pass filter (f_c = 6.9 kHz), which is required to eliminate the carrier frequency after the mixer. Using a spectrum analyzer, we measured an attenuation of approximately 50 dB at the carrier frequency (227.5 kHz), which compares well with the theoretical value of 56 dB obtained from Eq. (3.33). The analog outputs $U_{out\ sig\ A}$ and $U_{out\ sig\ B}$ of the demodulator are sampled with 12 bits resolution at a maximum sampling rate of 50 kHz (NB-MIO-16L card and LabView software from National Instruments) and then processed off-line to extract the phase information, which is the displacement of the surface under test.

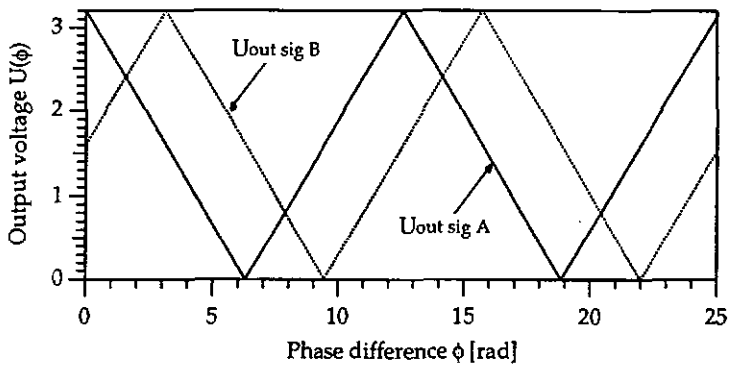


Fig. 3.25 Output signals of the FDPD as a function of the phase difference between the reference and the probe signals.

To conclude this paragraph devoted to the analog version of the full-dynamic phase-demodulator, we would like to mention one important drawback inherent to its principle of operation. The low frequency intensity fluctuations of the laser are transmitted through the phase demodulator and will produce phase noise on the FDPD output signals. When the laser intensity is not constant, the output voltage of the detector is given by [see Eq. (3.1)]

$$V(t) = V_0(t)\{1 + m\cos[\Delta\omega t + \phi(t)]\}, \tag{3.37}$$

where $V_0(t)$ is proportional to the laser intensity fluctuations and $\Delta\omega/2\pi$ is the heterodyne frequency (455 kHz). The phase-demodulator inputs being ac-coupled to the limiter amplifiers (see Fig. 3.16), the fluctuations of the laser intensity influence the zero-crossings of signals A and B. Therefore, the phase of signals C, D and E (see Fig. 3.17) are modified accordingly. If the input signals have similar fluctuations, the phase of the signals C and E are also modified similarly. Consequently, the laser intensity fluctuations do not affect the duty-cycle of $U_{out\ sig\ A}$. On the contrary, because the signal D is derived from the inverted signal $\overline{A}$, the phase of the signal D changes in the opposite sense with respect to the phase of E. Thus, the duty-cycle of signal G = D·E is influenced by these laser intensity fluctuations, resulting in phase noise on the output signal $U_{out\ sig\ B}$. This different behavior of the two FDPD outputs in presence of intensity laser noise is illustrated in Fig. 3.26, which shows the power spectrum of $U_{out\ sig\ A}$ and $U_{out\ sig\ B}$ in presence of laser intensity noise. This measurement was carried out with the same optical signal connected two both demodulator inputs.

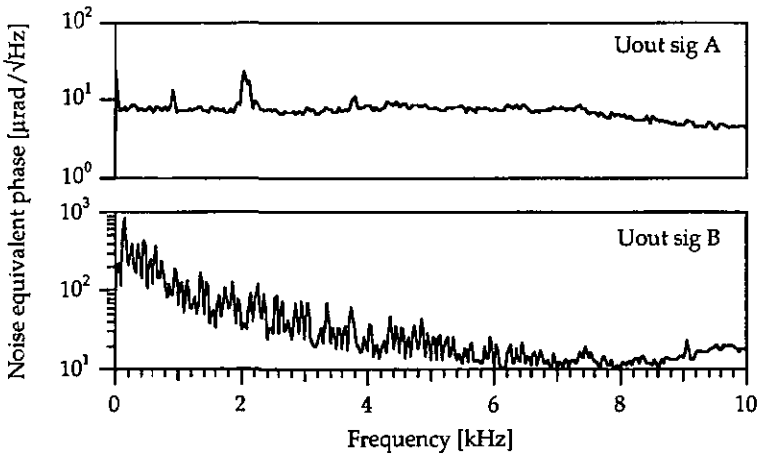


Fig. 3.26 Power spectrum of the phase demodulator output signals in presence of laser intensity noise (measured with 30 Hz bandwidth, converted to 3 kHz) for identical reference and input signals.

Even if the reference signal has a constant dc-value and is noise free, the laser intensity fluctuations influence the duty-cycle of signals C and D and will be transmitted through the FDPD, resulting in phase noise on both demodulator output signals.

One solution to overcome this drawback would be to use a band-pass filter at the detector output. Of course, the bandwidth of this filter must be sufficiently wide to take into account of the bandwidth associated to PM-signal (see § 3.5.2) and the phase-slope of the filter must be linear in the band of interest to avoid phase distortion (see § 3.1.2). Using a band-pass filter also reduces the noise level at the phase-demodulator inputs, resulting in a reduction of the input sensitivity.

3.3.2 Digital full-dynamic phase-demodulator

To overcome the most limiting factor of the analog FDPD, that is the limited frequency response of the system, an improved version of the full-dynamic phase-demodulator with a high throughput speed has been developed (see Fig 3.27).

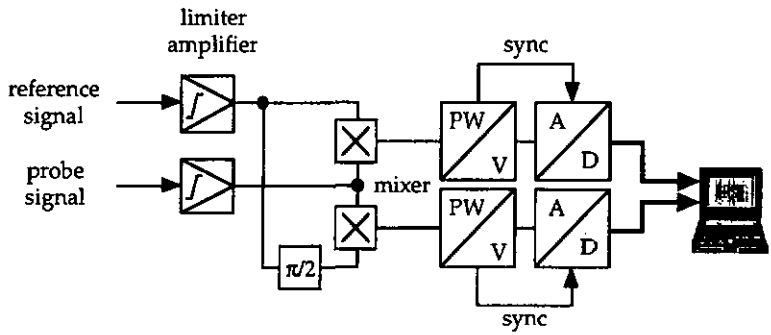


Fig. 3.27 Digital version of the full-dynamic phase-demodulator.
PW/V: pulse-width to voltage converter.

The idea is to replace the low-pass filter in Fig. 3.13 by a pulse-width to voltage converter (PW/V) followed by a synchronized A/D converter working at the carrier frequency (455 kHz/2 = 227.5 kHz). For this digital version with an extended frequency response, a new value for the phase difference between the modulated and the reference signals is obtained every 4.4 μsec. For the acquisition of the two 12-bit output signals, we used a high-speed 32-bit parallel digital I/O interface board (National Instruments card NB-DIO-32F) synchronized to the A/D converters. We would recall, that the mixer output signals F and G are pulse-width modulation waves, where the phase difference between the reference and probe signals defines the duration of individual pulses (see Fig. 3.19). The phase information is thus contained in the width of the individual pulses. The principle to convert the pulse width into a constant voltage is relatively clear; it is

shown in Fig. 3.28 for the channel A. A capacitance of 1 nF is charged with a constant current as long as the pulse is high; then, this voltage is hold and synchronously sampled by the A/D converter (arrows in Fig. 3.28). This mechanism is realized practically by the circuit sketched in Fig. 3.29 and described in more detail in Annex A.

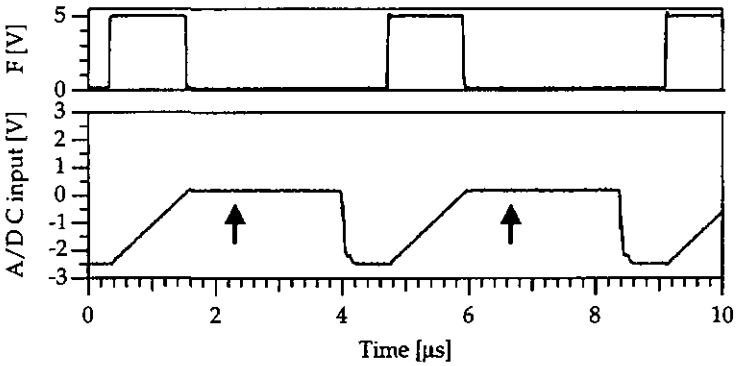


Fig. 3.28 The duration of the individual pulses is converted into a constant voltage, which is then synchronously sampled.

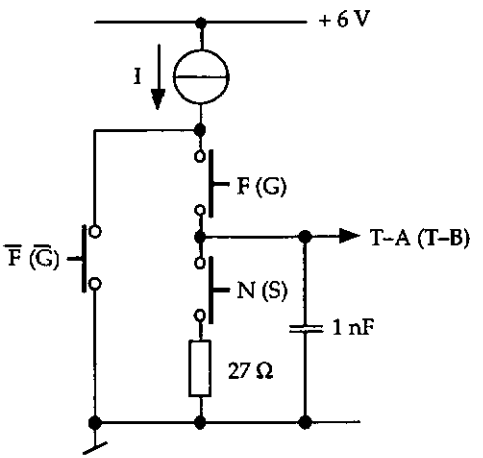


Fig. 3.29 Circuit which converts the pulse duration into a voltage. The variables in brackets are for the second channel B.

The signals F and $\overline{F}$ for channel A, respectively G and $\overline{G}$ for channel B, are used either to charge the capacitance or to hold the voltage. The signal N (S for channel B) is employed to discharge the capacitance once the signal T-A (T-B for channel B) was sampled by the A/D converter and the conversion is completed. The signals N and S are obtained from the signals D and C, respectively (see Annex A).

An amplifier stage is used to match the signals T-A and T-B to the input characteristics of the A/D converter (± 2.5 V). Then, these signal are digitized by a 12-bit analog-to-digital converter (Burr Brown ADS7819). This chip is a complete 12-bit sampling A/D converter using CMOS structure. It contains a complete switched-capacitor based A/D with inherent sample and hold capability, a reference voltage source, and an interface with three-state output driver for microprocessor use. The ADS7819 is specified at an 800 kHz sampling rate. This converter outputs full parallel data in Binary Two's Complement data format.

Figures 3.30 and 3.31 show the A/D conversion timing diagrams for the channels A and B, respectively. When C, respectively D for the second channel, goes down, a conversion is initiated by the falling edge of R/$\overline{C}$. Then, $\overline{busy}$ will go low and stay low until the conversion is completed and the output registers are updated. $\overline{busy}$ going high is used to transfer the data into the computer. For the acquisition of the two 12-bit digital output signals, we used a high-speed 32-bit parallel digital I/O interface board (National Instruments card NB-DIO-32F). Thus, for this digital version, a new value for the phase difference between the modulated and the reference signals is obtained every 4.4 µsec, that is at half the heterodyne frequency. Of course, it is not always necessary to collect data at such a high speed. We can also reduce the sampling rate by using the signals $\overline{P}/2$ or $\overline{P}/4, \ldots,$ or $\overline{P}/16$ ($\overline{N}/2, \overline{N}/4, \ldots, \overline{N}/16$ for channel B) instead of $\overline{P}$ ($\overline{N}$ for channel B) to initiate an A/D conversion. This gives us the possibility to choose, according to the experiment and the frequency range to be studied, a sampling frequency of either 227.5 kHz, 113.75 kHz, 56.88 kHz, 28.44 kHz, or 14.22 kHz.

The output characteristics of the digital version of the FDPD are identical to those of the analog version, shown in Fig. 3.25. From these two signal in quadrature, the phase information on the vibrating object can be recovered using a simple algorithm, described in paragraph 3.3.3.

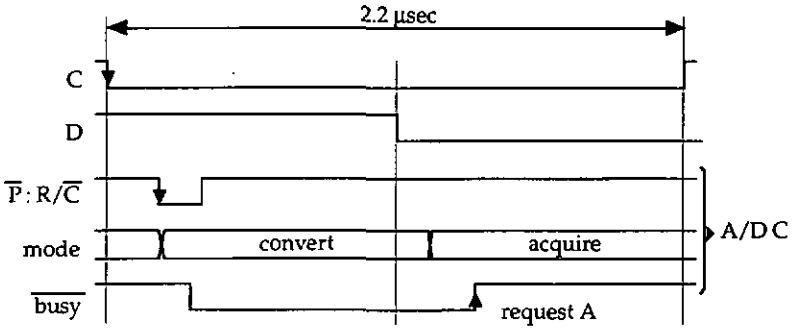


Fig. 3.30 Conversion timing for channel A.

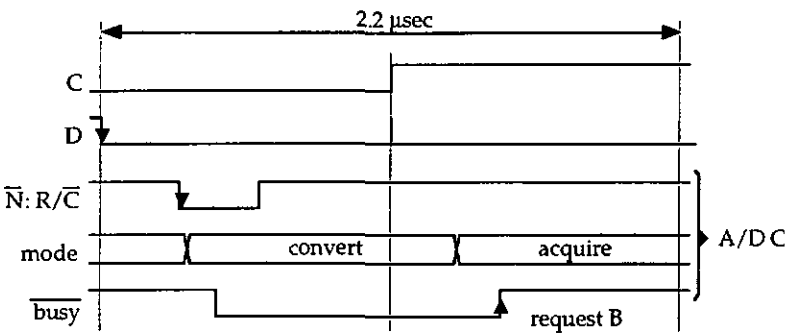


Fig. 3.31 Conversion timing for channel B.

3.3.3 Phase recovery

3.3.3.1 Signal processing for phase reconstruction

As already described in the previous paragraphs, the two output signals in quadrature coming either from the analog or digital version of the FDFD are sampled using a Macintosh computer and National Instruments acquisition boards. The phase information, that is the position of the vibrating object, is then recovered off-line. For that purpose, we wrote the programs given in Annex B, using the software Igor from WaveMetrics. This scientific package allows us not only to reconstruct the movements of the object under study, but also to make classical analyses by exploiting all the available mathematical tools (FFT, statistics functions, etc.).

Figures 3.32 and 3.33 show the FDPD output characteristics for the channels A and B, respectively. The name and the physical meaning of the different variables employed by the signal processing program are also outlined in these figures. We would like to emphasize that in these figures, the phase corresponds to the optical phase difference between the reference and the probe channels. That is why these signals are 4π-periodic. Indeed, because the input optical signals at the heterodyne frequency are first divided by two, an input phase difference of ϕ is equivalent to a phase difference of $\phi/2$ at half the heterodyne frequency.

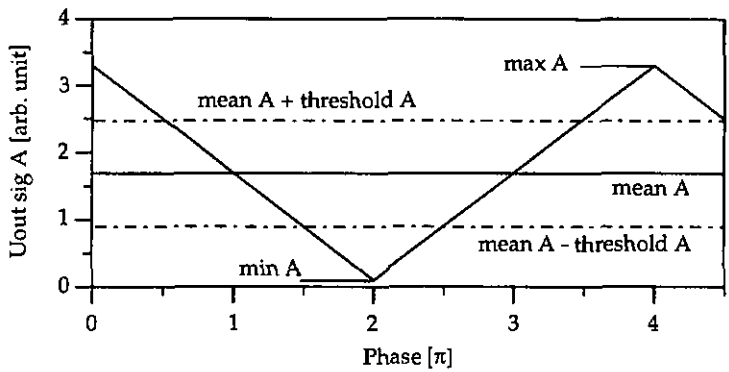


Fig. 3.32 FDPD output signal (channel A) with the variables used in the signal processing program.

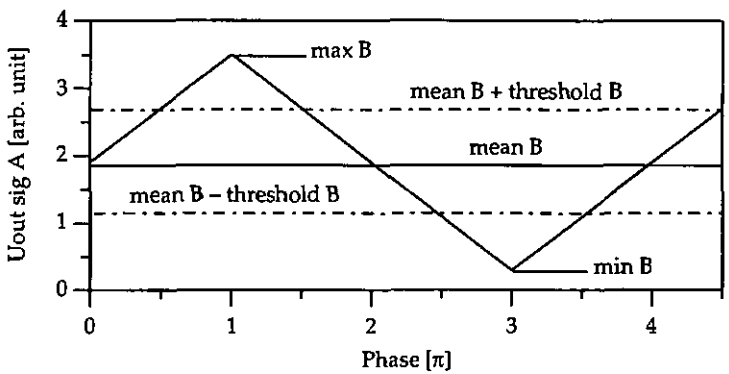


Fig. 3.33 FDPD output signal (channel B) with the variables used in the signal processing program.

One very important point is, that the phase demodulator must be calibrated. In other words, before recovering the phase information, we must introduce the maximum (maxA and maxB) and minimum (minA and minB) values reached by the output signals. Then, we can calculated the other variables needed (with the function Data_Init(maxA, minA, maxB, minB), see Annex B). The calibration procedure and the influence of wrong calibration on the measurement accuracy will be presented in paragraph 3.3.3.2.

The concept behind the proposed signal processing to recover the phase information (with the function Phase_Recovery(WaveNameA, WaveNameB), see Annex B) is simple, efficient, fast and accurate. It can be separated in three parts, as summarized in Fig. 3.34. First, using both channels, the program determines in which quadrant is the phase of the 4π-periodic output signals (see Figs. 3.32 and 3.33). This is carried out by looking at the sampled signals A and B to see if they are above or below the mean values meanA and meanB, respectively.

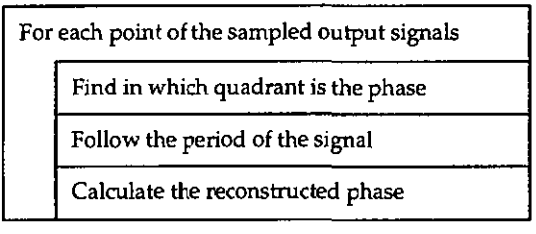


Fig. 3.34 Principle of operation of the signal processing for phase recovery.

Second, the number of periods followed by the reconstructed phase is determined by tracking the transitions from the fourth to the first quadrant and from the first to the fourth quadrant. This is equivalent to the well-known method of incremental fringe counting in interferometry using two signals in quadrature [3.17]. Of course, it is assumed that the number of periods is zero when the system is at rest (initial condition).

Finally, the absolute value of the reconstructed phase is calculated using at once the information on the number of periods and the quadrant. For this final step, either the channel A or the channel B is employed to compute the reconstructed phase (see Annex B). Indeed, the program use the channel A if the output signal

$U_{out\ sig\ A}$ is found in the interval [mean – threshold; mean + threshold] (see Figs 3.32 and 3.33), otherwise the channel B is utilized. Thus, only the linear part of the output characteristics are used by the signal processing program. We define the threshold value as a fourth of the peak-to-peak signal amplitude. Obviously, at least one output voltage must be in this interval, but this is always the case because these two signal are in quadrature.

3.3.3.2 FDPD calibration

The aim of the calibration procedure is to measure the minimum and maximum values of the output signals of the phase-demodulator. These values will be employed by the signal processing for the phase recovery. One way to obtain this information, and probably the easiest, is to vary the optical phase difference of the interferometer on several π, for example by driving the PZT with a voltage ramp. From Fig. 3.35, where the two output signals in quadrature $U_{out\ sig\ A}$ and $U_{out\ sig\ B}$ are plotted versus the time, it is elementary to determine the max. and min. of both signals.

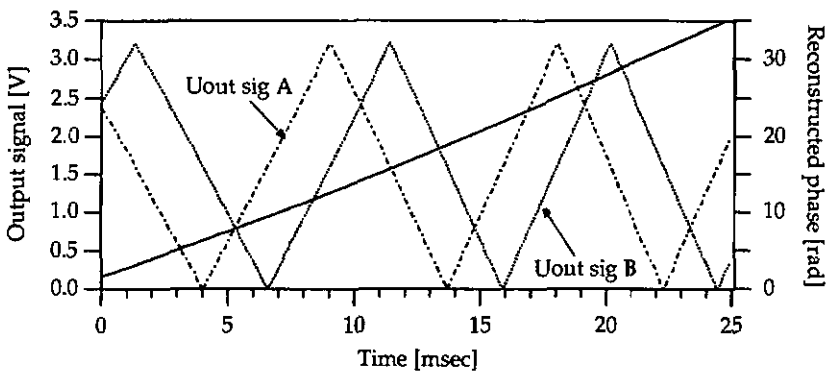


Fig. 3.35 Measured output signals used for the calibration procedure of the analog FDPD. The reconstructed phase obtained from the two signal in quadrature is also shown (plain line).

It is also apparent in Fig. 3.35 that the signal processing presented in the paragraph 3.3.3.1 works absolutely well. Indeed, the reconstructed phase presents no discontinuity and is perfectly smooth.

The phase reconstruction works accurately, as long as the calibration parameters maxA, minA, maxB and minB remain correct. If these parameters are wrong , the reconstructed phase shows some phase discontinuities (jumps), which occur when the program switches from one channel to the other.

Practically, the minimum and maximum values reach by the output signals of the phase-demodulator are given by the low-pass filter components for the analog FDPD (see Fig. 3.23), and by the pulse-width to voltage conversion for the digital FDPD (see Fig. 3.29). More precisely, the calibration values depend on the electronic components (resistances, capacitances, ...) utilized to build the demodulator. Thus, if electronic components of good quality are used, the calibration procedure described in the previous paragraph has to be carried out only once, because the calibration parameters will remain constant as a function of time.

The reconstructed phase for a sinusoidal excitation of 4 rad obtained from computer simulations is shown in Fig. 3.36, first when the calibration parameters are correct (upper trace) and second when they are wrong (lower trace). Thus, a bad calibration will introduce some amount of distortion, leading to an error on the vibration amplitude at the fundamental frequency.

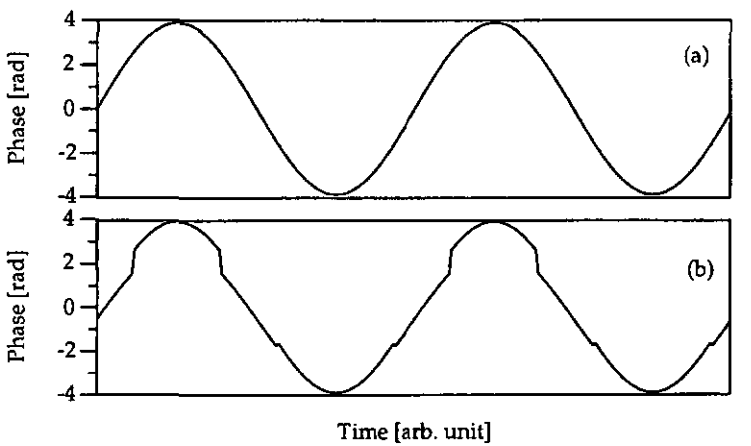


Fig. 3.36 Simulation of the phase reconstruction when the calibration parameters are (a) correct and (b) inexact.

To characterize the influence of calibration errors on the reconstructed phase, we used the well-known notion of harmonic distortion d given by [3.18]

$$d = \frac{\sqrt{A_2^2 + A_3^2 + A_4^2 + \ldots}}{A_1}, \tag{3.38}$$

where A_1 and A_n (n = 2, 3, ...) are the rms amplitude of the fundamental and of harmonics, respectively. We carried out computer simulations for a sinusoidal vibration (phase amplitude = 4 rad, frequency = 5 Hz) and with two classes of calibration errors. First, when only one calibration parameter is incorrect (maxA) and second, when two calibration parameters are wrong (maxA and maxB). We employed FFT analyses to calculate the harmonic distortion of the output signal.

The result of these simulations are shown in Fig. 3.37, where the harmonic distortion d is plotted as a function of the calibration error ε defined as

$$\varepsilon = \frac{\max A(\text{or } B)\big|_{\text{inexact}} - \max A(\text{or } B)\big|_{\text{exact}}}{\max A(\text{or } B)\big|_{\text{exact}}}. \tag{3.39}$$

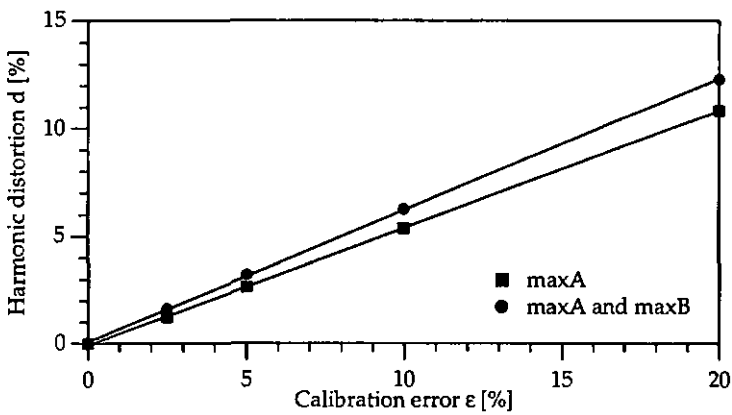


Fig. 3.38 Harmonic distortion of the output signal when one (squares) and two (circles) calibration parameters are inexact.

Practically, a calibration error of less than 1 % can be attained and so a distortion of the output signal below 0.5 % can be expected. We would like to emphasize

that this distortion is due to the signal processing principle itself and is the result of a wrong calibration of the FDPD.

3.3.3.3 Examples of phase reconstruction

This paragraph illustrates how our full-dynamic phase-demodulator is able to recover interference phases in different practical situations, for example when the vibration amplitude is small, or when the displacement is very slow but with a considerable amplitude, or even when the surface under test undergoes fast and abrupt movement greater than the optical wavelength.

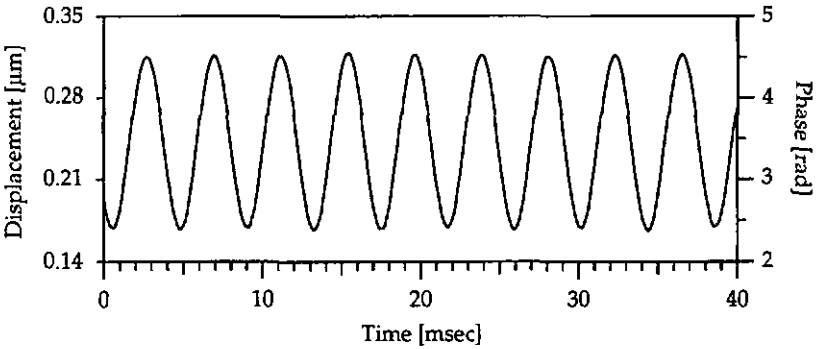


Fig. 3.38 Phase reconstruction with the analog FDPD for a sinusoidal motion of approximately 150 nm at a frequency of 230 Hz.

For the first example, the PZT was driven with a sinusoidal voltage at approximately 230 Hz. Figure 3.38 shows the reconstructed interference phase and displacement measured with the analog version of the phase-demodulator. We measured a peak-to-peak amplitude of about 150 nm. As already mentioned, the amplitude sensitivity of the phase-demodulation principle is independent of the vibration frequency and can thus be employed even for quasi-static vibration measurements. This is illustrated in Fig. 3.39 when the PZT was driven by a voltage ramp with a temporal period of approximately 2.5 sec on which a small oscillation of 20 Hz is superposed. The output signals of the analog FDPD are shown in the upper part of Fig. 3.39. The reconstructed interference phase and the displacement of the mirror are shown in the bottom part of Fig. 3.39. Here, the total displacement of the surface is nearly 2.5 µm with a vibration amplitude of 150 nm peak-to-peak.

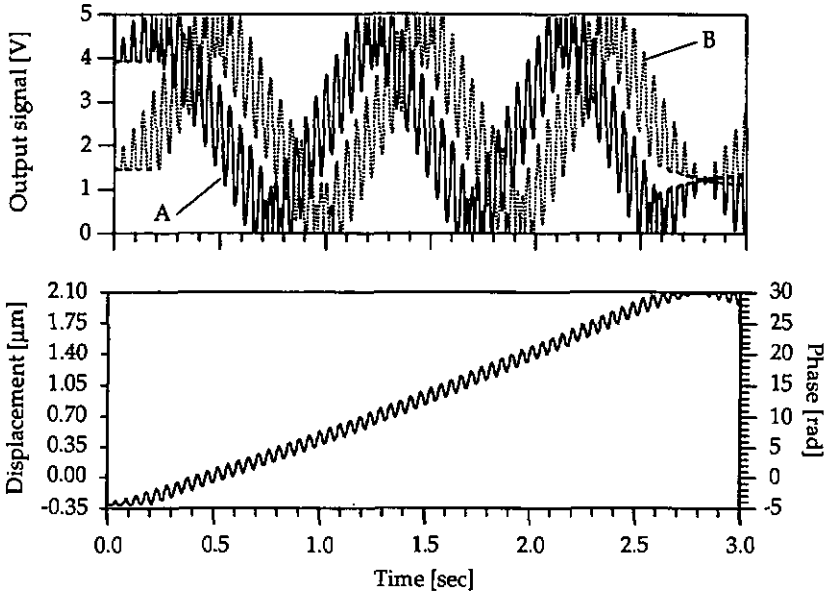


Fig. 3.39 Example of phase reconstruction with the analog FDPD when the displacement of the vibrating object is greater than the wavelength of the light.

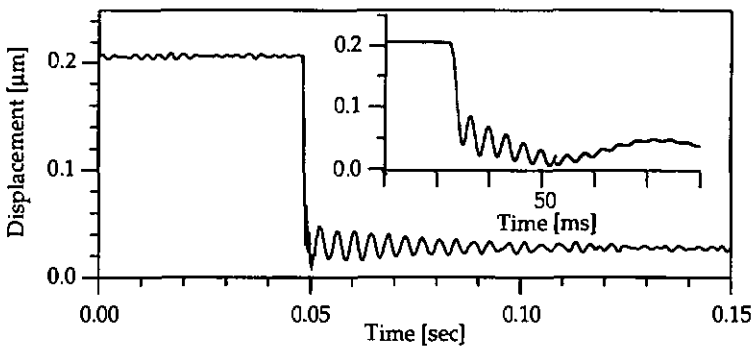


Fig. 3.40 Example of phase reconstruction using the digital FDPD.

The last example, shown in Fig. 3.40, is the response of the PZT to an abrupt step. Here we used the digital version of the FDPD. It should be indicated that the phase change is smaller than π between two samples (4.4 μsec), which allows the

system to follow accurately the phase variations. The damped mechanical oscillations of the PZT can be readily observed. The frequency of these oscillations is approximately 250 Hz. As shown in the insert of Fig. 3.40, additional fast oscillations of about 3 kHz occur just after the step during approximately 4 ms. Oscillations as small as a few nm can readily be measured and analyzed with our digital FDPD.

3.4 System performance

3.4.1 Noise measurements

In order to characterize both types of phase-demodulator, we measured for several different CNRs the spectrum of the noise equivalent phase, that is the minimum detectable phase excursion as a function of the vibration frequency. We used as input signals an unmodulated electrical carrier at 455 kHz connected to both modulator inputs. This is equivalent to simulate a zero phase difference between the two input signals. We employed a FFT spectrum analyzer (Stanford Research 760) to measure the analog outputs $U_{out\ sig\ A}$ and $U_{out\ sig\ B}$. The FFT analysis of the digital signals was realized off-line using the software Igor Pro 3.0 from WaveMetrics.

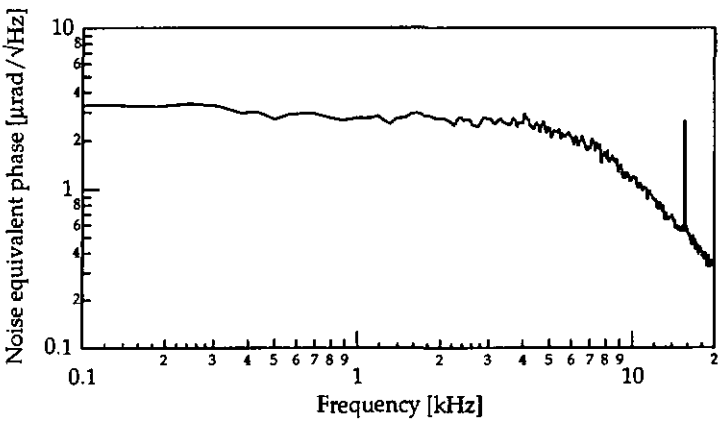


Fig. 3.41 Noise equivalent phase measured at the analog phase-demodulator output (unmodulated electrical carrier, CNR = 70 dB).

The phase noise level for a CNR = 70 dB (measured with 30 Hz bandwidth, converted to 3 kHz bandwidth) as a function of frequency is shown in Fig. 3.41 for the analog phase-demodulator. The peak at 15.6 kHz is due to electrical pick-up from the noisy environment. The shape of the curve in Fig. 3.41 is mainly given by the transfer function of the active low-pass filter (see § 3.3.1.3). Below the cut-off frequency (6.9 kHz), the noise equivalent phase is nearly constant and depends, as expected, of the CNR as summarized in Table 3.2. The right-side column of Table 3.2 is the equivalent phase noise in case of a limitation by the CNR calculated from Eq. (3.18). Therefore, we can conclude that the analog FDPD is limited, for CNR below 70 dB (3 kHz bandwidth), by the applied input noise (limitation due to the CNR). Moreover, the internal demodulator noise could be estimated to be smaller than 1 µrad/√Hz.

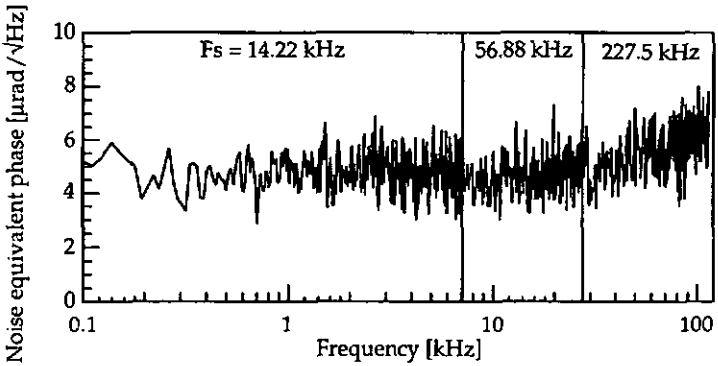


Fig. 3.42 Noise equivalent phase measured at the digital phase-demodulator output (unmodulated electrical carrier with CNR = 70 dB, F_S: sampling frequency).

	$\delta\varphi^{PM}_{noise}$ [µrad/√Hz]		
CNR [dB]	Analog outputs	Digital outputs	CNR limitation
50	47	61	40
60	12	19	13
70	3	5	4

Table 3.2 Measured noise equivalent phase (unmodulated electrical carrier) and phase noise limitation due to the CNR for various input CNR (3 kHz bandwidth).

As shown in Fig. 3.42 for the digital version, the spectrum of the noise equivalent phase is approximately flat for frequencies below 30 kHz, then the noise increases slightly up to the Nyquist frequency of 113.75 kHz. The spectrum of the noise equivalent phase depends on the input CNR as reported in Table 3.2. We would like to mention that we acquired the digital signals at reduced sampling frequencies F_S (see § 3.3.2) to study the low-frequency part of the noise spectrum.

Quantizing noise is introduced in the digital phase-demodulation by the A/D converter. If we assume a quantizing process with an uniform step ΔV, the root-mean-square value of the quantizing noise δV_q is given by [3.19]

$$\delta V_q = \frac{\Delta V}{\sqrt{12}} = \frac{V_{max}}{2^n \sqrt{12}}, \tag{3.40}$$

where V_{max} is the full scale range and n the number of bits of the A/D converter. For V_{max} = 5 V and n = 12 bits, we obtain ΔV = 1.22 mV, which leads to a quantizing noise of $\delta V_q \cong 350$ µV or $\delta\phi_q \cong 500$ µrad. It should be noted that uniform quantizing yields a noise which is virtually independent of the signal and, as a first approximation, that it is reasonable to take the total quantizing noise power, given by $\delta V_q{}^2$, as a measure of the noise to be expected in the frequency band of interest, after demodulation [3.19].

On the other hand, the noise equivalent phase $\delta\varphi_{noise}^{PM}$ (see Table 3.2) was calculated from 8192 samples. From statistical considerations it can be estimated that the quantizing noise equivalent phase is $\delta\phi_q/\sqrt{8192} \cong 5.5$ µrad. This value is in good agreement with the one reported in Table 3.1 for CNR = 70 dB. Therefore, for high CNR, the digital FDPD system seems to be limited by the quantizing noise introduced by the 12-bit A/D converter.

Finally, we measured the phase noise of the FDPD when the same optical signal is connected to both demodulator inputs. The power spectrum of this signal is shown in Fig. 3.43. At 470 kHz, we have a CNR of approximately 68 dB (measured within 30 Hz bandwidth, converted to 3 kHz). The noise equivalent phase spectrum of the analog FDPD is shown in Fig. 3.44. From this figure, we obtained a noise equivalent phase of about $\delta\varphi_{noise}^{PM} \cong 8$ µrad/√Hz from dc to 7 kHz. The additional noise on channel B is due to the laser intensity noise which is transmitted through the FDPD and which produced some phase noise on the output signal $U_{out\ sig\ B}$, but not on the signal $U_{out\ sig\ A}$ (see § 3.3.1). For the digital

version of the FDPD, we obtained a noise equivalent phase of approximately $\delta\varphi_{noise}^{PM} \cong 10\ \mu rad/\sqrt{Hz}$ (from dc to 120 kHz).

These values, measured with optical signals, are approximately two times larger than those obtained with electrical signals (see Table 3.2).

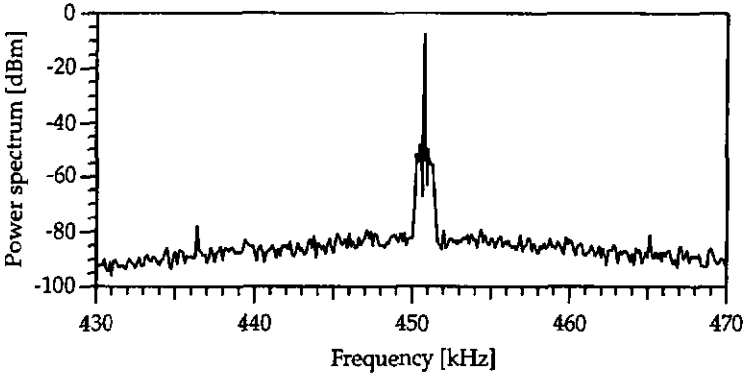


Fig. 3.43 Power spectrum of the optical input signal (30 Hz bandwidth) used to measure the phase noise of the FDPD.

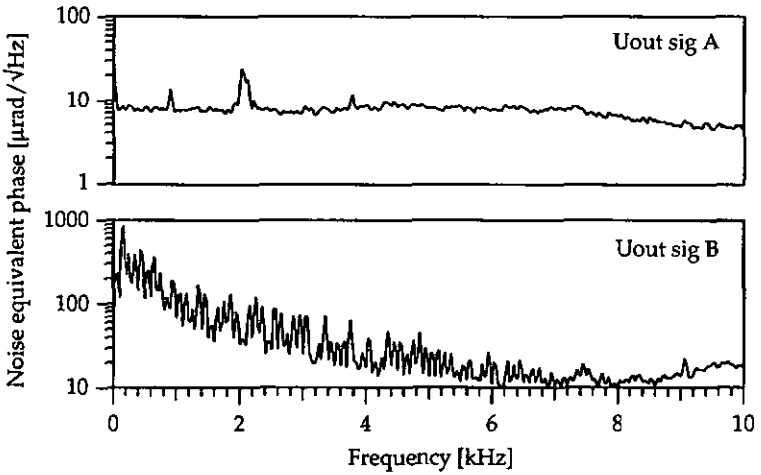


Fig. 3.44 Noise equivalent phase of the analog FDPD measured with the optical signal shown in Fig. 3.43.

3.4.2 Reference signal produced electronically or optically

We have seen in paragraph 3.1.1 that the phase-demodulation principle needs a stable reference signal. It can be created either optically with the setup shown in Fig. 3.3 or electronically from the two electrical driving signals of the acousto-optic modulators. To compare the performance of both methods, we measured the drift of the interferometer by sampling the analog outputs of the FDPD at a rate of 2 Hz during 60 minutes with either the optical or the electrical reference signal connected to the FDPD. The recovered phase as a function of time is plotted in Fig. 3.45. These measurements were carried out with CNRs of approximately 56 dB for the optical signals and 53 dB for the electrical reference (measured at 455 kHz within 30 Hz bandwidth, converted to 3 kHz bandwidth).

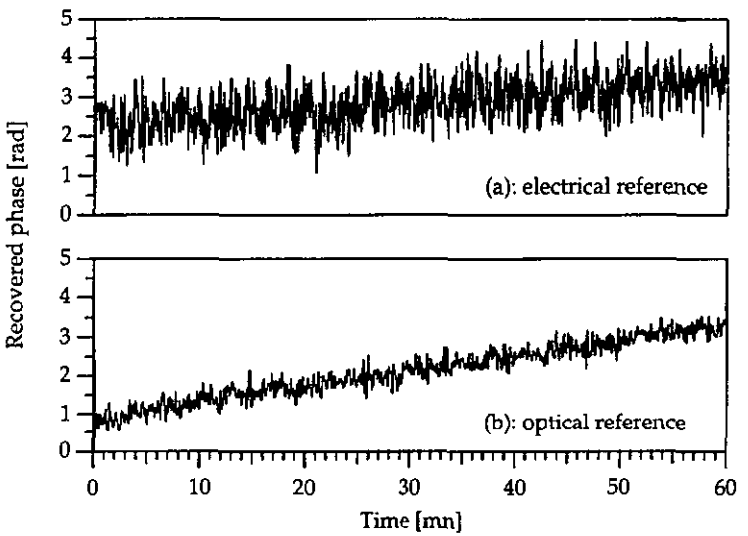


Fig. 3.45 Long term stability of the interferometer when the reference signal is produced either electronically [curve (a)] or optically [curve (b)].

From Fig. 3.45, we observe that the drift during one hour is comparable for both methods but it is obvious that the phase fluctuations are smaller when the optical reference is used instead of the electrical one. The Allan variance is one of the most common stability quantities for the characterization of the stability of

oscillators. It provides information on the fluctuations averaged over an integration time interval T [3.20]. In our context, we can use the Allan deviation (square root of the Allan variance) of the reconstructed phase to quantify the phase fluctuations of the optical setup. The Allan variance is given by the average value of the standard variances between two consecutive samples. From the measurements presented in Fig. 3.45, we obtained phase fluctuations of about $\sigma_{allan} \equiv 120$ mrad for the electrical reference and $\sigma_{allan} \equiv 30$ mrad for the optical reference.

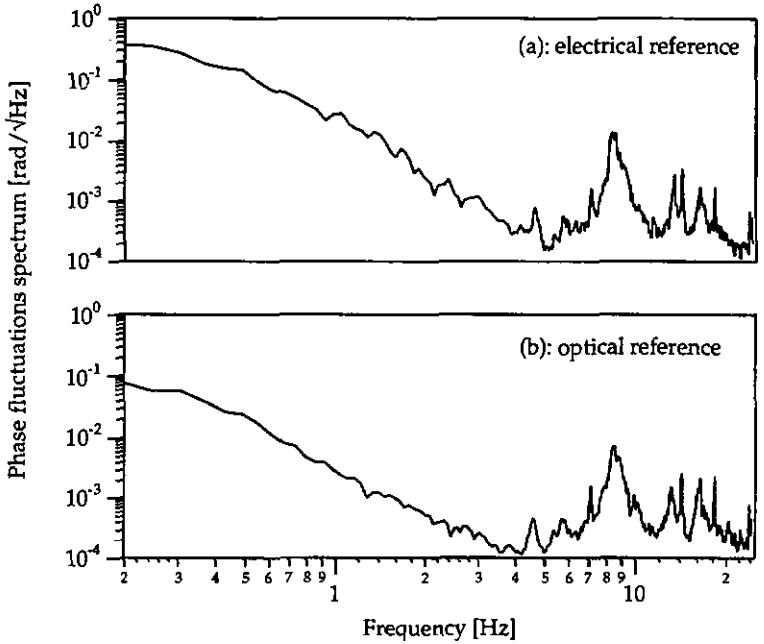


Fig. 3.46 Spectra of phase fluctuations, for optical reference [curve (a)] and electrical reference [curve (b)].

Figure 3.46 shows spectra of the phase fluctuations between 0.2 Hz and 25 Hz (measured within 0.061 Hz bandwidth) for electrical or optical reference signals. For these low-frequency measurements, we used a FFT spectrum analyzer (Stanford Research 760). As can be seen in Fig. 3.46, these spectra look very similar, except for frequencies below 5 Hz where the phase noise is smaller for the optical reference. This difference is due to a better rejection of low-frequency

vibrations produced by the noisy environment when using the setup shown in Fig. 3.3 (see § 3.1.1).

3.5 Vibration measurements

3.5.1 Small sinusoidal vibration amplitudes

When the vibration amplitude is smaller than $\lambda/2 \equiv 316$ nm (for $\alpha = 0°$, see Fig. 3.3), one channel, either channel A or B, can be used to recover the phase information in real-time. Indeed, from Fig. 3.26, we observe that when the optical phase amplitude is smaller than 2π, which corresponds to half a period of the output signals, one of the FDPD output voltage is proportional to the phase difference between the reference and probe signals. Of course, the mean phase difference must be adjusted in such a way that the used output signal is not folded. Moreover, the experimental setup has to be mechanically stable to avoid an important drift of the output signal during the experiment.

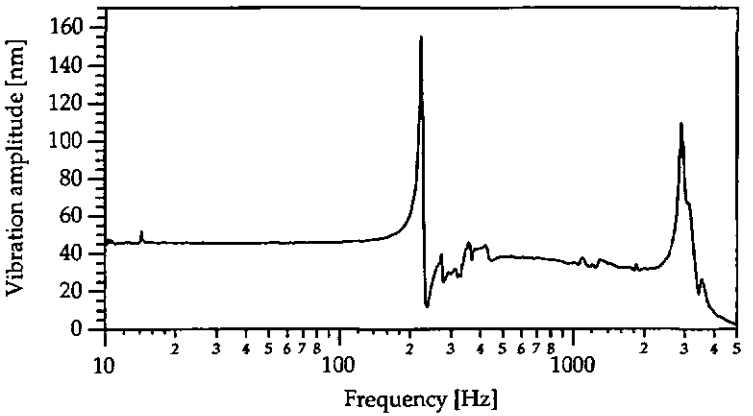


Fig. 3.47 Vibration amplitude as a function of the excitation frequency for a piezo electric transducer. The data was collected in real time employing a Network Analyzer (HP 4195).

To illustrate this possibility of on-line vibration analysis, we measured the vibration amplitude of a piezo electric transducer for 400 hundred excitation frequencies between 10 Hz and 5000 Hz. For this experiment, we used a Network

analyzer (HP 4195), to drive the PZT with a constant excitation voltage and to measure the vibration amplitude of the piezo electric transducer. The vibration amplitude as a function of the excitation frequency is shown in Fig. 3.47. The measuring time was about 10 minutes and we checked during this period that the output signal did not fold.

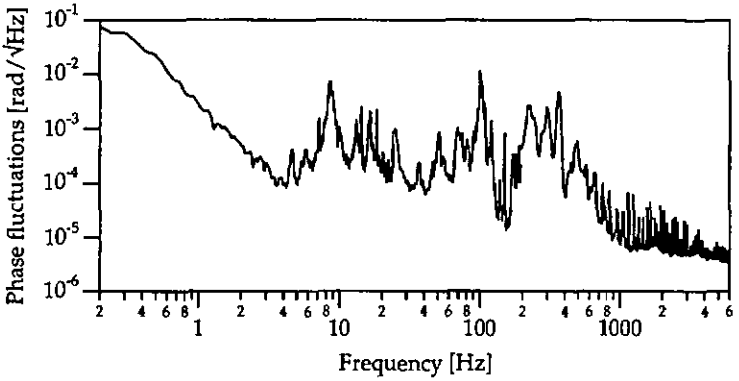


Fig. 3.48 Phase fluctuations spectrum at the analog FDPD output. The reference and probe signals were created optically with the setup shown in Fig. 3.3.

For small vibration amplitudes, the phase-demodulator is limited by the fluctuations of the output signals of the FDPD. The phase fluctuation spectrum of the analog phase-demodulator is shown in Fig. 3.48 between 0.2 Hz to 6 kHz. This spectrum was measured using a FFT spectrum analyzer (Stanford Research SR760) for CNRs of approximately 58 dB and 52 dB (within 30 Hz bandwidth, converted to 3 kHz bandwidth) for the reference and probe optical signals, respectively. The important phase fluctuations between 10 Hz and 700 Hz are due to external vibrational sources (machine-tools, electric fans, nearby road transport, etc.), to vibrations of optical elements and mechanical mounts of the experimental setup and, to air currents in the laboratory. For frequencies below 4 Hz, the phase fluctuations increase as $1/f^2$. This behavior is characteristic to the compliance curve for a typical optical table top [3.21]. The piezo electric transducer was then driven by a sinusoidal voltage of small amplitude at a frequency of 25 Hz. The reconstructed phase as a function of time is shown in Fig. 3.49. As shown in Fig. 3.50, we also measured the phase spectrum (within 0.244 Hz bandwidth) with the same sinusoidal motion. The peak amplitude at 25 Hz corresponds to

approximately 195 mrad, which compares well with the 200 mrad obtained from Fig. 3.49.

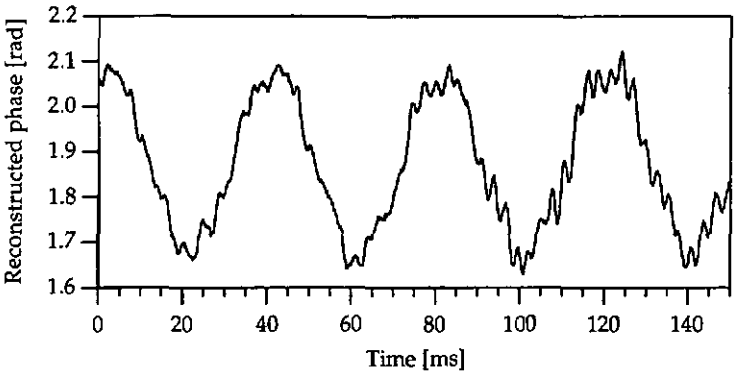


Fig. 3.49 Phase reconstruction with the analog FDPD for a sinusoidal vibration of approximately 200 mrad at a frequency of 25 Hz.

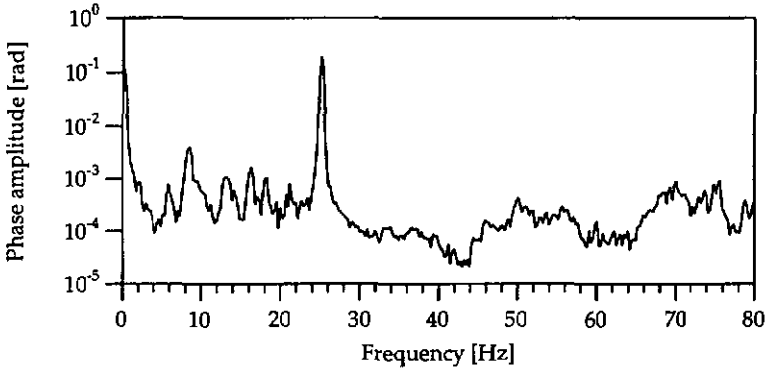


Fig. 3.50 Phase spectrum corresponding to the sinusoidal motion shown in Fig. 3.49. The peak amplitude at 25 Hz is about 195 mrad.

In order to study the influence of the input CNR on the signal-to-noise ratio (SNR) at the demodulator output, the CNR of the probe signal was changed by attenuating the optical power in the object beam. Indeed, we have seen in paragraph 3.2.2 that the CNR value for shot-noise limited detection is

proportional to the light power in the object beam [see Eq. (3.30)]. In Fig. 3.51, the signal-to-noise ratio (SNR) measured at the FDPD output is given as a function of the CNR of the probe signal, for a vibration frequency of 25 Hz [curve (a)] and 2.5 kHz [curve (b)]. The vibration amplitude was approximately 0.9 rad at 25 Hz and 0.6 rad at 2.5 kHz. The CNR of the reference signal was approximately 58 dB. In Fig. 3.51, the CNR and the SNR are given for a detection bandwidth of 3 kHz and 0.977 Hz, respectively.

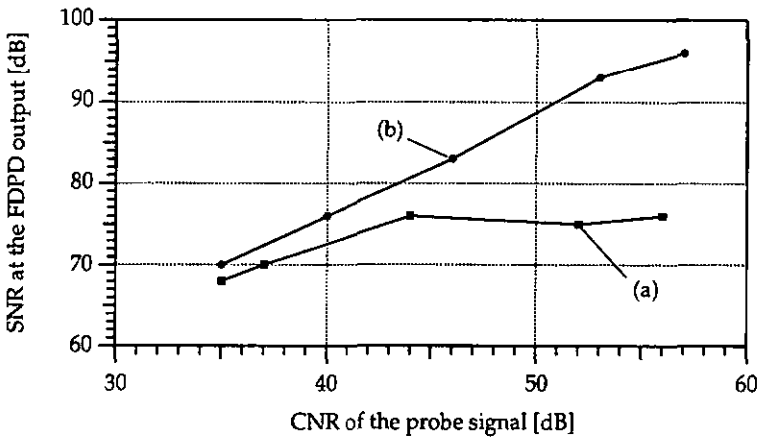


Fig. 3.51 SNR measured at the FDPD output as a function of the CNR of the probe signal for a vibration frequency of 25 Hz [curve (a)] and 2.5 kHz [curve (b)]. The CNR of the reference signal was approximately 58 dB. The CNR and the SNR are given for a detection bandwidth of 3 kHz and 0.977 Hz, respectively.

From Fig. 3.51, we observe that for a vibration frequency of 25 Hz, the output SNR is nearly constant for CNRs of the probe signal larger than about 45 dB. Indeed, even for high CNRs the noise at the demodulator output is not given by the applied input noise (limitation due to the CNR, see § 3.2.1) or by the internal demodulator noise, but rather by the phase fluctuations due to external vibrations observed between 10 Hz and 700 Hz on the spectrum in Fig. 3.50. On the other hand, when the phase demodulator is limited by the applied input noise, the output SNR (in dB) is approximately proportional to the input CNR (in dB), as expected from Eq. (3.18).

3.5.2 High frequency and large amplitude vibrations

The experimental results presented in this paragraph illustrate the use of the digital FDPD to measure large amplitudes at high vibration frequencies. In this situation, the main limitation is given by the condition that the phase change between two samples must be smaller than π to allow the system to follow accurately the phase variations. This leads to a maximum phase rate

$$\dot{\phi}_{max} = \frac{\pi}{T_s}, \tag{3.41}$$

where T_s is the sampling period. Elsewhere, the phase rate $\dot{\phi}$ is related to the frequency excursion Δf of the interference signal by

$$\dot{\phi} = 2\pi\Delta f. \tag{3.42}$$

Substituting Eq. (3.42) into Eq. (3.41), the maximum frequency excursion is then given by

$$\Delta f_{max} = \frac{1}{2T_s} = \frac{F_s}{2}, \tag{3.43}$$

where F_s is the sampling frequency. Therefore, the maximum frequency excursion Δf_{max} is equal to the Nyquist frequency. From Eq. (3.2), the maximum phase rate $\dot{\phi}_{max}$ for a periodic movement is

$$\dot{\phi}_{max} = \beta u\Omega = 2\pi\beta uF = 2\pi\phi_0 F, \tag{3.44}$$

where F is the vibration frequency and ϕ_0 is the optical phase amplitude. Using Eqs. (3.41) and (3.44), the maximum optical phase amplitude-frequency product is

$$(\phi_0 \cdot F)_{max} = \frac{1}{2T_s} = \frac{F_s}{2}. \tag{3.45}$$

One can see that the maximum optical phase amplitude-frequency product is equal to the Nyquist frequency.

From Eqs. (3.43) and (3.45), we can see that

$$\phi_0 = \frac{\Delta f}{F}, \tag{3.46}$$

where $\Delta f/F$ is commonly called the modulation index of a FM wave [3.22]. The bandwidth occupied by the interference signal is given by the Carson's rule, namely [3.23]

$$B \cong 2(\Delta f + F) = 2(\phi_0 + 1)F. \tag{3.47}$$

We emphasize that Eq. (3.47) is an approximation for the bandwidth of an FM- and PM-wave. In paragraph 3.3.1, we have mentioned that the use of band-pass filters at the demodulator inputs prevents the laser intensity noise to be transmit through the FDPD and also reduces the input sensitivity of the FDPD. From Eqs. (3.43), (3.45) and (3.47), we can see that the band-pass input filters must be designed for a central frequency equal to the heterodyne frequency Δf_h and with a bandwidth of $B = 2F_s = \Delta f_h$, because the A/D converters of digital FDPD works at one half of the heterodyne frequency.

From Eq. (3.45), the maximum amplitude-frequency product is given by

$$(u \cdot F)_{max} = \frac{1}{2\beta T_s} = \frac{F_s}{2\beta}, \tag{3.48}$$

that is, for $\beta = 14\ \mu m^{-1}$ [Eq. (3.3), Figs. 3.1 and 3.3, $2\alpha = 90°$] and $F_s = 227.5$ kHz, $(u \cdot F)_{max} = 8$ mm·Hz. Using Eqs. (3.2) and (3.41), the maximum acceleration for a periodic movement is

$$\ddot{u}_{max} = u\Omega^2 = \frac{\dot{\phi}_{max}}{\beta}\Omega. \tag{3.49}$$

Substituting Eq. (3.41) into Eq. (3.49), the maximum acceleration is then given by

$$\ddot{u}_{max} = \frac{2\pi^2}{\beta}\frac{F}{T_s} = \frac{2\pi^2 F_s}{\beta}F. \tag{3.50}$$

Following Eqs. (3.48) and (3.50), the maximum admitted amplitude u_{max} and acceleration $\ddot{u}_{max}$ as a function of the vibration frequency F is plotted in Fig. 3.52 for $\beta = 14\ \mu m^{-1}$. The curve (a) is calculated for the highest available sampling frequency F_s = 227.5 kHz and the curve (b) for the lowest available sampling frequency F_s = 14.22 kHz (see § 3.3.2).

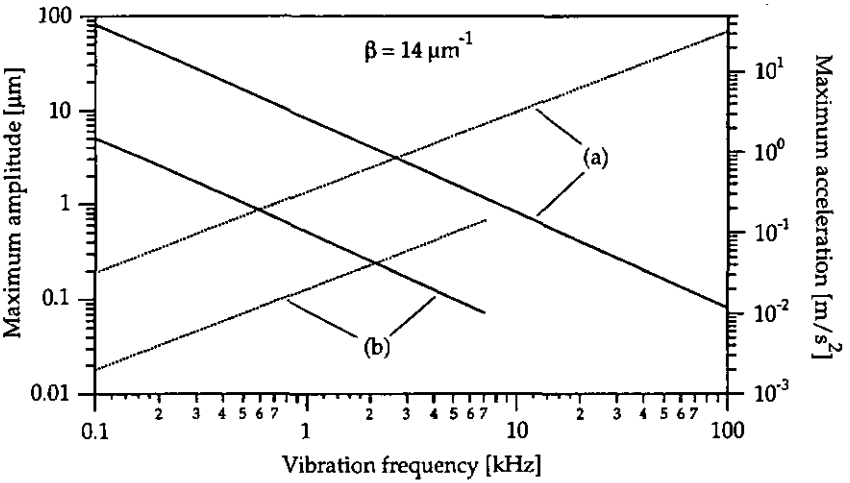


Fig. 3.52 Maximum vibration amplitude (plain line) and acceleration (doted line) as a function of the vibration frequency for the digital FDPD with F_s = 227.5 kHz [curve (a)] and F_s = 14.22 kHz [curve (b)].

For the experimental test, we employed as vibrating object a compact piezoelectric transducer with a high resonance frequency of 22 kHz (Physik Intrumente P 810.10). To avoid a reduction of this resonance frequency, we utilized as reflecting surface a piece of polished silicon (1.5 x 1.5 mm^2) fixed on the PZT by the capillary force of an oil film. The two 12-bit digital output signals were acquired at a sampling rate of 227.5 kHz and processed off-line to recover the phase information. This sampling rate leads to a maximum phase rate of 0.7 rad/μsec ($\beta \cong 14\ \mu m^{-1}$) or equivalently to a maximum speed of the vibrating surface $v_{max} \cong 50$ mm/sec.

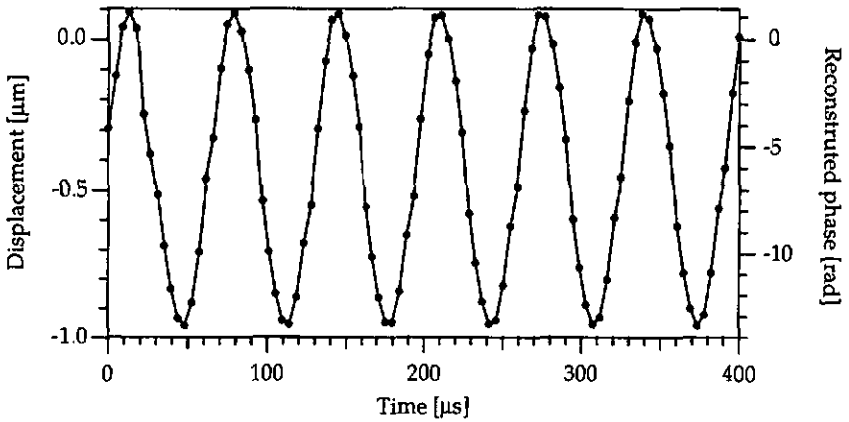


Fig. 3.53 Sinusoidal vibration of a piezo electric transducer at 15 kHz.

The reconstructed phase and the normal displacement of the surface are shown in Fig. 3.53 when the PZT was driven by a sinusoidal excitation of frequency of about 15 kHz. The peak-to-peak vibration amplitude is about 1 μm, which is nearly the maximum permitted amplitude at this excitation frequency (see Fig 3.52). Beyond the maximum phase amplitude-frequency product, the signal processing does not work properly and the reconstructed phase looks like the one shown in Fig. 3.54.

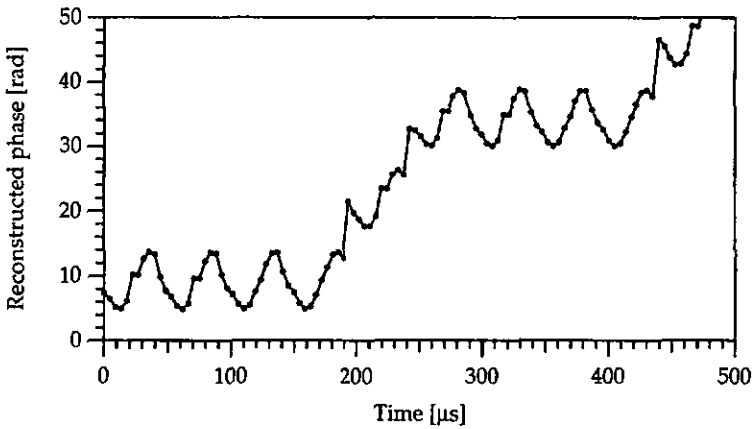


Fig. 3.54 Reconstructed phase when the vibration amplitude exceeds the maximum admitted value given by Eq. (3.43).

We also measured the vibration amplitude of the PZT at various vibration frequencies (5, 10, 15 and 20 kHz) and for several excitation voltages (1, 2, 4, 8 and 11 V). The results are shown in Fig. 3.55. The dashed curve corresponds to the maximum admitted amplitude u_{max} for a sampling frequency of F_s = 227.5 kHz (see Eq. 3.45)

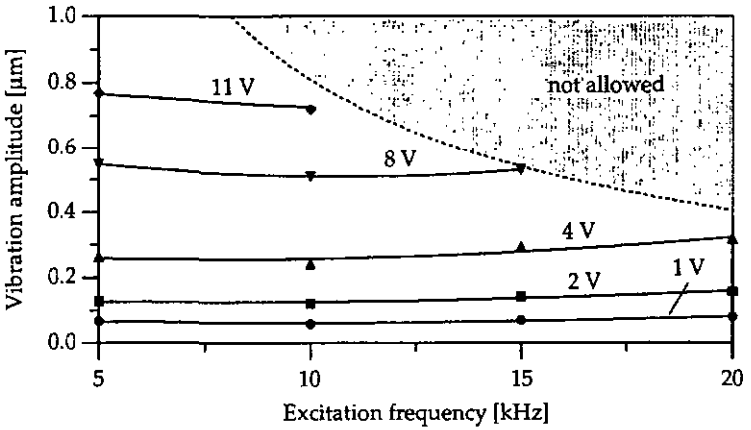


Fig. 3.55 Vibration amplitude as a function of the vibration frequency for excitation voltages of 1, 2, 4, 8 and 11 V, respectively.

The situation is quite different with the analog FDPD. Indeed, when the optical phase amplitude-frequency product becomes high, the output folded signals get very complicated. The spectrum of these signals occupies a wide frequency range and the higher frequency components will be affected by the low-pass filters. Therefore, the bandwidth of the folded signal, which is given by the phase amplitude-frequency product, should be kept sufficiently narrow to avoid distortion of the output signals. From Eqs. (3.43) and (3.45), an estimation of the maximum allowed phase amplitude-frequency product is given by

$$(\phi_0 \cdot F)_{max} = \frac{1}{2\tau} \cong \pi f_c, \tag{3.51}$$

where the sampling period τ_s of the digital FDPD has been replaced by $\tau \cong 1/(2\pi f_c)$, which represents the time constant (averaging time) of the low-pass filter. Thus, for a cut-off frequency of f_c = 6.9 kHz, the maximum phase amplitude-frequency is equal to approximately $(\phi_0 \cdot F)_{max} \cong 22$ rad · kHz .

3.6 Conclusion

In this chapter, we have presented a powerful approach for performing demodulation of heterodyne interferometric signals. For this purpose, a novel full-dynamic phase-demodulator (FDPD) for heterodyne interferometric vibration measurements has been developed and tested. Contrary to a conventional phase-demodulator, which has a limited dynamic range of ± 90°, the proposed phase-demodulator has an unlimited dynamic range. Indeed, the signal processing program recovers (off-line) the phase information, that is the displacement of the surface under test from two output signals in quadrature. It has been proven, theoretically and experimentally, that phase-demodulation is superior to frequency-demodulation at low vibration frequencies, because the displacement sensitivity is constant with vibration frequency for PM-demodulation and does not decrease as for FM-demodulation.

Two versions of the FDPD have been developed. The analog version uses active low-pass filters (cut-off frequency = 6.9 kHz) to eliminate the carrier frequency (227.5 kHz) after mixing. Furthermore, an improved version, which uses synchronized A/D converters working at one half of the heterodyne frequency (227.5 kHz), has been realized. This digital version has an extended frequency response, since a new value of the phase difference between probe and reference signals is obtained every 4.4 µsec. Consequently, the analog and digital versions allow to measure and to analyze complex and arbitrary movements of objects for vibration frequencies ranging from dc to approximately 115 kHz. This upper limit is not inherent to the principle of operation of the FDPD but is fixed by the choice of the heterodyne frequency (actually 455 kHz).

The input sensitivity of the FDPD is about 5 mV and the dynamic range is more than 80 dB. Noise measurements have shown that for a CNR = 68 dB (3 kHz bandwidth), an equivalent phase noise measured with optical signals of approximately 8 µrad/√Hz and 10 µrad/√Hz can be obtained for the analog and digital version, respectively. Actually, the performance of the FDPD is limited, for vibration frequencies below 1 kHz by the mechanical vibrations of the optical setup due to the noisy environment and for frequencies above 1 kHz by the applied input noise (limitation due to the CNR).

4. Noise resistant signal processing for white-light interferometry

Most parts of this chapter have been published as regular papers in international journals or presented at conferences [4.1, 4.2, 4.3].

In white-light interferometry, a broadband source is used instead of a highly coherent laser [4.4]. Thus, white-light interferometry allows to make absolute measurements rather than incremental measurements, overcoming the phase ambiguity limitation from which one-wavelength interferometers suffer [4.5, 4.6, 4.7]. Moreover, electronically scanned white-light interferometry [4.1, 4.2], where the interferogram is scanned in space with a photodiode array, rather than in time, permits to increase both speed and mechanical stability of the measurement system.

In this chapter, we present an original method to find the center position of the white-light fringes. It works in two steps. First, the center of gravity of the power of the interference signal is calculated to better than half a fringe period, second,. the phase of the fringe signal is determined to get the exact center position of the white-light fringes. For the implementation of this method we introduce a novel optimum sampling concept, which is based on synchronous sampling with four samples per fringe period. The theoretical predictions for systematic and statistical errors will be compared with computer simulations. Experimental results show how this novel signal processing is particularly adequate for white-light fringes scanned by linear detector arrays. Finally, we propose to apply this signal processing scheme to absolute force measurements, using the induced birefringence in an optical monomode fiber. Actually, many different white-light interferometers using either single-mode or highly birefringent fibers have been studied as force [4.2], strain [4.8] and pressure [4.9] sensors. But here, we introduce an novel solution for the receiving interferometer, which uses a Wollaston prism to compensate the phase retardation induced by the force [4.1, 4.2, 4.8].

The theoretical analysis and experimental results presented in this chapter show that this proposed method to analyze white-light interferograms is simple to operate, fast, accurate and extremely noise resistant.

4.1 Signal processing

4.1.1 Center position of white-light fringes

A typical white-light fringe signal, as shown in Fig. 4.1, can be described by

$$s(x) = s_{dc}(x) + s_{ac}(x) = s_{dc}(x) + F(x - x_s)\cos\left[\frac{2\pi(x - x_s)}{\Lambda}\right] + g(x,\sigma), \qquad (4.1)$$

where Λ is the fringe period and F(x) is the fringe envelope, which is given by the modulus of the complex degree of coherence (see § 2.2.2). The dc-part of the signal s_{dc} can change slightly from one pixels to another because the light intensity is not constant through the photodiode array. In Eq. (4.1) the noise is taken into account by g(x,σ), which is assumed to be Gaussian noise with variance σ^2.

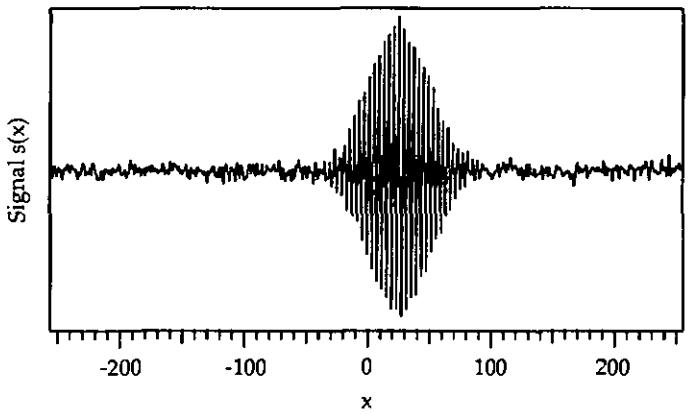


Fig. 4.1 Typical white-light fringe signal s(x) with Gaussian noise as described by Eq. (4.1).

This signal is sampled at the positions $x_n = n \cdot b$, where n is an integer with $-N/2 \leq n \leq N/2$ and b is the center-to-center spacing of the photodiode array pixels used to sampled the interferogram. To determine the center position x_s of the white-light fringes, the center of gravity x_{s1} of the power of the ac-part s_{ac} is calculated from its total power A and its first moment B, using the equations

$$A = \sum_n [s_{ac}(x_n)]^2, \tag{4.2}$$

$$B = \sum_n n \cdot [s_{ac}(x_n)]^2, \tag{4.3}$$

and

$$x_{s1} = \frac{B}{A}. \tag{4.4}$$

Note that both values x_s and x_{s1} are in general fractional values of the sample positions. For a signal without noise, we get for the center position exactly $x_{s1} = x_s$. However, for a signal with noise, x_{s1} has a statistical error and a systematic error, as explained in the next paragraph.

4.1.2 Statistical and systematic errors

In this paragraph, we only consider the ac-part $s_{ac}(x)$ of the white-light fringes as described by Eq. (4.1). This signal can be decomposed in two parts. First, the fringe signal without noise f(x) given by

$$f(x) = F(x - x_s)\cos\left[\frac{2\pi(x - x_s)}{\Lambda}\right], \tag{4.5}$$

where Λ is the fringe period and F(x) is the fringe envelope. Second, the Gaussian noise g(x, σ) with variance σ^2 (see Annex C). We get the total power A of this signal and the first moment B of the signal power by

$$A = \sum_n [f(x_n) + g(x_n)]^2, \tag{4.6}$$

and

$$B = \sum_n n[f(x_n) + g(x_n)]^2, \tag{4.7}$$

with $-N/2 \le n \le N/2$ and where $f(x_n)$ and $g(x_n)$ are the sampled values of signals f(x) and g(x), respectively.

The center position x_s of s_{ac} is given by

$$x_s = \frac{B}{A}. \tag{4.8}$$

The average of A is equal to

$$\begin{aligned}\langle A\rangle &= \left\langle \sum_n \left[f(x_n)+g(x_n)\right]^2 \right\rangle = \sum_n \left\langle \left[f(x_n)+g(x_n)\right]^2 \right\rangle \\ &= \sum_n f(x_n)^2 + \sum_n \left\langle g(x_n)^2 \right\rangle + 2\sum_n f(x_n)\langle g(x_n)\rangle .\end{aligned} \tag{4.9}$$

Using the results obtained in Annex C for the statistical properties of sampled Gaussian noise, we can substitute Eqs. (C.2) and (C.10) into Eq. (4.9), to find that

$$\langle A\rangle = A_f + (N+1)\cdot\sigma^2, \tag{4.10}$$

where A_f is the total power of the ac signal without noise, given by

$$A_f = \sum_n f(x_n)^2 \tag{4.11}$$

Similarly, we get for the average of the first moment of the signal power B, using Eqs. (C.2) and (C.17),

$$\begin{aligned}\langle B\rangle &= \left\langle \sum_n n\left[f(x_n)+g(x_n)\right]^2 \right\rangle = \sum_n \left\langle n\left[f(x_n)+g(x_n)\right]^2 \right\rangle \\ &= \sum_n nf(x_n)^2 + \sum_n n\left\langle g(x_n)^2 \right\rangle + 2\sum_n nf(x_n)\langle g(x_n)\rangle = B_f ,\end{aligned} \tag{4.12}$$

where

$$B_f = \sum_n nf(x_n)^2 \tag{4.13}$$

is the first moment of the ac signal power.

Using Eqs. (4.10) and (4.12), the average value of the center position x_s is given by

$$\langle x_s \rangle = \frac{\langle B \rangle}{\langle A \rangle} = \frac{B_f}{A_f + (N+1)\sigma^2}. \tag{4.14}$$

To calculate the statistical error δx_s, we use the standard equations for the propagation of errors [4.10]

$$\delta x_s{}^2 = \left\langle \Delta A^2 \right\rangle \left(\frac{\partial x_s}{\partial A} \right)^2 + \left\langle \Delta B^2 \right\rangle \left(\frac{\partial x_s}{\partial B} \right)^2 + 2\langle \Delta A \Delta B \rangle \left(\frac{\partial x_s}{\partial A} \right) \left(\frac{\partial x_s}{\partial B} \right), \tag{4.15}$$

with $< \Delta A \Delta B >$ being the covariance between A and B. This yields

$$\delta x_s{}^2 = \frac{1}{\langle A \rangle^2} \left[\langle x_s \rangle^2 \left\langle \Delta A^2 \right\rangle - 2\langle \Delta A \Delta B \rangle \langle x_s \rangle + \left\langle \Delta B^2 \right\rangle \right]. \tag{4.16}$$

After some calculations and using the statistical properties of sampled Gaussian noise (see Annex C), we find for the variance of A

$$\left\langle \Delta A^2 \right\rangle = 4\sigma^2 A_f + 2(N+1)\sigma^4, \tag{4.17}$$

and for the variance of B

$$\left\langle \Delta B^2 \right\rangle = 4\sigma^2 C_f + \frac{N(N+1)(N+2)}{6}\sigma^4, \tag{4.18}$$

where

$$C_f = \sum_n n^2 f(x_n)^2, \tag{4.19}$$

is the second moment of ac signal power. The covariance between A and B is found to be equal to

$$\langle \Delta A \Delta B \rangle = \left\langle (A - \langle A \rangle) \cdot (B - \langle B \rangle) \right\rangle = 4\sigma^2 B_f. \tag{4.20}$$

From Eqs. (4.11), (4.13) and (4.19), it is helpful to express A_f, B_f and C_f as a function of the center position x_s, namely

$$A_f(x_s) = A_f(0) = A_{f0} = A_f \quad \forall\, x_s, \tag{4.21}$$

$$B_f(x_s) = A_f \cdot x_s \;\Rightarrow\; B(0) = B_{f0} = 0, \tag{4.22}$$

and

$$C_f(x_s) = C_f(0) + A_f \cdot x_s^2 \;\Rightarrow\; C_f(0) = C_{f0} \neq 0 \text{ but } C_f(x_s) > C_{f0} \;\forall\, x_s \neq 0, \tag{4.23}$$

where A_{f0}, B_{f0} and C_{f0} are the values of A_f, B_f and C_f for $x_s = 0$. We would like to emphasize that Eqs. (4.21) to (4.23) are valid for all fringe envelope functions and that A_{f0} and C_{f0} depend only on the fringe signal f(x) (signal without noise).

For the further calculations we assume that the shape of the envelope F(x) is Gaussian, given by

$$F(x) = m \exp\left[\frac{(x - x_s)^2}{\Delta x^2}\right], \tag{4.24}$$

where Δx is the width of the envelope and m is the fringe contrast. We can now find an analytical expression for A_f with

$$A_f = \sum_{n=-N/2}^{N/2} f(x_n)^2 \cong \int_{-\infty}^{\infty} f(x)^2 dx \qquad \text{if } N >> 1, \tag{4.25}$$

which yields then

$$A_f = A_{f0} \cong \frac{m^2 \Delta x}{2} \sqrt{\frac{\pi}{2}}. \tag{4.26}$$

Similarly, we find

$$C_{f0} \cong \frac{m^2 \Delta x^3}{8} \sqrt{\frac{\pi}{2}} = \frac{\Delta x^2}{4} A_f. \tag{4.27}$$

Substituting Eq. (4.22) into Eq. (4.14) we find

$$\langle x_s \rangle = \frac{\langle B \rangle}{\langle A \rangle} = \frac{x_s}{1+(N+1)\dfrac{\sigma^2}{A_f}}. \tag{4.28}$$

If $(N+1)(\sigma^2/A_f) << 1$, which is usually the case, we obtain finally

$$\langle x_s \rangle \cong x_s \left[1-(N+1)\frac{\sigma^2}{A_f}\right]. \tag{4.29}$$

Equation (4.29) shows that $\langle x_s \rangle$ has a systematic error Δx_s given by

$$\Delta x_s = \langle x_s \rangle - x_s \cong -(N+1)\frac{\sigma^2}{A_f} x_s. \tag{4.30}$$

This systematic error is proportional to x_s and to the noise power. Thus, for $x_s = 0$, the systematic error becomes zero and the statistical average of the center position $\langle x_s \rangle$ is equal to x_s.

Substituting Eqs. (4.10), (4.17), (4.18) and (4.20) into Eq. (4.16), we find

$$\delta x_s^2 = \frac{4\sigma^2 A_f}{\left[A_f+(N+1)\sigma^2\right]^2}\left[\left(\langle x_s \rangle - x_s\right)^2 + \frac{12(N+1)\langle x_s \rangle^2 + N(N+1)(N+2)}{24}\frac{\sigma^2}{A_f} + \frac{C_{f0}}{A_f}\right]. \tag{4.31}$$

The statistical error δx_s becomes then

$$\delta x_s \cong \frac{2\sigma\sqrt{A_f}}{A_f+(N+1)\sigma^2}\left[\Delta x_s^2 + \frac{12(N+1)x_s^2 + N(N+1)(N+2)}{24}\frac{\sigma^2}{A_f} + \frac{C_{f0}}{A_f}\right]^{1/2}, \tag{4.32}$$

where $\Delta x_s^2 = \left(\langle x_s \rangle - x_s\right)^2$ is given for a Gaussian envelope by Eq. (4.30) and $\langle x_s \rangle \cong x_s$. For $x_s = 0$, the statistical error δx_s is reduced to

$$\delta x_s(x_s = 0) = \frac{2\sigma \cdot \sqrt{A_f}}{A_f + (N+1)\sigma^2} \left[\frac{N(N+1)(N+2)}{24} \cdot \frac{\sigma^2}{A_f} + \frac{C_{f0}}{A_f} \right]^{1/2}, \tag{4.33}$$

since for $x_s = 0$, we have $\langle x_s \rangle = 0$. The statistical error δx_s is thus minimal when $x_s = 0$.

The validity of this theoretical analysis was checked by comparison with computer simulations using the scientific software Igor from WaveMetrics. In Figs. (4.2) and (4.3), the mean and the variance of x_s are represented as a function of the noise standard deviation σ when $x_s = 15$ is fixed. To compute $\langle x_s \rangle$ and $\langle \Delta x_s^2 \rangle$, we used the following values, namely: N = 256, m = 0.9, $\Delta x = 25$ and $\Lambda = 4$. It is interesting to note that for $\sigma \neq 0$, the mean value of the center position $\langle x_s \rangle$ is always smaller than the true value x_s, as explicitly given by Eq. (4.30). The systematic and statistical error are shown in Figs. (4.4) and (4.5), when the center position is changed for a constant noise of $\sigma = 3$ %. The other fixed values are the same as previously. We would like to emphasize again that for $x_s = 0$, the systematic error Δx_s becomes zero and the statistical error δx_s is minimal. As we can see from Figs. (4.2) to (4.5), the agreement between the theoretical predictions and the computer simulations is very good.

The quality of the fringe signal can be defined by the signal-to-noise ratio (SNR) at the maximum of the envelope. From Eqs. (4.1) and (4.24), we get SNR = m^2/σ^2. Figure 4.6 shows $\Delta x_s/x_s$ and δx_s (x_s) as a function of the SNR calculated with N = 160 and $\Delta x = 35$. Some values of the relative systematic error $\Delta x_s/x_s$ and the statistical error δx_s are shown in Table 4.1 for $\Delta x = 35$ pixels and for different SNR.

	SNR [dB]			
Error	20	30	40	50
$\Delta x_s/x_s$ (N = 360)	0.14	0.016	0.0016	0.00016
δx_s (N=360)	1.27	0.27	0.076	0.024
δx_s (N=160)	0.78	0.24	0.075	0.024

Table 4.1 Theoretical values of the relative systematic error $\Delta x_s/x_s$ and the statistical error δx_s (pixels) for a Gaussian envelope of width $\Delta x_s = 35$ pixels.

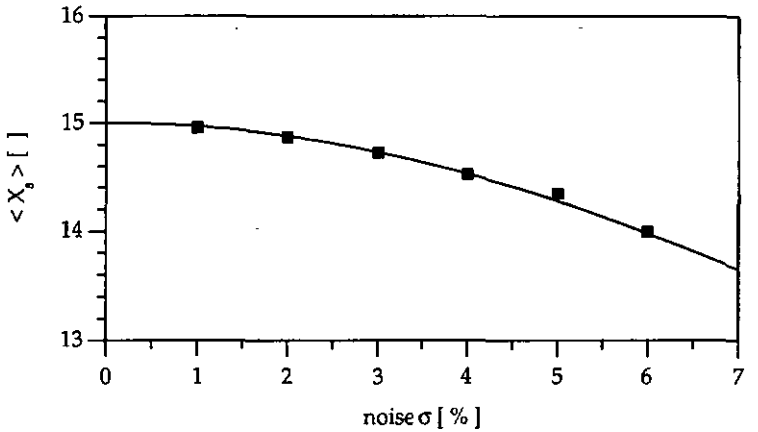


Fig. 4.2 Theoretical prediction (plain line) and computer simulations (dots) for the mean value of the center position x_s as a function of σ.

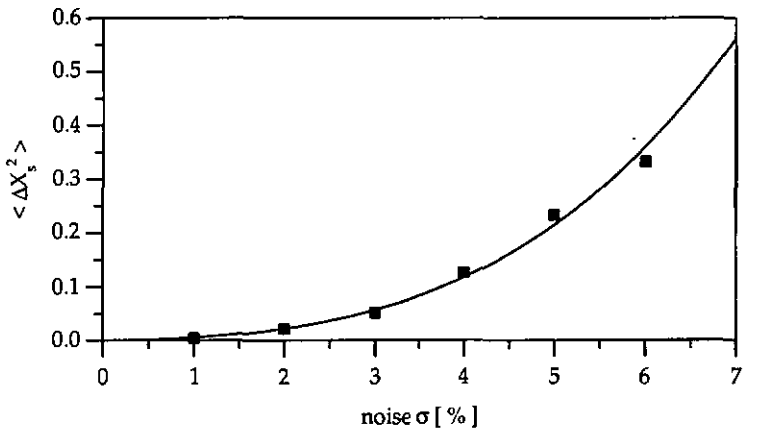


Fig. 4.3 Theoretical prediction (plain line) and computer simulations (dots) for the variance of the center position x_s versus σ.

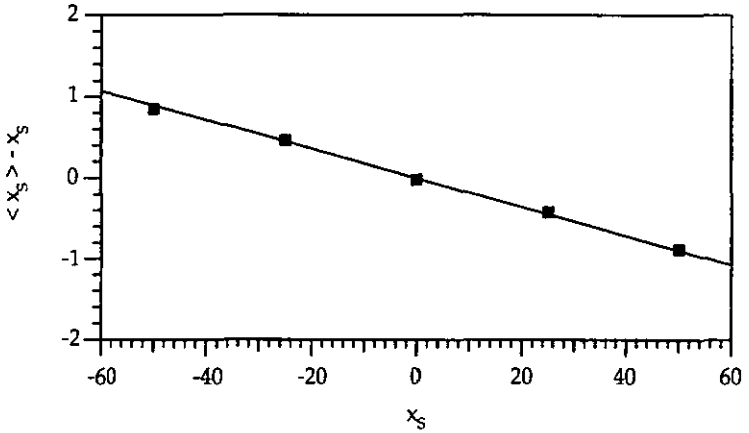


Fig. 4.4 Theoretical prediction (plain line) and computer simulations (dots) for the systematic error as a function of x_S.

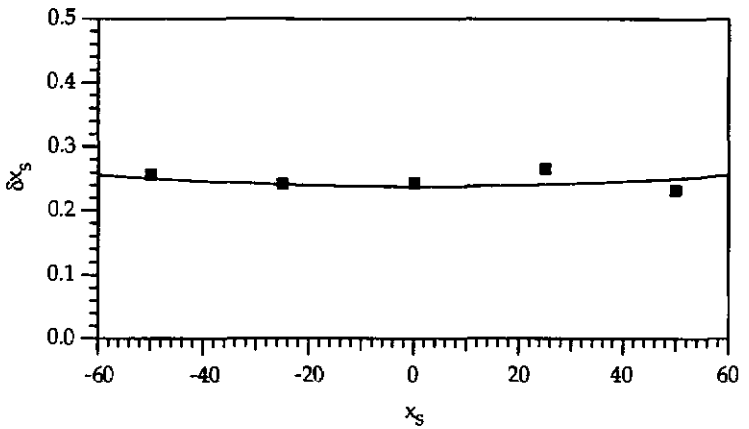


Fig. 4.5 Theoretical prediction (plain line) and computer simulations (dots) for the statistical error as a function of x_S.

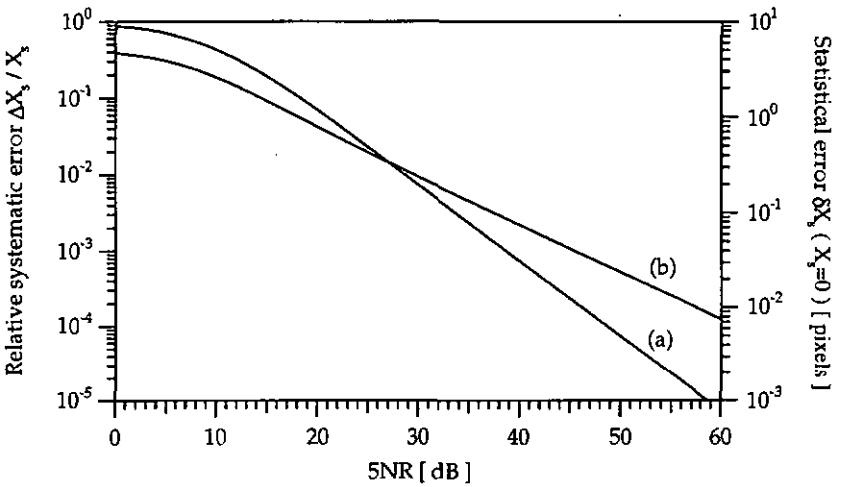


Fig. 4.6 Systematic and statistical errors as a function of the signal to noise ratio with N = 160 and Δx = 35. (a) relative systematic error and (b) statistical error.

Equations (4.29) and (4.32) show that for $x_s = 0$ the systematic error Δx_s disappears and the statistical error δx_s becomes minimum. Therefore the accuracy can be improved by iteration. First, we calculate the center position x_{s1} of the signal as described above. Second, we take as new origin the sample position x_{02} next to x_{s1}. Thus, the new center position x_{s2} is close to zero. This procedure reduces both the systematic and the statistical error. The statistical error δx_s can further be reduced by limiting the number of samples (N+1) around the new origin x_{02}. However, the number of samples must be large enough to cover the whole range where the fringes are visible and symmetrically distributed around the new origin x_{02}. Once the error for the center of gravity is smaller than half a fringe period Λ, the phase $\Phi = -2\pi x_s/\Lambda$ of the white-light fringes can be determined without ambiguity.

4.1.3 Optimum sampling concept

For the implementation of the described method we introduced an optimum sampling concept, which is based on synchronous sampling with four samples per fringe period, i.e. $b = \Lambda/4$ as shown in Fig. 4.7. The white-light fringe signal without noise is given, as previously mentioned, by

$$s(x) = s_{dc}(x) + F(x - x_s)\cos\left[\frac{2\pi x}{\Lambda} + \Phi\right], \tag{4.34}$$

with $\Phi = -(2\pi x_s)/\Lambda$. The range R of the sampled signal can be written as

$$R = 2\pi M \frac{b}{\Lambda}, \tag{4.35}$$

where M is the number of available pixels on the photodiode array. Using a small number of pixels per fringe period is advantageous, because it allows to enhance the range of measurement R and to reduce the cost of the detector.

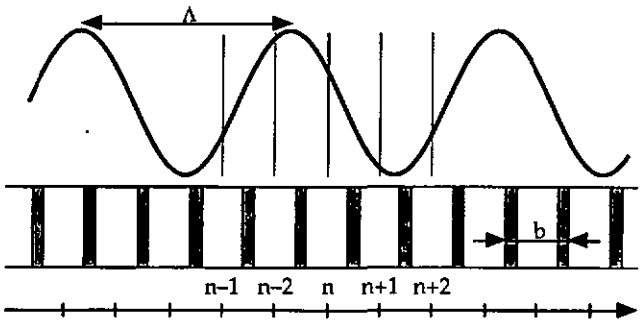


Fig. 4.7 Synchronous sampling with four samples per interference fringe period.

Two local values of the signal s(x) are calculated per period, namely

$$S_n = \frac{1}{2}[s(x_{n-1}) - s(x_{n+1})] \cong F(x_n)\sin\left(n\frac{\pi}{2} + \Phi\right), \tag{4.36}$$

and

$$C_n = \frac{1}{2}\left[s(x_n) - \frac{1}{2}[s(x_{n+2}) + s(x_{n-2})]\right] \cong F(x_n)\cos\left(n\frac{\pi}{2} + \Phi\right). \tag{4.37}$$

The local power of the ac signal is then given by

$$F_n^2 = F(x_n)^2 = C_n^2 + S_n^2, \tag{4.38}$$

and F_n is a good approximation for the local value of the envelope. If we use F_n in Eqs. (4.2) to (4.4) instead of $s_{ac}(x_n)$ to calculate the center position x_s, the systematic error given in Eq. (4.30) is even reduced by a factor of 7/16 (see Annex D). The phase Φ of the fringe signal is determined with the help of the complex value

$$C_n + iS_n = F_n \exp\left\{\frac{in\pi}{2}\right\} \exp\{i\Phi_n\}, \tag{4.39}$$

from which the local phase Φ_n is obtained through

$$\Phi_n = \arg\left\{(-i)^n (C_n + iS_n)\right\}. \tag{4.40}$$

Now, the phase Φ is calculated as the average of Φ_n over as many periods as possible with a reasonable SNR. The center position of the fringes expressed in pixels is then

$$x_\Phi = \frac{2\Phi}{\pi}, \tag{4.41}$$

since 4 pixels are equal to one fringe. Therefore, in Fig. 4.6 and in Table 4.1, an error of $\delta x_s = 1$ corresponds now to a quarter of a fringe and for $\delta x_s \leq 0.66$ the probability to find the center of gravity x_s within the correct fringe period is better than 99.7 %.

4.2 Experimental validation of the theoretical model

4.2.1 Optical setup

This novel signal processing method for white-light interferometry has been investigated with the help of the polarization interferometer shown in Fig. 4.8. The broad-band source was a LED emitting at $\lambda \cong 880$ nm (ABB HAFO 1A225) with a spectral width $\Delta\lambda$ of about 50 nm (coherence length $\cong 16$ µm). After the lens L_1 the beam is collimated with a diameter of approximately 1 mm. The linearly polarized light is oriented at 45° with respect to the birefringence axis of the Soleil-Babinet compensator (SBC). Therefore, two linear polarizations shifted

in phase by the retardation of the SBC are obtained at its output. The beam is then enlarged by the lens L_2 and L_3 to fill the whole aperture (15 mm) of the Wollaston prism (W). The retardation varies linearly across the Wollaston. This causes the relative phase between the two polarizations to change linearly across the beam section. Interference between this two polarizations is then obtained by using a dichroic sheet polarizer (DSP), oriented at 45° with respect to these polarizations.

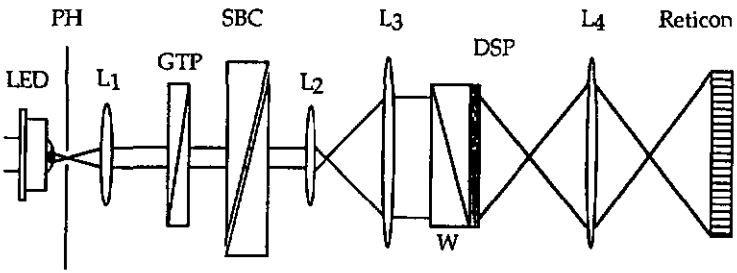


Fig. 4.8 Experimental set-up: PH: pinhole; GTP: Glan-Thomson polarizer.

The detector was a linear photodiode array of 512 pixels with a center-to-center spacing of 25 μm (EG & G Reticon 5 Serie) and a line scan time of 30 ms. The Wollaston was designed to produce fringes of about 100 μm separation. Therefore synchronous sampling with 4 pixels per period was obtained for 1:1 imaging by the lens L_4. By adjusting the distance between the lens L_4 and the detector, the imaging can be exactly adjusted, however, the results showed that this adjustment is not critical. The width of the envelope of the fringe signal (half power points) was equal to about 15 fringes, which corresponds to $\Delta x = 35$ pixels in Eq. (4.24).

4.2.2 Experimental results

The fringe signal s(x) obtained from the photodiode array is shown in Fig. 4.9. The average irradiance is not uniform across the array and therefore the number of sampled pixels was limited to within the range between the half irradiance points of the light spot on the detector ($N \equiv 360$).

The first measurement, with the origin $x_{01} = 0$, gives the center position x_{s1}. Then, to suppress the systematic error and to lower the statistical error, a new

origin x_{02} close to x_{s1} is defined and the number of samples around x_{02} is reduced to a width slightly larger than the fringe signal (N = 160). With these conditions a new value x_{s2} for the center is obtained. The center position of the fringes x_Φ, expressed in pixels, is then calculated following Eqs. (4.40) and (4.41) from the average of the local phases Φ_n over about 40 pixels around x_{s2}. The non-uniformity of the light spot on the photodiode array involves a distortion of the signal which influences the center of gravity but not the phase of the fringes. Therefore, it is not necessary to compensate this non-uniformity, unless the resulting error for the center of gravity becomes larger than half a fringe period.

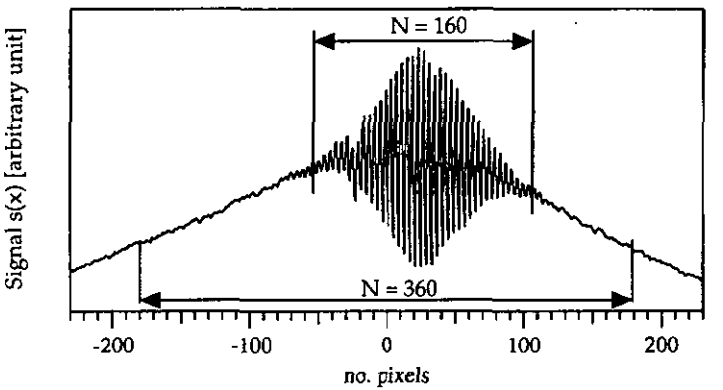


Fig. 4.9 Fringe signal s(x) obtained from the photodiode array (SNR ≅ 63 dB).

The experimental results for the center of gravity x_{s1} and x_{s2} versus the SBC retardation are shown in Fig. 4.10 by the squares and circles, respectively. The statistical reproducibility was measured to be $\delta x_{s1} \cong 0.26$ pixels and $\delta x_{s2} \cong 0.18$ pixels, which corresponds to a SNR of about 31 dB (see Table 4.1). The observed deviations from the linear fit are reproducible and have a periodicity of 2π in the retardation of the SBC. This effect could not be explained yet, but simulations showed that it seems to be independent of the parameters involved in the signal processing.

The experimental results for the calculated phase Φ, expressed in numbers of pixels x_Φ, are shown in Fig. 4.10 by the crosses. The linearity is nearly perfect and the statistical reproducibility $\delta_{x\Phi}$ was about $4 \cdot 10^{-3}$ pixels. The necessity for

iteration is also apparent in Fig. 4.10 from the difference between the squares (before iteration) and the circles (after iteration), which is nearly four pixels or one fringe order.

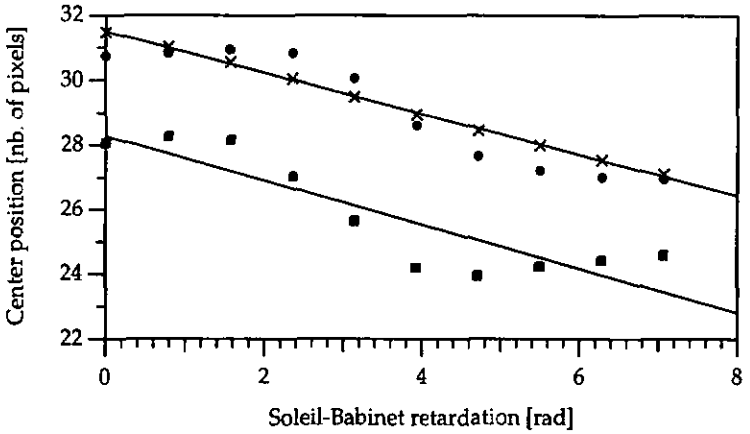


Fig. 4.10 Center position of the white-light signal versus SBC retardation for SNR ≡ 31 dB: squares, center of gravity x_{s1}; circles, center of gravity x_{s2} (after one iteration); crosses, zero fringe position x_Φ (phase).

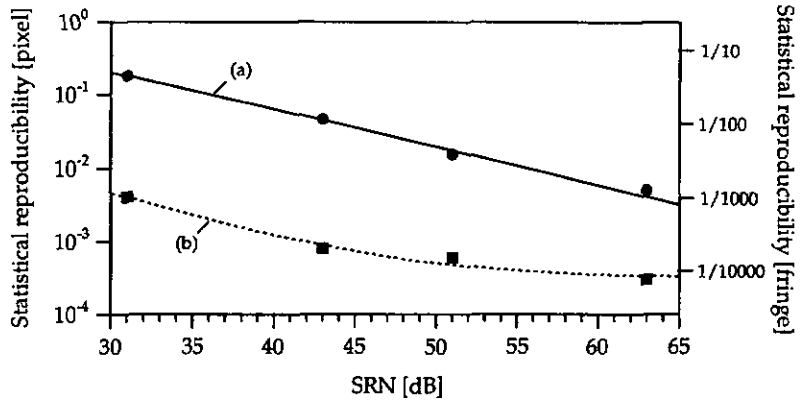


Fig. 4.11 Statistical reproducibility of the center position x_{s2} [curve (a)] and the phase Φ [curve (b)] as a function of the signal-to-noise ratio SNR.

The statistical reproducibility of the center position x_{s2} and the phase Φ was also measured for different SNRs, which were obtained by attenuating the light intensity with gray filters, with the assumption that the noise is mainly electrical noise from the photodiode array. The results are shown in Fig. 4.11 and, for the center of gravity, agree well with the theory. As can be seen from Figs. 4.10 and 4.11, the presented signal processing works perfectly well, even with a SNR as low as 30 dB.

As we will see in the next paragraph, we used this method to measure the absolute force from the induced birefringence in an optical fiber, which replaces the Soleil-Babinet compensator in Fig. 4.8.

4.3 Force measurement by white-light interferometry

4.3.1 Polarizing interferometer

The heart of the force sensor is a polarizing interferometer, as shown in Fig. 4.12. Two orthogonal linear polarizations are injected into a single-mode fiber. The applied force induces a linear birefringence. Thus, two orthogonal polarizations shifted in phase are obtained at the fiber output. A linear polarizer is then used to cause the two polarizations to interfere, upon which they are detected.

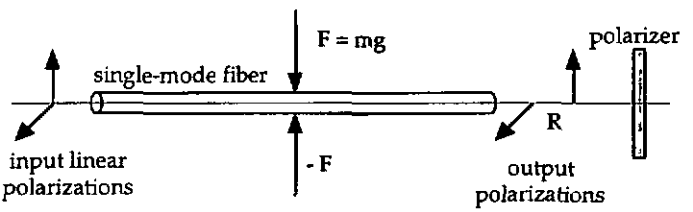


Fig. 4.12 Polarizing interferometer (birefringence).

The retardation Φ between the two orthogonal polarizations is given by [4.11]

$$\Phi = \Phi_1 - \Phi_2 = \frac{8Cmg}{\lambda r}, \tag{4.42}$$

where C is the stress-optic coefficient, m is the mass, g is the gravity constant, λ is the wavelength of light and r is the fiber radius. For fused silica and for λ = 633 nm, the stress-optic coefficient has a value of $C = -3.7 \cdot 10^{-12}$ m^2/N [4.12]. From Eq. (4.42), the retardation dispersion is given by

$$\frac{1}{\Phi}\frac{d\Phi}{d\lambda} = \frac{1}{C}\frac{dC}{d\lambda}\frac{1}{\lambda}, \tag{4.43}$$

and for λ = 633 nm, we have $(1/\Phi)d\Phi/d\lambda \cong -1.5\ \mu m^{-1}$ [4.13]. Using Eqs. (4.42) and (4.43), the sensor sensitivity at λ = 880 nm is then, $\Phi/m \cong 7.23$ rad/kg (2r = 80 μm) or $\Phi/m \cong 4.62$ rad/kg (2r = 125 μm).

The force sensor is shown in Fig. 4.13. This setup is similar to that described in chapter 4.2 (see Fig. 4.8), with a single-mode optical fiber replacing the Soleil-Babinet compensator. The polarization of the light is oriented at 45° with respect to the induced birefringence axis. Since the source is encapsulated in a case with a built-in glass lens, the use of a supplementary optical system to launch light into the fiber is not necessary. The optical power injected into the fiber was approximately 1 μW. At the output of the polarizing interferometer, we measured a SNR of approximately 44 dB.

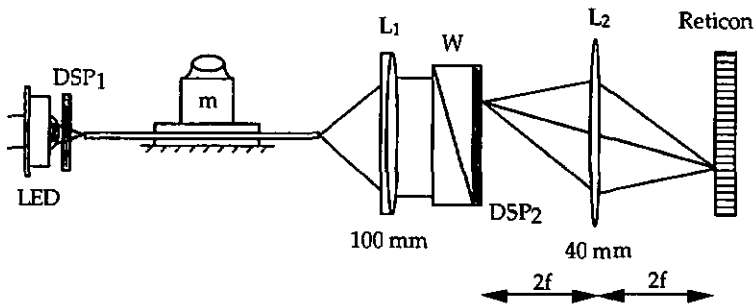


Fig. 4.13 Optical setup for absolute force measurement. DSP: dichroic sheet polarizer; W: Wollaston prism.

Contrary to previous experiments [4.12], which used a signal processing based on two signals in quadrature to count the number of periods and to interpolate the phase accurately (incremental measurement), the method proposed here allows

to make absolute (non-incremental) force measurements from the induced birefringence in an optical fiber.

The signal processing method is based on the determination of the center position x_s of the detected white-light fringe signal shown in Fig. 4.14. This is achieved in three steps (see § 4.2.2). First, the center of gravity x_{s1} is calculated with the origin $x_{01} = 0$. The number of sampled pixels N is limited to within the range between the half irradiance points of the light spot on the detector (N ≅ 360 pixels). Then to suppress the systematic error and to reduce the statistical error, a new origin x_{02} close to x_{s1} is defined and the number of samples M is reduced to a width only slightly larger than the fringe signal (M = 160 pixels). With these conditions, a new value x_{s2} for the center of gravity is obtained.

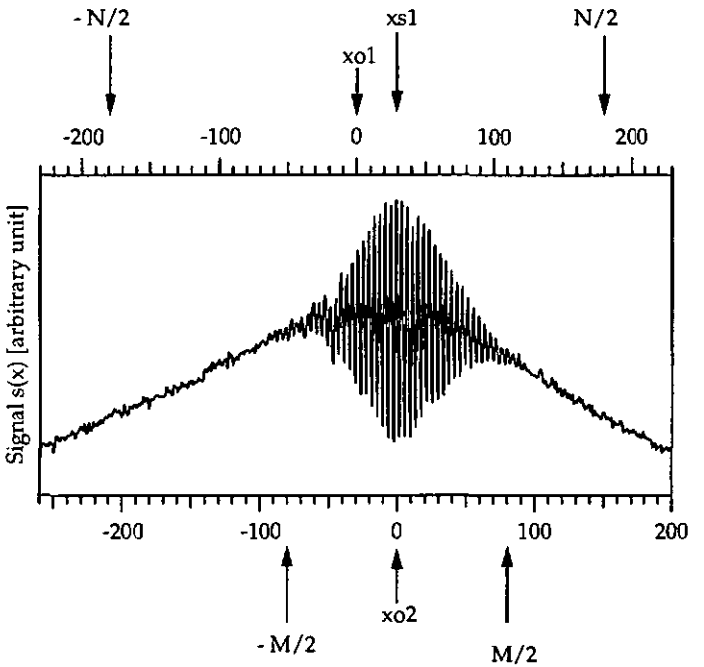


Fig. 4.14 Fringe signal obtained from the photodiode array.

The center position of the fringes x_Φ, expressed in pixels, is then calculated following Eqs. (4.40) and (4.41) from the average of the local phases Φ_n over about

40 pixels around x_{s2} as shown in Fig. 4.15. This curve is approximately flat around the origin x_{s2}, because the condition of four pixels per period was obtained by adjusting the distance between the lens L_4 and the detector (see Fig. 4.8).

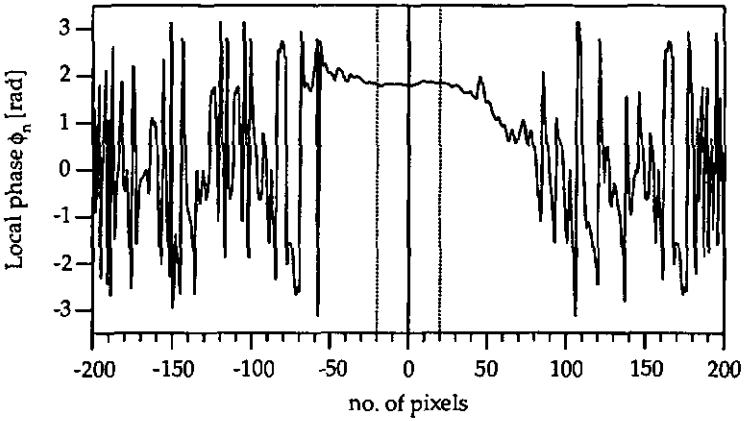


Fig. 4.15 Local phase of the fringe signal.

4.3.2 Experimental results

Several types of optical fibers have been investigated and tested in view of their utilization in an optical fiber force sensor. As a preliminary remark, we would like to mention that the employment of certain fibers was not possible for the white-light force sensor by reason of the presence of speckles within the output light spot. For example, we tried to used fibers where the coating was first removed to avoid relaxation effects when the force is applied on the fiber. But in this situation we always observed speckles at the fiber output.

For each type of fiber, the center position x_{s1} and x_{s2} (after one iteration) have been measured as a function of the applied mass on the fiber. Moreover, the zero fringe positions x_ϕ have also been determined. For a given mass, the measurement has been repeated ten times without removing the load. Then, the average and the variance of measurements have been calculated. The statistical reproducibility is so given by the square root of the variance. The sensor sensitivity has been deduced from the slope of the curve x_ϕ versus the mass. All these measurements have been made with coated fibers.

4.3.2.1 Fiber SM800 (3M)

The fiber SM800 from 3M is a standard single-mode fiber with a cut-off wavelength [4.10] of approximately 800 nm. Thus we are sure that at 880 nm this fiber is monomode. Its diameter is 80 μm and so a sensitivity of approximately 4.6 pixels/kg can be expected.

The experimental results obtained with this fiber are shown in Fig. 4.16 where the center position x_{s1}, x_{s2} and the zero fringe position x_ϕ versus the applied mass are plotted. The observed deviation from the linear fit for masses lower than 2 kg is due to the intrinsic fiber birefringence. Indeed, single-mode optical fibers are birefringent as a result of elliptical core cross section and residual mechanical stress. This intrinsic fiber birefringence can be avoided by using low-birefringence fibers. A summary of the experimental results collected with this fiber is given in Table 4.2.

The measured sensor sensitivity is lower than the theoretical value given by Eqs. (4.42) and (4.43). Since the force is not directly applied to the fiber itself but to the coating, a reduced sensor sensitivity was expected. The reproducibility $\delta x_\phi = 1/570$ fringe is quite good, thanks to the high SNR of 44 dB. This value corresponds very well with theoretical calculations (see Fig. 4.6).

Signal-to-noise ratio:	SNR ≡ 44 dB
Sensor sensitivity:	$4\Phi/2\pi m \equiv -4.04$ pixels/kg
Reproducibility:	$\delta x_{s1} \equiv 0.053$ pixel (1/75 fringe)
	$\delta x_{s2} \equiv 0.045$ pixel (1/89 fringe)
	$\delta x_\phi \equiv 0.007$ pixel (1/570 fringe)

Table 4.2 Experimental results obtain with the fiber SM800 (3M).

We also investigated the stability of the sensor during a period of five hours. The fiber was loaded with a constant mass of 1 kg, applied at time t = 0. Four times per minute x_{s1}, x_{s2} and x_ϕ were measured. The results are shown in Fig. 4.17, as curves (a), (b) and (c), respectively.

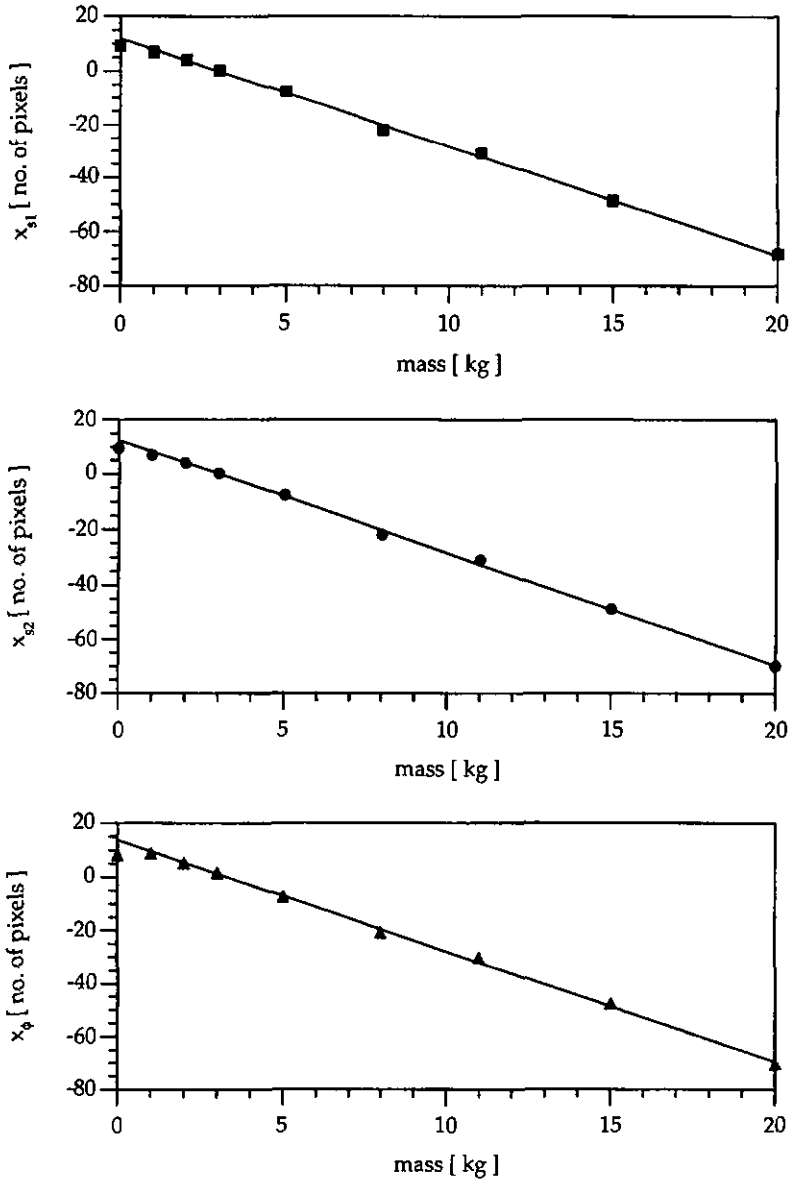


Fig. 4.16 Center position versus mass for fiber SM800 (3M): squares, center position x_{s1}; circles, center position x_{s2} (after one iteration); triangles, zero fringe position x_{ϕ} (phase).

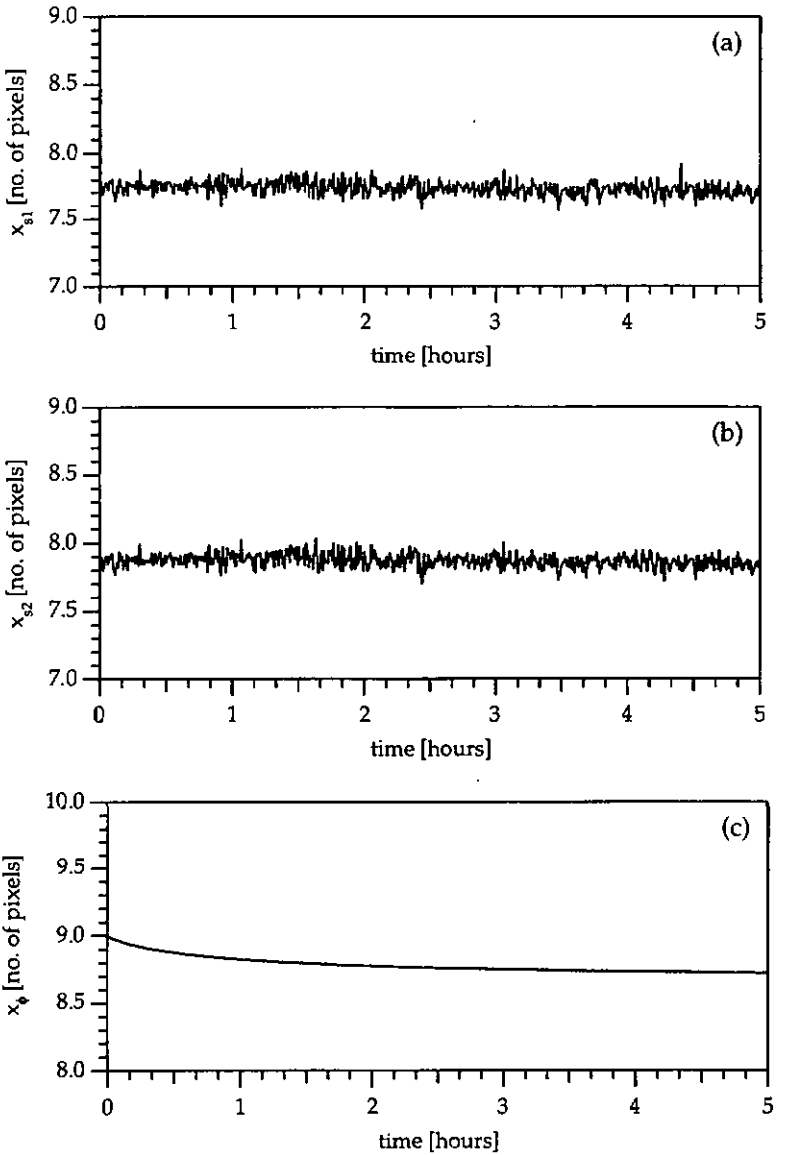


Fig. 4.17 Long-term stability for fiber SM800 (3M): center position x_{s1} [curve (a)]; center position x_{s2} (after one iteration) [curve (b)]; zero fringe position x_ϕ (phase) [curve (c)].

In Fig. 4.17, the difference between x_{s2} and x_ϕ is due to the nonuniformity of the light spot on the detector. This nonuniformity (see Fig. 4.14) is caused, first, by the fixed pattern noise which is due to the fact that alternate photodiodes are sampled on different clock phases, second, by the different sensitivities of the individual pixels generated by process variations and, third, by vignetting effects of the optics. This nonuniformity involves a distortion of the detected signal, which influences the center of gravity x_s but not the phase x_ϕ of the fringes. Therefore, it is not necessary to compensate for this nonuniformity, unless the resulting error for the center of gravity becomes larger than half a fringe period (2 pixels). We can see that x_{s1}, x_{s2} and x_ϕ undergo a drift of approximately –0.011 pixel/hour. The difference between x_ϕ at t = 0 and t = 300 minutes was about 1/16 of a fringe. This drift is due to the relaxation effects into the fiber coating. Figure 4.17 [curve (c)] shows also, that the determination of the zero fringe position x_ϕ from the detected white-light fringe signal is remarkably stable and accurate.

4.3.2.2 Fiber LB800 (York)

This fiber is a low-birefringence optical fiber with a core diameter of 125 μm and a cut-off wavelength of about 800 nm [4.14]. The expected sensor sensitivity is 3 pixels/kg. The results for x_{s1}, x_{s2} and x_ϕ versus mass are shown in Fig. 4.18 as curves (a), (b) and (c), respectively. The linearity is nearly perfect even for masses lower than 2 kg. A summary of the experimental results collected with this fiber are given in Table 4.3.

Signal-to-noise ratio:	SNR ≅ 37 dB
Sensor sensitivity:	$4\Phi/2\pi m \cong -3.06$ pixels/kg
Reproducibility:	$\delta x_{s1} \cong 0.17$ pixel (1/24 fringe)
	$\delta x_{s2} \cong 0.10$ pixel (1/40 fringe)
	$\delta x_\phi \cong 0.036$ pixel (1/110 fringe)

Table 4.3 Experimental results obtain with the fiber LB800 (York).

We used an optical setup slightly different than the one used previously. In fact, we employed three lenses instead of one to fill the whole aperture of the Wollaston prism. Thus, a lower optical power reached the detector and we measured a lower SNR of 37 dB.

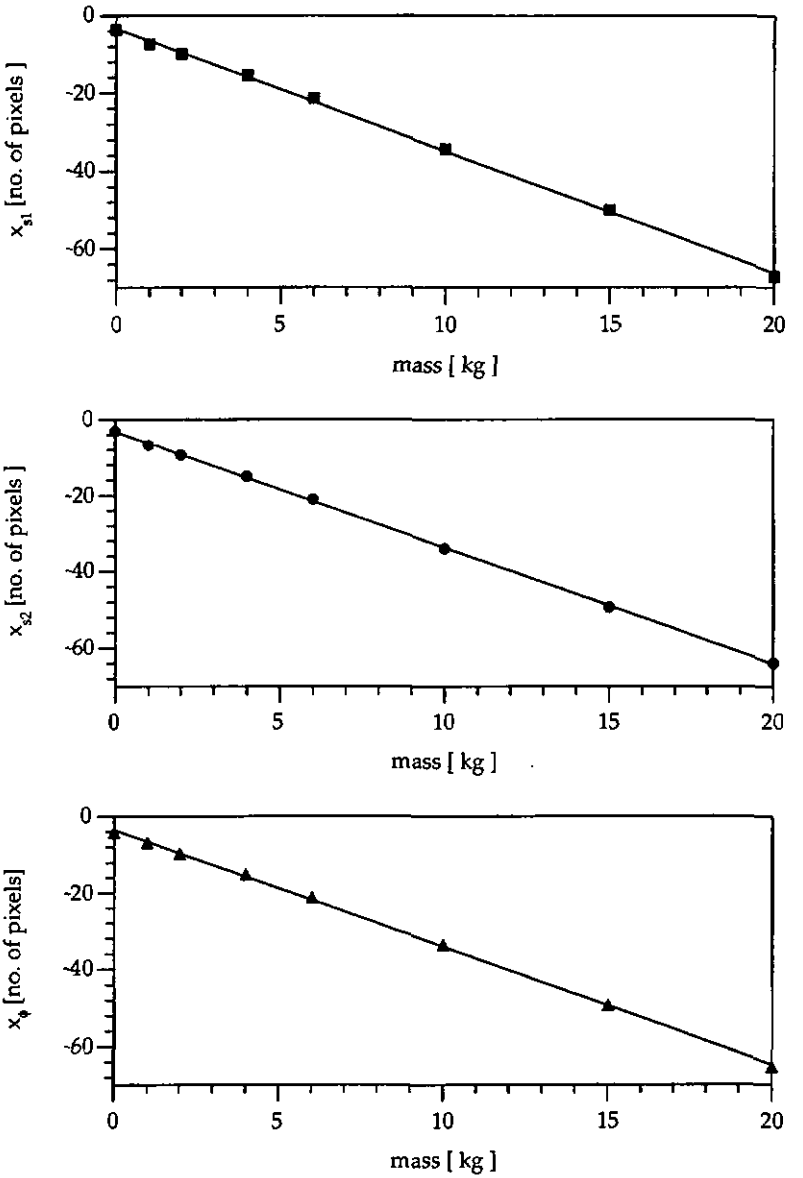


Fig. 4.18 Center position versus mass for fiberLB800 (York), 125 μm: squares, center of gravity x_{s1}; circles, center of gravity x_{s2} (after one iteration); triangles, zero fringe position x_{ϕ} (phase).

4.4 Conclusion

Through this chapter and with an outstanding signal processing, we have demonstrated that white-light interferometry is a powerful alternative to classical interferometry for numerous optical sensors. In fact, white-light interferometers exhibit a number of advantages over conventional systems, among others: a large unambigous operating range and, therefore, the possibility to make absolute measurements; the potentiality to identify the interference order when the interferometer is switched on and off; a high resolution and the possibility to employ inexpensive sources like LED or superluminescent diodes.

A signal processing method for white-light fringes, based on the calculation of the center of gravity of the fringe power (square modulus of the coherence function), has been presented. In addition, an optimum sampling concept has been developed, which is based on synchronous sampling with four samples per fringe period. This concept is particularly adequate for white-light fringes scanned by linear array detectors. The theoretical analysis and the experimental results show, that the proposed method is simple to operate, fast, accurate and extremely noise resistant.

We have described a simple polarizing interferometer for measuring absolute force or pressure, employing the induced birefringence in an optical monomode fiber. For this application, we introduce an novel solution for the receiving interferometer, which uses a Wollaston prism to compensate the phase retardation induced by the force. The interferogram generated at the output of the white-light interferometer is analyzed by a novel signal processing. We have demonstrated that the center position of the fringe signal can be located with a precision of about 1/500 fringe for a SNR of 44 dB. Moreover, the described signal processing could be utilized, with some minor changes, in any electronically scanned white-light interferometer.

The goal of the CERS research project with the company Hilti AG (Schaan, Liechtenstein) was the development of a sensor for surveying the force applied to bolts. For this application, the annular sensor shown in the top of Fig. 4.19 has an embedded optical fiber and remains permanently in the place of measurement, like a washer. The force acting on the sensor induces a birefringence in the monomode optical fiber which is measured with the polarizing interferometer

(see Fig. 4.13). The interferometer is mounted in the sensor head, as shown in the bottom of Fig. 4.19.

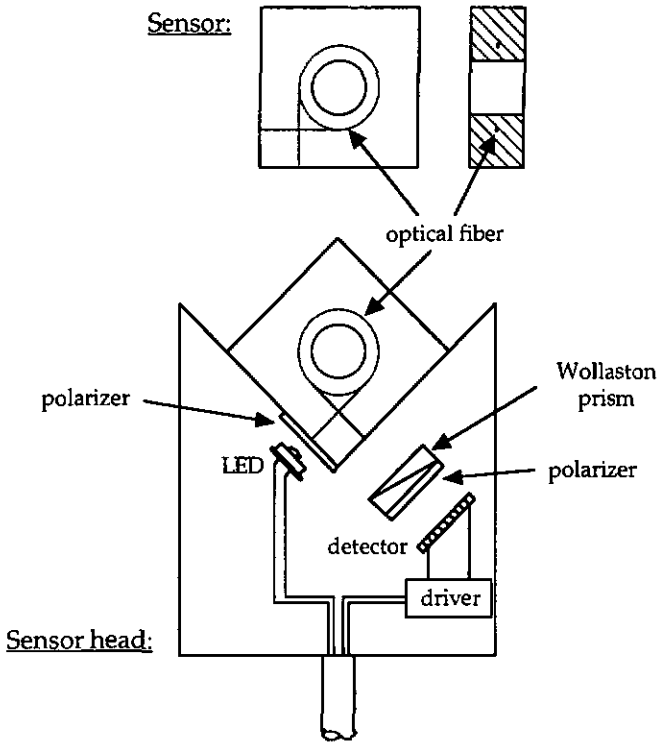


Fig. 4.19 Fiber optic force sensor.

The problem of embedding the fiber had not been solved satisfactorily. Hence, further investigations had to be carried out to increase the reproducibility of the measurements. The investigation were focused on two geometry of sensors: the sandwich geometry and the groove geometry. As material for embedding the fiber, epoxy with powder of quartz has been used. The principal problem of this process is the induced birefringence due to the local mechanical stress of the fiber. This stress causes an uncontrollable deviation of the expected linear behavior of the induced retardation in function of the applied force [see Eq. (4.42)]. The choice of the fiber (diameter, intrinsic birefringence) has, as well, an influence on the performance of the sensor.

Finally, an issue that needs to be discussed shortly is the dispersion of optical components. Dispersion refers here to the variation in the speed of light as a function of the wavelength (chromatic dispersion) [4.15]. In contrast to inteferometry with highly coherent lasers, dispersion has to be considered in white-light interferometry, particularly when optical waveguides and fibers are involved [4.16].

The practical implications of dispersion can bcome quite significant. For example, de Groot had shown theoretically and experimentally, that neglecting chromatic dispersion when the light propagates through fused silica can lead to an error on the measured distance of approximately 14 μm/mm of glass thickness [4.16].

In the theoretical analysis developed in this thesis we made the assumption that the maximum of the coherence envelope $F(x-x_s)$ has the same position x_s as the phase of the zero-order fringe $\Phi = -(2\pi x_s)/\Lambda$ [see Eq. (4.1)]. This is true only for strictly nondispersive sensing and receiving interferometers. For dispersive sensing or receiving interferometers, the measurand-induced displacements of the coherence envelope and the interference fringe are no longer identical. Urbanczyk and al. showed that the position of the coherence function and the zero-order fringe are determined, respectively, by the group and phase imbalanced optical path difference [4.17]. Moreover, the fringe envelope may present some asymmetry [4.5].

The signal processing presented in this chapter calculates first the center of gravity of the power of the interference signal to better than half a fringe period (2 pixels). Then, the phase of the fringe signal is determined to get the exact center position of the white-light fringes. Therefore, if the dispersion of optical components induces an error on the position of the fringe envelope greater than 2 pixels, the phase of the zero-order fringe calculated in the second step will be wrong (by an integer multiple of half a fringe period). Thus, dispersion effects should be carefully analyzed for white-light interferometric sensors before applying any kind of high resolution signal processing.

References

Chapter 1

[1.1] R. J. Pryputniewicz, G. M. Brown and W. P. O. Jüptner, eds., *Laser Interferometry VIII: Applications*, Proc. SPIE **2861** (1996).

[1.2] J.-F. Willemin, "Interférométrie hétérodyne à speckles: Application à la mesure de vibrations mécaniques microscopiques", Ph.D.thesis, University of Neuchâtel, Switzerland (1983).

[1.3] J.-F. Willemin, R. Dändliker and S. M. Khanna, "Heterodyne interferometer for submicroscopic vibration measurements in the inner ear", J. Accoust. Soc. Am. **82**, 787-795 (1988).

[1.4] S. M. Khanna, "Cellular vibration and mobility in the organ of Corti: an introduction and summary", Acta Otolaryngol, Suppl. **467**, 7-18 (1989).

[1.5] R. Dändliker, W. Odoni, H.-P. Schad and E. Schweizer, "Verfahren und Einrichtung zur faseroptischen Kraftmessung", Patent application DE 4037077, CH, D, UK, FL, USA, Application date: 7.11.91.

[1.6] R. Dändliker, W. Odoni, H.-P. Schad and E. Schweizer, "Verfahren zur faseroptischen Kraftmessung", Patent application DE 4202185, Application date: 18.1.93.

Chapter 2

[2.1] H. J. Tiziani, "Optical methods for precision measurements", Optical and Quantum Elect. **21**, 253-282 (1989).

[2.2] P. Hariharan and B. J. Thompson, eds., *Selected Papers on Interferometry*, SPIE milestone series MS **28** (1991).

[2.3] P. Hariharan and D. Malacara, eds., *Selected Papers on Interference, Interferometry, and Interferometric Metrology*, SPIE milestone series **MS 110** (1995).

[2.4] A. D. Kersey, "Recent progress in interferometric fiber sensor technology", in *Fiber Optic and Laser Sensors VIII*, R. P. DePaula and E. Udd, eds., Proc. SPIE **1367**, 2-12 (1990).

[2.5] A. J. den Boef, "Applications of laser diodes in interferometry", in *Industrial Inspection*, D. W. Braggins, ed., Proc. SPIE **1010**, 10-17 (1988).

[2.6] A. Abou-Zeid, "Diode lasers for interferometry", Precision Eng. **11**, 139-144 (1989)

[2.7] T. Spirig, P. Seitz, O. Vietze and F. Heiger, "The lock-in CCD: Two-dimensional synchronous detection of light", IEEE J. of Quantum Elect. **31**, 1705-1708 (1995).

[2.8] R. Dändliker, E. Zimmermann and Y. Salvadé, "Absolute distance measurement", in *Optics for Science and New Technology*, J.-S. Chang, J.-H. Lee, S.-Y. Lee and C. H. Nam, eds., Proc. SPIE **2778**, 1074-1075 (1996).

[2.9] A. Lassila, E. Ikonen and K. Riski, "Interferometers for calibration of length standards", Opt. Eng. **34**, 2619-2623 (1995).

[2.10] C. C. Williams and H.K. Wickramasinghe, "Absolute optical ranging with 200 nm resolution", Opt. Lett. **14**, 542-544 (1989).

[2.11] A. Koch and R. Ulrich, "Fiber-optic displacement sensor with 0.02 μm resolution by white-light interferometry", Sensors and Actuators A **25-27**, 201-207 (1991).

[2.12] R. Dändliker, K. Hug, J. Politch and E. Zimmermann, "High-accuracy distance measurements with multiple-wavelength interferometry", Opt. Eng. **34**, 2407-2412 (1995).

[2.13] U. Schnell, E. Zimmermann and R. Dändliker, "Absolute distance measurement with synchronously sampled white-light channeled spectrum interferometry", Pure and Appl. Opt. **4**, 643-651 (1995).

[2.14] E. P. Tomasini, ed., *Vibration Measurements by Laser Techniques: Advances and Applications*, Proc. SPIE **2868** (1996).

[2.15] J. C. Wyant, "How to extend interferometry for rough-surface tests", in *Laser Focus World* , pp. 131-135 (Sept. 1993).

[2.16] B. S. Lee and T. C. Strand, "Profilometry with a coherence scanning microscope", Appl. Opt. **29**, 3784-3788 (1990).

[2.17] P. de Groot, "Interferometric laser profilometer for rough surfaces", Opt. Lett. **16**, 357-359 (1991).

[2.18] S. Chen, A. J. Rogers and B. T. Meggitt, "Electronically scanned optical-fiber Young's white-light interferometer", Opt. Lett. **16**, 761-763 (1991).

[2.19] M. N. Charasse, M. Turpin and J.-P. Le Pesant, "Dynamic pressure sensing with a side-hole birefringent optical fiber", Opt. Lett. **16**, 1043-1045 (1991).

[2.20] R. J. Adrian, ed., *Selected Papers On Laser Doppler Velocimetry*, SPIE milestone series **MS 78** (1993).

[2.21] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics* (John Wiley & Sons, New York, 1991), chap. 5.

[2.22] B. E. A. Saleh and M. C. Teich, op. cit., chap. 2.

[2.23] B. E. A. Saleh and M. C. Teich, op. cit., chap. 10.

[2.24] H. Marguerre, "Optical phase-sensitive detection", in *Sensors vol. 6: Optical Sensors*, E. Wagner, R. Dändliker and K. Spenner, eds. (VCH, Weinheim, 1992), pp. 307-332.

[2.25] J. C. Wyant, "Testing aspherics using two-wavelength holography", App. Opt. 10, 2113-2118 (1971).

[2.26] C. Polhemus, "Two-wavelength interferometry", Appl. Opt. **12**, 2071-2074 (1973).

[2.27] K. Creath and J. C. Wyant, in *Optical Shop Testing 2nd Ed.*, D. Malacara, ed. (John Wiley, New York, 1978), pp. 612-617.

[2.28] A. F. Fercher, H. Z. Hu and U. Vry, "Rough surface interferometry with a two-wavelength heterodyne speckles interferometer", Appl. Opt. **24**, 2181-2188 (1985).

[2.29] P. de Groot, "Three-color laser-diode interferometer", Appl. Opt. **30**, 3612-3616 (1991).

[2.30] L. Deck and F. Demarest, "Two-color light-emitting-diode source for high precision phase-shifting interferometry", Opt. Lett. **18**, 1899-1901 (1993).

[2.31] R. Dändliker, R. Thalmann and D. Prongué, "Two-wavelength laser interferometry using superheterodyne detection", Opt. Lett. **13**, 339-341 (1988).

[2.32] E. Gelmini, U. Minoni and F. Docchio, "Tunable, double-wavelength heterodyne detection interferometer for absolute-distance measurements" Opt. Lett. **19**, 213-215 (1994).

[2.33] E. Zimmermann, Y. Salvadé and R. Dändliker, "Stabilized three-wavelength source calibrated by electronic means for high accuracy distance measurement", Opt. Lett. **21**, 531-533 (1996).

[2.34] P.A. Flournoy, R. W. McClure and G. Wyntjes, "White-light interferometric thickness gauge", Appl. Opt. **11**, 1907-1915 (1972).

[2.35] K. T. V Grattan, A. W. Palmer and Y. N. Ning, "The use of low-coherence light sources in fiber-optic interferometric systems", in *Sensors vol. 6: Optical Sensors*, E. Wagner, R. Dändliker and K. Spenner, eds. (VCH, Weinheim, 1992), pp. 529-550.

[2.36] T. Dresel, G. Häusler and H. Venzke, "Three-dimensional sensing of rough surfaces by coherence radar", Appl. Opt. **31**, 919-925 (1992).

[2.37] J. Schwider and L. Zhou, "Dispersive interferometric profilometer", Opt. Lett. **19**, 695-697 (1994).

[2.38] R. Dändliker, E. Zimmermann and G. Frosio, "Electronically scanned white-light interferometry: a novel noise-resistant signal processing", Opt. Lett. **17**, 679-681 (1992).

[2.39] P. J. Caber, "Interferometric profiler for rough surfaces", Appl. Opt. 32, 3438-3441 (1993).

Chapter 3

[3.1] E. Zimmermann and R. Dändliker, "Full-dynamic phase-demodulator for heterodyne interferometric vibration measurements", *Proceedings of International Workshop on Interferometry* (International Commission for optics and Optical Society of Japan, Waco, 1996), pp. 79-80.

[3.2] E. Zimmermann and R. Dändliker, "Full-dynamic phase-demodulator for heterodyne interferometric vibration measurements", in *Vibration Measurements by Laser Techniques: Advances and Applications*, E. P. Tomasini, ed., Proc. SPIE 2868, 488-499 (1996).

[3.3] H. Marguerre, "Optical phase-sensitive detection", in *Sensors vol. 6: Optical Sensors*, E. Wagner, R. Dändliker and K. Spenner, eds. (VCH, Weinheim, 1992), pp. 307-332.

[3.4] K. Creath, "Temporal phase measurement methods", in *Interferogram Analysis: Digital Fringe Pattern Measurement Technique*, D. W. Robinson, G. T. Reid, eds. (Institute of Physics Publishing, Philadelphia, 1993), chap. 4.

[3.5] T. Spirig, P. Seitz, O. Vietze and F. Heiger, "The lock-in CCD: Two-dimensional synchronous detection of light" IEEE J. of Quantum Elect. 31, 1705-1708 (1995).

[3.6] R. Dändliker, E. Zimmermann and Y. Salvadé, "Absolute distance measurement", in *Optics for Science and New Technology*, J.-S. Chang, J.-H. Lee, S.-Y. Lee and C. H. Nam, eds., Proc. SPIE 2778, 1074-1075 (1996).

[3.7] J.-F. Willemin, R. Dändliker and S. M. Khanna, "Heterodyne interferometer for submicroscopic vibration measurements in the inner ear", J. Accoust. Soc. Am. 82, 787-795 (1988).

[3.8] J.-F. Willemin, S. M. Khanna and R. Dändliker, "Heterodyne interferometer for cellular vibration measurement", Acta Otolaryngol, Suppl. 467, 35-42 (1989).

[3.9] S. M. Khanna, J.-F. Willemin and M. Ulfendahl "Measurement of optical reflectivity in cells of the inner ear", Acta Otolaryngol, Suppl. 467, 69-75 (1989).

[3.10] S. Haykin, *Communication Systems* (John Wiley & Sons, New York, 1978), pp. 396-407.

[3.11] S. Haykin, Op. cit., pp. 325-328.

[3.12] A. Yariv, *Optical Electronics 3rd ed.* (Holt-Saunders, New-York, 1985), pp. 306-320.
[3.13] S. Haykin, Op. cit., pp. 445-454.
[3.14] M. Schwartz, W. R. Bennet and S. Stein, *Communication Systems and Techniques* (McGraw-Hill, New York, 1966), pp. 248-255.
[3.15] P.-G. Fontalliet, Systèmes de Télécommunication (Presses Polytechniques Romandes, Lausanne, 1983), pp. 296-299.
[3.16] J. Blot, *Electronique Linéaire* (Dunod Université, Paris, 1993), pp. 211-212.
[3.17] P. Hariharan, "Interferometric metrology: current trends and future prospects", in *Interferometric Metrology*, N. A. Massie, ed., Proc. SPIE **816**, 2-18 (1987).
[3.18] M. Girard, *Amplificateur de Puissance* (McGraw Hill , Paris, 1988). pp. 22-23.
[3.19] K. W. Cattermole, *Principles of Pulse Code Modulation* (Iliffe Books, London, 1969), chap. 3.
[3.20] D. W. Allan, "Statistics of atomic frequency standards", Proc. IEEE **54**, 221-230 (1966).
[3.21] Spindler & Hoyer GmbH, *Catalog on Optics and Mechanics*, Q2-Q3 (1994).
[3.22] S. Haykin, Op. cit., pp. 321-323.
[3.23] S. Haykin, Op. cit., pp. 333-334.

Chapter 4

[4.1] R. Dändliker, E. Zimmermann and G. Frosio, "Noise-resistant signal processing for electronically scanned white-light interferometry", *Proceeding of Eighth Optical Fiber Sensors Conference* (Optical Society of America, Monterey, 1992) pp. 53-56.
[4.2] R. Dändliker, E. Zimmermann and G. Frosio, "Electronically scanned white-light interferometry: a novel noise-resistant signal processing", Opt. Lett. **17**, 679-681 (1992).
[4.3] R. Dändliker, K. Hug, E. Zimmermann and U. Schnell, "Multiple-wavelength and white-light interferometry", *Proceedings of Second Japan-France Congress on Mechatronics* (Japan Society of Precision Engineering, Takamatsu, 1994), pp. 305-309.
[4.4] K. T. V Grattan, A. W. Palmer and Y. N. Ning, "The use of low-coherence light sources in fiber-optic interferometric systems", in *Sensors vol. 6: Optical Sensors*, E. Wagner, R. Dändliker and K. Spenner, eds. (VCH, Weinheim, 1992), pp. 529-550.

[4.5] T. Dresel, G. Häusler and H. Venzke, "Three-dimensional sensing of rough surfaces by coherence radar", Appl. Opt. 31, 919-925 (1992).

[4.6] P. J. Caber, "Interferometric profiler for rough surfaces", Appl. Opt. 32, 3438-3441 (1993).

[4.7] P. de Groot and L. Deck, "Surface profiling by analysis of white-light interferograms in the spatial frequency domain", J. of Mod. Opt. 42, 388-391 (1995)

[4.8] W. J. Bock, W. Urbanczyk and M. B. Zaremba, "Electronically scanned white-light interferometric strain sensor employing highly birefringent fibers", Opt. Commun. 101, 157-162 (1993).

[4.9] M. N. Charasse, M. Turpin and J.-P. Le Pesant, "Dynamic pressure sensing with a side-hole birefringent optical fiber", Opt. Lett. 16, 1043-1045 (1991).

[4.10] P. R. Bevington, *Data Reduction and Analysis for the Physical Sciences* (McGraw-Hill Inc., New-York, 1969), chap. 4.

[4.11] Y. Mamihia, K. Kudo and Y. Mushiaka, "Effect of mechanical stress on the transmission characteristics of optical fibers", Trans. IECE Jpn. 60, 107-113 (1977).

[4.12] A. Bertholds and R. Dändliker, "High-resolution photoelastic pressure sensor using low-birefringence fiber", Appl. Opt. 25, 340-343 (1986).

[4.13] A. Bertholds and R. Dändliker, "Determination of the individual strain-optic coefficients in single-mode optical fibers", J. of Ligthwave Technol. 6, 17-20 (1988).

[4.14] A. W. Snyder, "Understand monomode optical fiber", Proc. IEEE 69, 6-13 (1981).

[4.15] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics* (John Wiley & Sons, New York, 1991), chaps. 5and 22.

[4.16] P. de Groot, "Chromatic dispersion effects in coherent absolute ranging", Opt. Lett. 17, 898-900 (1992).

[4.17] W. Urbanczyk and W. J. Bock, "Analysis of disperson effects for white-light interferometric fiber-optic sensors", Appl. Opt. 33, 124-129 (1994).

Annexes

A. Phase demodulator: electronic circuits

In this Annex, the detailed circuit diagrams of the full-dynamic phase demodulator (FDPD) are given. These circuits was originally designed by J.-F. Willemin.

The concept of the FDPD is schematically shown in Fig. A1. The analog version consists of limiter amplifiers, mixers and active low-pass filters to eliminate the carrier frequency after mixing. In the digital version of the FDPD, the low-pass filter is replaced by a pulse-width to voltage converter followed by a synchronized A/D converter working at one half of the carrier frequency.

- Input sensitivity: 5 mV
- Dynamic range: more than 80 dB
- Frequency response: dc – 6.9 kHz (analog FDPD)
  dc – 113.5 kHz (digital FDPD)

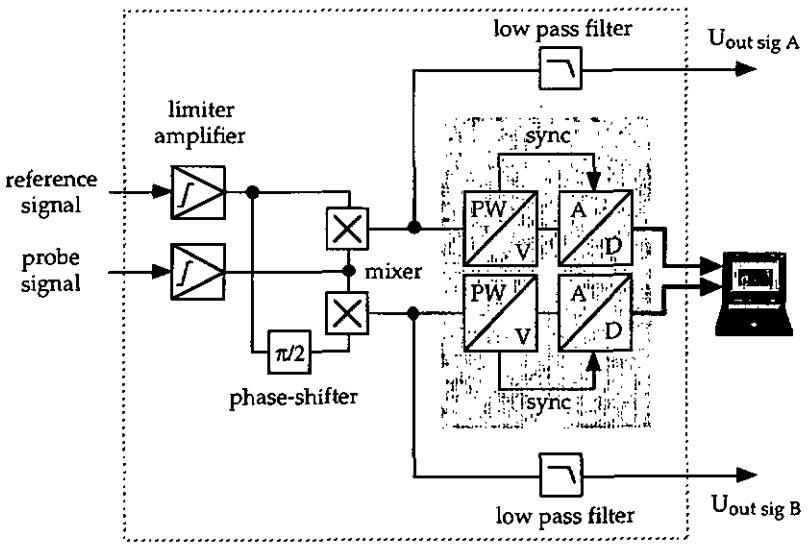


Fig. A1 Full-dynamic phase-demodulator.

Practically, the FDPD was build on three printed Eurocards mounted in a 19" rack.

Card #1: Analog FDPD, channel A

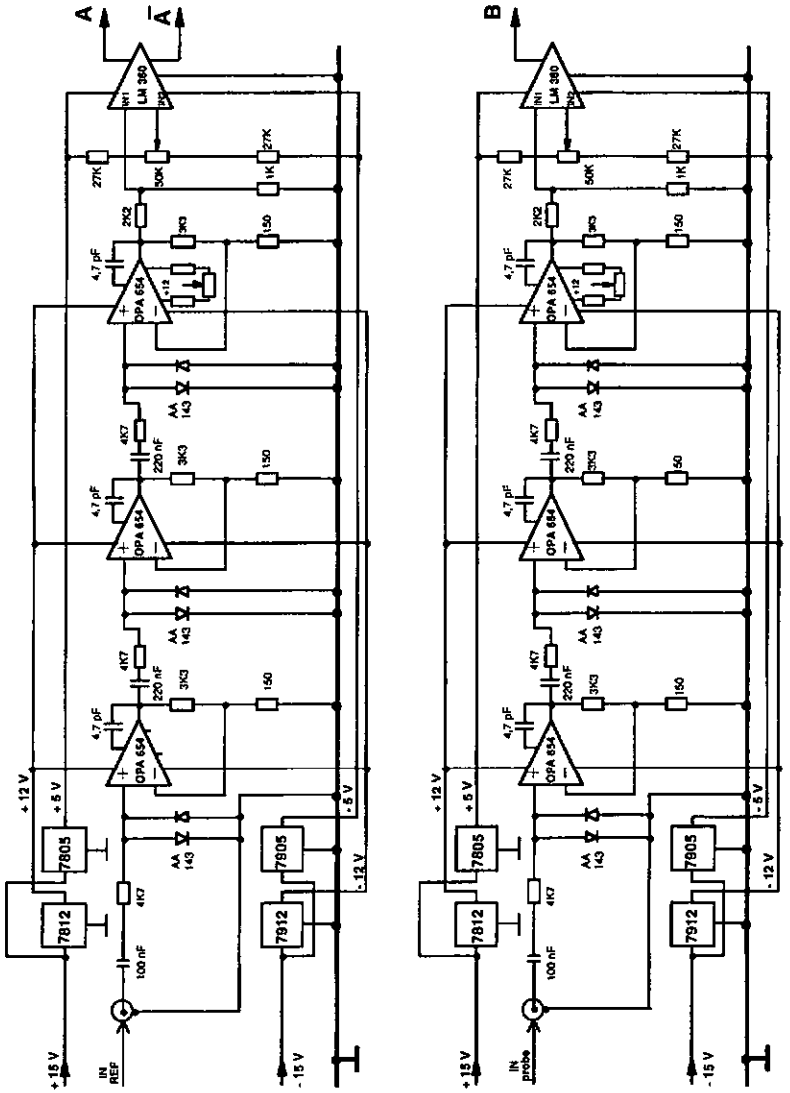
- Input stage for the reference signal, (see p. 111).
- Phase-shifter module, mixer and active low-pass filter (see p. 112).
- Pulse-width to voltage converter and command circuit for the A/D converter (see p. 112).

Card #2: Analog FDPD, channel B

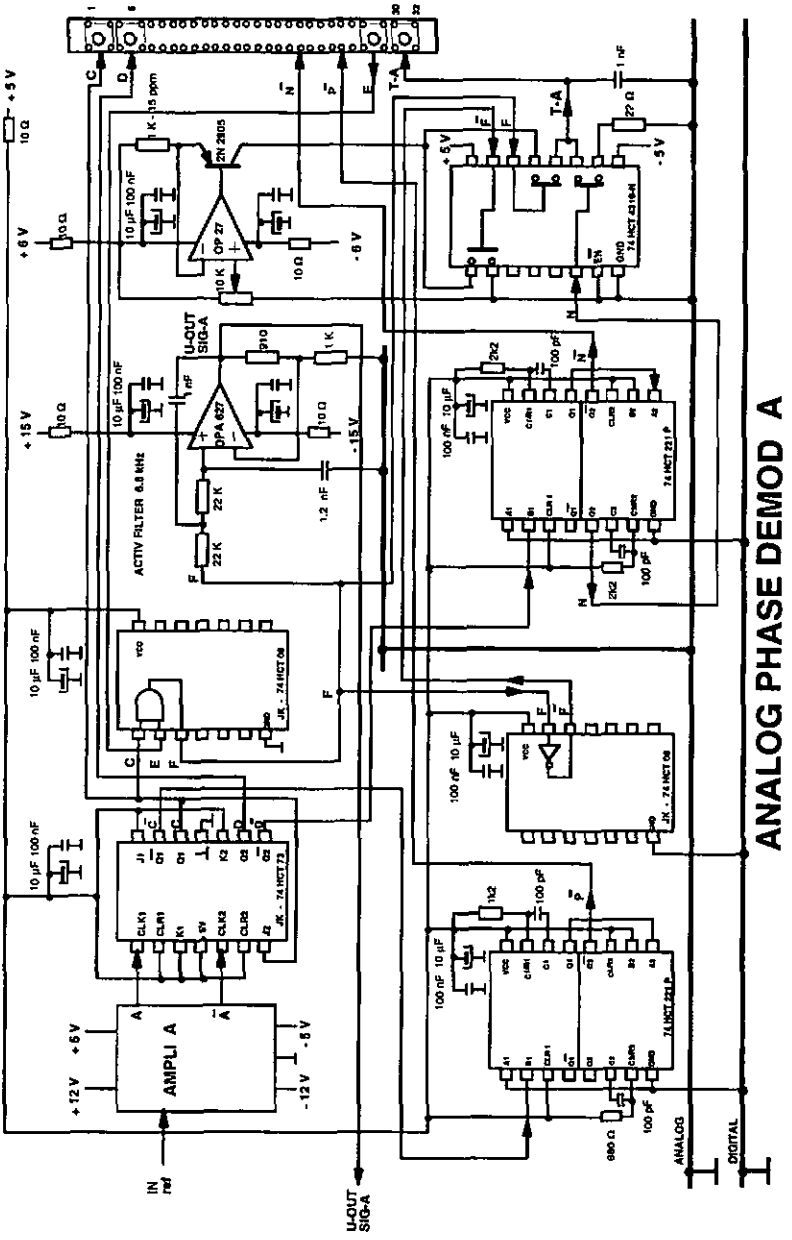
- Input stage for the probe signal (see p. 111).
- Phase-shifter module, mixer and Active low-pass filter (see p. 113).
- Pulse-width to voltage converter and command circuit for the A/D converter (see p. 113).

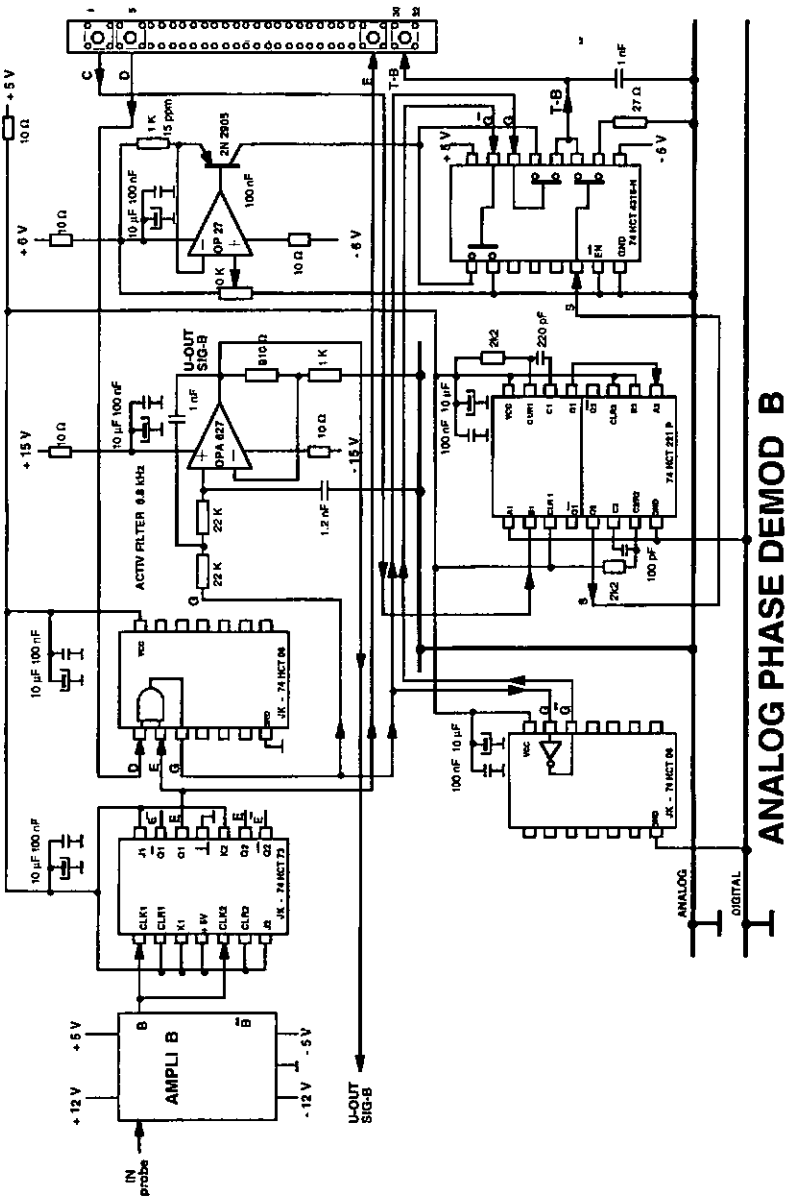
Card #3: Digital FDPD, channel A and B

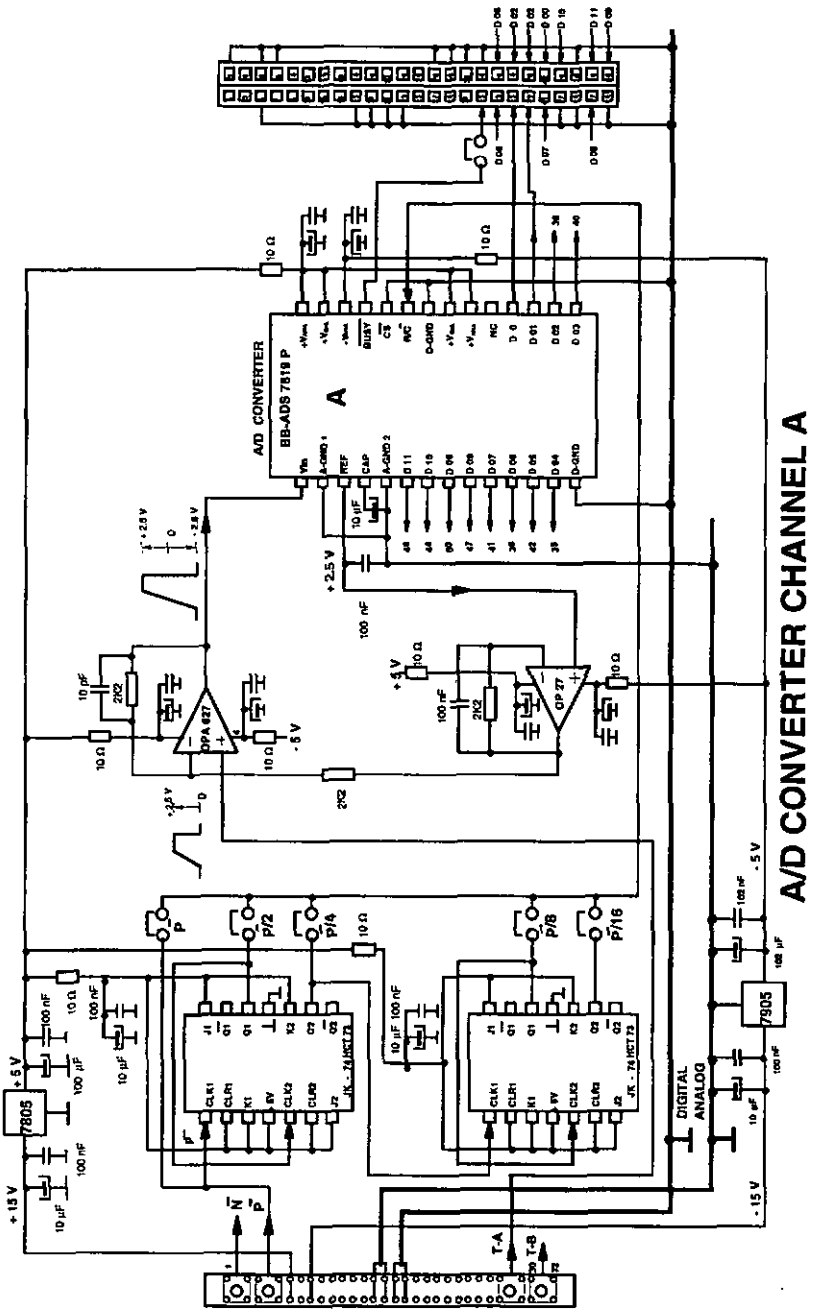
- Synchronized A/D converter, channel A (see p. 114).
- Synchronized A/D converter, channel B (see p. 115).
- Computer interface (see pp. 114 and 115).

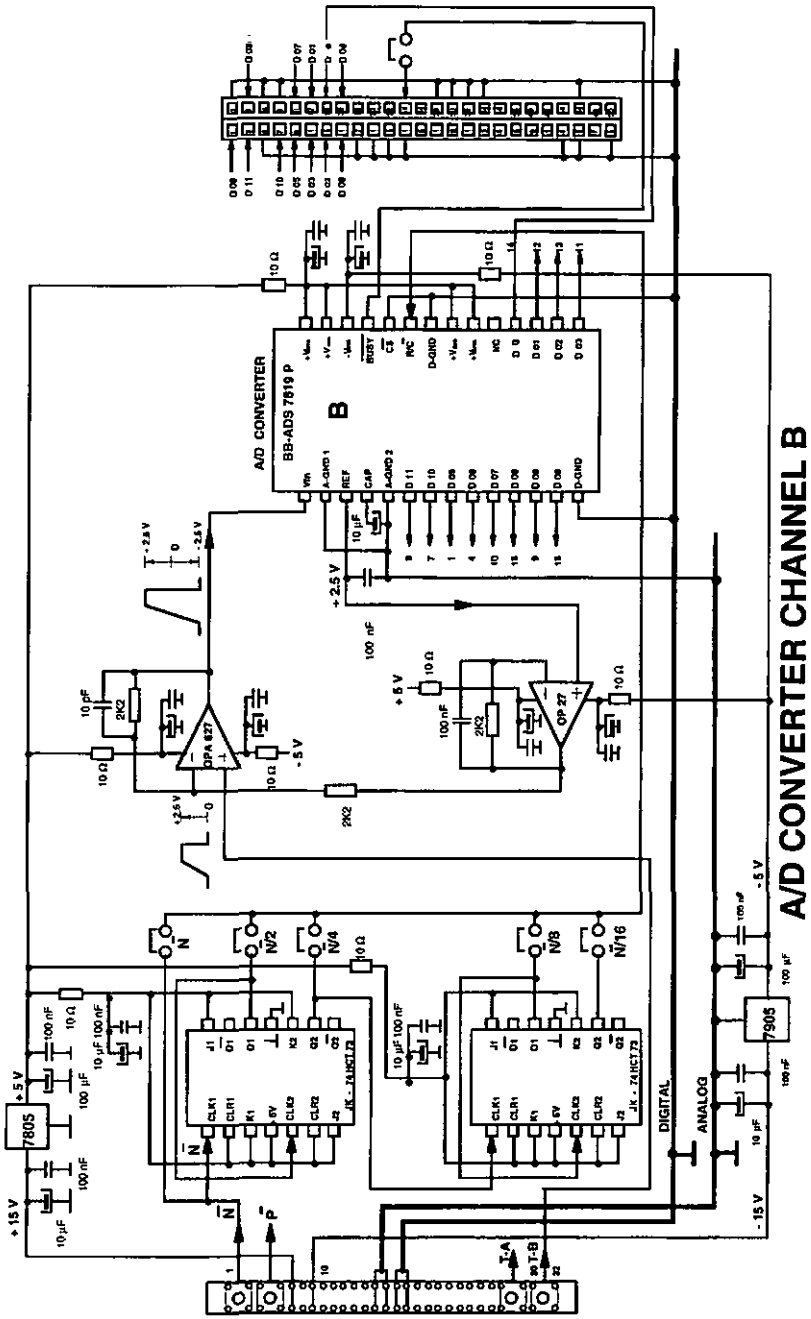


LIMITER AMPLIFIER


ANALOG PHASE DEMOD A
ANALOG
DIGITAL
AMPLI A
ACTIV FILTER 6.8 MHz
U-OUT SIG-A
IN ref
OP 27
OPA 627
2N 2905
74 HCT 221 P
74 HCT 4316 M
74 HCT 73
74 HCT 08
JK - 74 HCT 08
JK - 74 HCT 73
+5 V
-5 V
+6 V
-6 V
+15 V
-15 V
+12 V
-12 V
10 Ω
1 K - 15 ppm
10 µF 100 nF
10 K
22 K
1.2 nF
1 nF
100 pF
100 nF 10 µF
2k2
1k2
680 Ω


ANALOG PHASE DEMOD B
ANALOG
DIGITAL
AMPLI B
IN probe
U-OUT SIG-B
ACTIV FILTER 8.8 kHz
OPA 627
OP 27
2N 2905
74 HCT 221 P
74 HCT 4316 N
74 HCT 86
74 HCT 73
+ 5 V
+ 6 V
+ 15 V
+ 12 V
- 5 V
- 6 V
- 15 V
- 12 V


A/D CONVERTER CHANNEL A
A/D CONVERTER
BB-ADS 7819 P
A
OPA 627
OP 27
7805
7905
+ 15 V
+ 5 V
- 5 V
- 15 V
+ 2,5 V
DIGITAL
ANALOG
T-A
T-B


A/D CONVERTER CHANNEL B
A/D CONVERTER
BB-ADS 7819 P
B
OPA 627
OP 27
7805
7905
DIGITAL
ANALOG
+ 5 V
- 5 V
+ 15 V
- 15 V
+ 2,5 V
- 2,5 V
10 Ω
100 nF
10 µF
100 µF
2K2
10 pF
N
N/2
N/4
N/8
N/16
JK - 74HCT73

B. Signal processing for phase recovery

The phase information, that is the displacement of the vibrating object, is recovered off-line from the two output signals in quadrature of the phase-demodulator. We utilized the software Igor Pro 3.0 from WaveMetrics to implement the functions described in this Annex.

```
//-------------------------------------------------------------------------------------------
//
//    SIGNAL PROCESSING FOR PHASE RECOVERY
//
//    by Eric Zimmermann, vers. 3.1, ©IMT, September 1996
//
//-------------------------------------------------------------------------------------------

Function Data_Init(maxA, minA, maxB, minB)                    // initialization
      Variable      maxA, minA, maxB, minB

      Variable/G    max_A=maxA, min_A=minA                    // signal A
      Variable/G    mean_A=minA+(maxA-minA)/2
      Variable/G    threshold_A=(maxA-minA)/4
      Variable/G    slope_A=(maxA-minA)/(2*pi)

      Variable/G    max_B=maxB, min_B=minB                    // signal B
      Variable/G    mean_B=minB+(maxB-minB)/2
      Variable/G    threshold_B=(maxB-minB)/4
      Variable/G    slope_B=(maxB-minB)/(2*pi)
End

Function Phi_SigA(voltage)                                    // phase vs output voltage:
      Variable voltage                                        // signal A

      NVAR max=max_A, slope=slope_A

      return -(voltage-max)/slope
End

Function Phi_SigB(voltage)                                    // phase vs output voltage:
      Variable voltage                                        // signal B

      NVAR mean=mean_B, slope=slope_B

      return (voltage-mean)/slope
End
```

```
Function Phase_Recovery(WaveNameA, WaveNameB)
	String WaveNameA, WaveNameB
	String WaveNamePhase= WaveNameA+"_PhaseRec"
	Duplicate/O $WaveNameA, $WaveNamePhase
	Variable pwave=0, period=1, n_old=1, n=1
	Variable lengthwave=numpnts($WaveNameA)
	Wave Awave=$WaveNameA, Bwave=$WaveNameB
	Wave Phasewave=$WaveNamePhase
	NVAR meanA=mean_A, thresholdA=threshold_A
	NVAR meanB=mean_B, thresholdB=threshold_B
	do                                                    // determine the quadrant
	if ((Awave[pwave]>meanA) %& (Bwave[pwave]>=meanB))
		n=1
	else
		if ((Awave[pwave]<=meanA) %& (Bwave[pwave]>meanB))
			n=2
		else
			if ((Awave[pwave]<meanA) %& (Bwave[pwave]<=meanB))
				n=3
			else
				if ((Awave[pwave]>=meanA) %& (Bwave[pwave]<meanB))
					n=4
				endif
			endif
		endif
	endif
	if ((n_old==4) %& (n==1))                             // compute the period
		period+=1
	endif
	if ((n_old==1) %& (n==4))
		period-=1
	endif
	n_old=n
	do                                                    // case structure
		if (n==1)                                         // first quadrant
			if (Bwave[pwave] <= meanB+thresholdB)
				Phasewave[pwave] =Phi_SigB(Bwave[pwave])
			else
				Phasewave[pwave] =Phi_SigA(Awave[pwave])
			endif
			break
		endif
		if (n==2)                                         // second quadrant
			if (Awave[pwave] >= meanA-thresholdA)
				Phasewave[pwave] =Phi_SigA(Awave[pwave])
			else
```

```
                    Phasewave[pwave] = 2*pi-Phi_SigB(Bwave[pwave])
                endif
                break
            endif
            if (n==3)                                   // third quadrant
                if (Bwave[pwave] >= meanB-thresholdB)
                    Phasewave[pwave] = 2*pi-Phi_SigB(Bwave[pwave])
                else
                    Phasewave[pwave] =4*pi-Phi_SigA(Awave[pwave])
                endif
            break
            endif
            if (n==4)                                   // fourth quadrant
                if (Awave[pwave] <= meanA+thresholdA)
                    Phasewave[pwave] =4*pi-Phi_SigA(Awave[pwave])
                else
                    Phasewave[pwave] =4*pi+Phi_SigB(Bwave[pwave])
                endif
                break
            endif
        while(0)
        Phasewave[pwave] = Phasewave[pwave]+4*pi*(period-1)
        pwave+=1
        while(pwave<=lengthwave)
End
```

C. Statistical properties of sampled Gaussian noise

The probability distribution p(g) of Gaussian noise is given by

$$p(g) = \frac{1}{\sqrt{2\pi}\cdot\sigma}\exp\left[-\frac{g^2}{2\sigma^2}\right]. \tag{C.1}$$

Using Eq. (C.1), the average <g> of the noise is equal to

$$\langle g\rangle = \int_{-\infty}^{\infty} g\cdot p(g)dg = 0. \tag{C.2}$$

For the variance and for the moments of order 3 and 4, we find

$$\left\langle \Delta g^2\right\rangle = \left\langle g^2\right\rangle = \int_{-\infty}^{\infty} g^2\cdot p(g)dg = \sigma^2, \tag{C.3}$$

$$\left\langle g^3\right\rangle = \int_{-\infty}^{\infty} g^3\cdot p(g)dg = 0, \tag{C.4}$$

and

$$\left\langle g^4\right\rangle = \int_{-\infty}^{\infty} g^4\cdot p(g)dg = 3\sigma^4. \tag{C.5}$$

Now, we consider Gaussian noise $g(x,\sigma)$ which is uniformly sampled at the positions x_n, where n is an integer with $-N/2 \le n \le N/2$. We define the average noise power as

$$A_g = \frac{1}{N+1}\sum_n g^2(x_n), \tag{C.6}$$

and the first moment of the noise power by

$$B_g = \frac{1}{N+1}\sum_n n g^2(x_n). \tag{C.7}$$

The center position (center of gravity) of the Gaussian noise is then equal to [see Eq. (4.4)]

$$x_g = \frac{B_g}{A_g}. \tag{C.8}$$

From the definition (C.6), the mean of A_g can calculated by

$$\langle A_g \rangle = \frac{1}{N+1}\left\langle \sum_{-N/2}^{N/2} g(x_n)^2 \right\rangle = \frac{1}{N+1}\sum_{-N/2}^{N/2} \langle g(x_n)^2 \rangle. \tag{C.9}$$

Substituting Eq. (C.3) into Eq. (C.9), we obtain

$$\langle A_g \rangle = \sigma^2. \tag{C.10}$$

The variance of the average noise power is given by

$$\langle \Delta A_g^2 \rangle = \langle A_g^2 \rangle - \langle A_g \rangle^2, \tag{C.11}$$

where $\langle A_g^2 \rangle$ is obtain from

$$\langle A_g^2 \rangle = \left\langle \left[\frac{1}{N+1} \sum_{-N/2}^{N/2} g(x_n)^2 \right]^2 \right\rangle. \tag{C.12}$$

Explicitely, Eq. (C.12) can be developed as

$$\langle A_g^2 \rangle = \frac{1}{(N+1)^2}\left\langle \left(\sum_{n=-N/2}^{N/2} g(x_n)^2 \right) \cdot \left(\sum_{m=-N/2}^{N/2} g(x_m)^2 \right) \right\rangle, \tag{C.13}$$

then

$$\left\langle A_g^2 \right\rangle = \frac{1}{(N+1)^2} \sum_{n=-N/2}^{N/2} \left\langle g(x_n)^4 \right\rangle + \frac{1}{(N+1)^2} \sum_{n \neq m} \left\langle g(x_n)^2 \right\rangle \left\langle g(x_m)^2 \right\rangle. \qquad (C.14)$$

Substituting Eqs. (C.3) and (C.5) into Eq. (C.14), we obtain

$$\left\langle A_g^2 \right\rangle = \frac{N+3}{N+1}\sigma^4. \qquad (C.15)$$

From Eqs. (C.10) and (C.15), the variance of the average noise power is finally given by

$$\left\langle \Delta A_g^2 \right\rangle = \frac{2\sigma^4}{N+1}. \qquad (C.16)$$

Using the definitions (C.7) and (C.8), we can calculate the average and the variance of the first moment of the noise power B_g. Without further details, we simply give the results, namely

$$\left\langle B_g \right\rangle = 0, \qquad \left\langle \Delta B_g^2 \right\rangle = \frac{N(N+2)}{6(N+1)} \cdot \sigma^4. \qquad (C.17)$$

For the mean and the variance of x_g, we find

$$\left\langle x_g \right\rangle = 0, \qquad \left\langle \Delta x_g^2 \right\rangle = \frac{N(N+2)}{6(N+1)} \cong \frac{N}{6} \quad \text{if} \quad N >> 1. \qquad (C.18)$$

It is important to note that we used the well-known propagation of errors equations to calculate $\left\langle x_g \right\rangle$ and $\left\langle \Delta x_g^2 \right\rangle$ (see § 4.1.2).

The validity of the theoretical analysis presented in this Annex was confirmed with computer simulations using the scientific software Igor from WaveMetrics. Only the most relevant results are presented here. Figs. C.1 and C.2 show the mean value and the variance of the average noise power A_g in function of the noise standard deviation σ, as described by Eqs. (C. 10) and (C. 16). Following Eq. (C.17), the mean value and the variance of B_g versus the noise standard deviation σ are plotted in Figs. C.3 and C.4, respectively. Finally, in Figs. C.5 and C.6, we can see the results for the mean value and the variance of x_g as described by Eq. (C.18). These Figures show that the theoretical formulas developed here agree well with the computer simulations and, thus, they can be employed with confidence.

The essential theoretical results of this Annex are summarized in Table (C.1).

$\langle A_g \rangle = \sigma^2$	$\langle B_g \rangle = 0$	$\langle x_g \rangle = 0$
$\langle A_g{}^2 \rangle = \frac{N+3}{N+1}\sigma^4$	$\langle B_g{}^2 \rangle = \frac{N(N+2)}{6(N+1)} \cdot \sigma^4$	$\langle x_g{}^2 \rangle = \frac{N(N+2)}{6(N+1)}$
$\langle \Delta A_g{}^2 \rangle = \frac{2\sigma^4}{N+1}$	$\langle \Delta B_g{}^2 \rangle = \langle B_g{}^2 \rangle$	$\langle \Delta x_g{}^2 \rangle = \langle x_g{}^2 \rangle$

Table C.1 Statistical properties of sampled Gaussian noise.

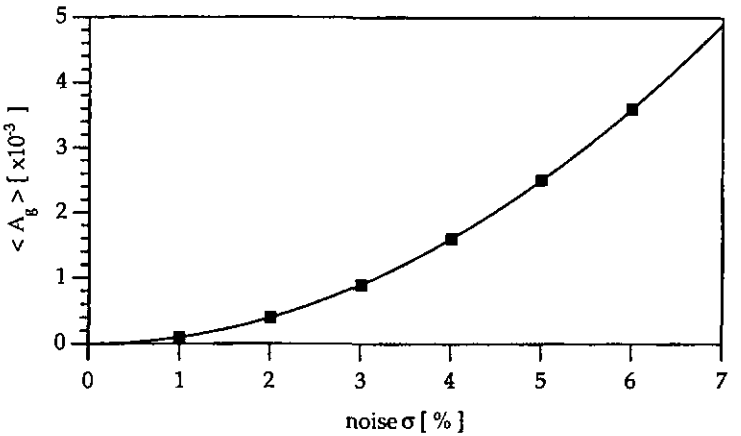


Fig. C.1 Theoretical prediction (plain line) and computer simulations (dots) for the mean value of the average noise power A_g as a function of σ. The number of samples is N = 256.

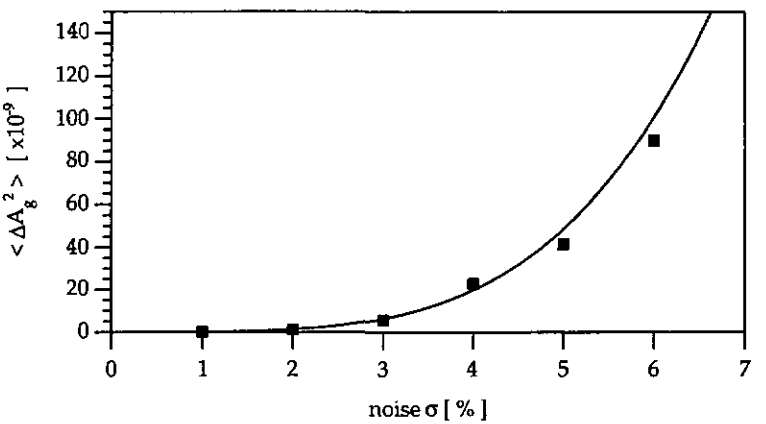


Fig. C.2 Theoretical prediction (plain line) and computer simulations (dots) for the variance of the average noise power A_g as a function of σ. The number of samples is N = 256.

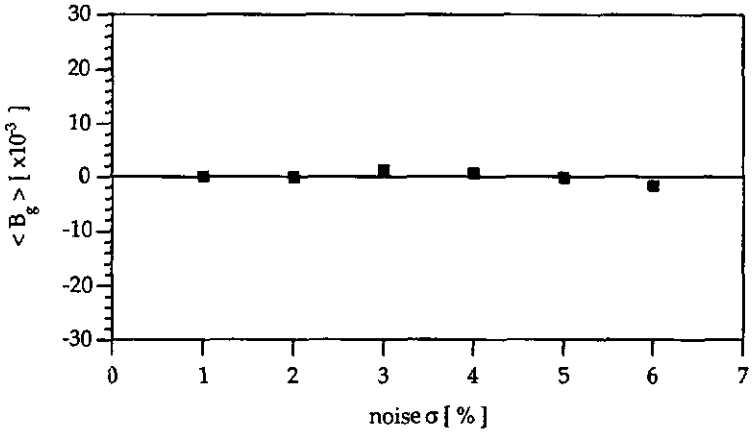


Fig. C.3 Theoretical prediction (plain line) and computer simulations (dots) for the mean value of the first moment of noise power B_g as a function of σ. The number of samples is N = 256.

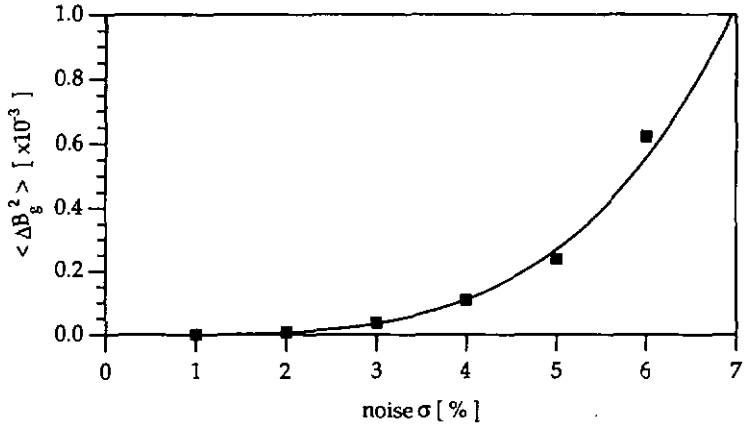


Fig. C.4 Theoretical prediction (plain line) and computer simulations (dots) for the variance of the first moment B_g of the noise power as a function of σ. The number of samples is N = 256.

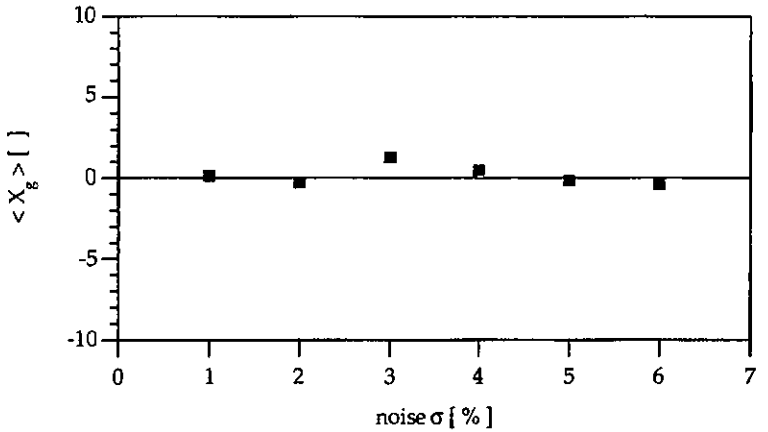


Fig. C.5 Theoretical prediction (plain line) and computer simulations (dots) for the mean value of the noise power center position x_g as a function of σ. The number of samples is N = 256.

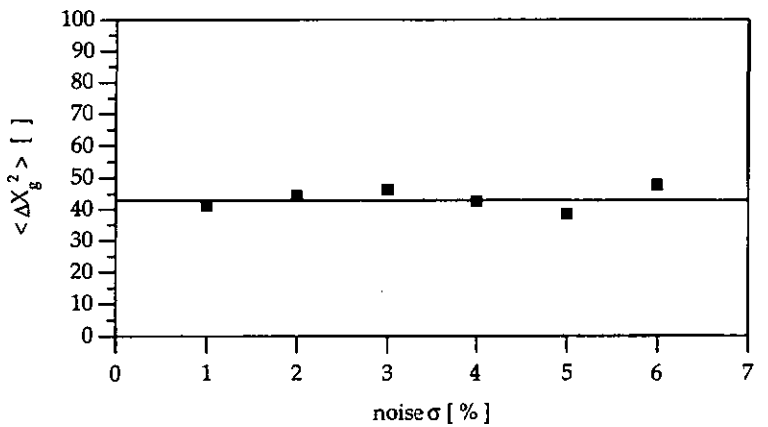


Fig. C.6 Theoretical prediction (plain line) and computer simulations (dots) for the variance of the noise power center position x_g as a function of σ. The number of samples is N = 256.

D. Center position using the fringe envelope: systematic error

In this Annex, we will see that the systematic error Δx_s depends of the method chosen to determine the center position x_s of the white-light fringe signal s(x). The first method consists of calculating the center position x_s directly from the sampled signal $s(x_n)$ (see Fig. D.1).

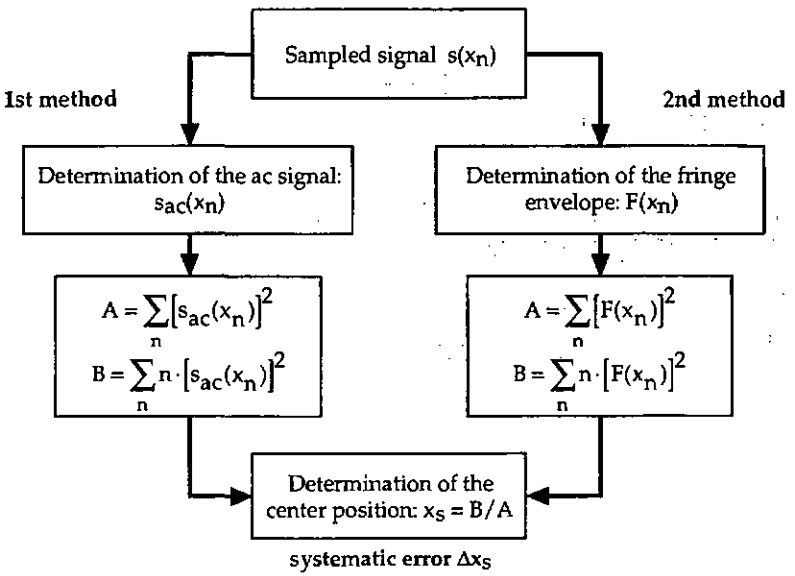


Fig. D.1 Two different methods can be employed to determine the center position x_s of the white-light fringe signal.

For this method the systematic error is given by [see Eq. (4.30)]

$$\Delta x_s^{(1)} \cong -(N+1)\frac{\sigma^2}{A_f^{(1)}} x_s. \qquad (D.1)$$

where σ^2 is the variance of the Gaussian noise and $A_f^{(1)}$ is defined as the total power of the ac signal without noise. For a Gaussian shape of the envelope, A_f can be expressed as [see Eq. (4.26)]

$$A_f^{(1)} \cong \frac{m^2 \Delta x}{2}\sqrt{\frac{\pi}{2}}. \tag{D.2}$$

Subsituting Eq. (D.2) into Eq. (D.1), we obtain for the systematic error

$$\Delta x_s^{(1)} \cong -(N+1)\frac{2\sigma^2 x_s}{m^2 \Delta x}\sqrt{\frac{2}{\pi}} \tag{D.3}$$

A second and more interesting procedure consists of calculating first the local value of the envelope $F(x_n)$, using Eqs. (4.36) to (4.38). Then, the center position x_s of the fringe envelope is obtained from

$$x_s^{(2)} = \frac{B}{A} \tag{D.4}$$

where A is the total power of the fringe envelope, given by

$$A = \sum_n F(x_n)^2, \tag{D.5}$$

and B is the first moment of the envelope power, namely

$$B = \sum_n nF(x_n)^2. \tag{D.6}$$

Applying the same approach as the one described in § 4.1.2, the mean value of the center position $x_s^{(2)}$ becomes then

$$\left\langle x_s^{(2)} \right\rangle = \frac{\langle B \rangle}{\langle A \rangle} = \frac{x_s}{1+\frac{7}{8}(N+1)\frac{\sigma^2}{A_f^{(2)}}}, \tag{D.7}$$

where $A_f^{(2)}$ is the total power of the fringe envelope $F(x_n)$. If $(N+1)(\sigma^2/A_f) << 1$, which is usually the case, we obtain finally

$$\left\langle x_s^{(2)} \right\rangle \cong x_s\left[1-\frac{7}{8}(N+1)\frac{\sigma^2}{A_f^{(2)}}\right]. \tag{D.8}$$

From Eq. (D.8), the systematic error is found to be

$$\Delta x_s^{(2)} = \langle x_s \rangle - x_s \cong -\frac{7}{8}(N+1)\frac{\sigma^2}{A_f^{(2)}} x_s. \tag{D.9}$$

For a Gaussian envelope

$$F(x) = m \exp\left[\frac{(x - x_s)^2}{\Delta x^2}\right], \tag{D.10}$$

where Δx is the width of the envelope and m is the fringe contrast, we find

$$A_f^{(2)} = \sum_{n=-N/2}^{N/2} F(x_n)^2 \cong \int_{-\infty}^{\infty} F(x)^2 dx \quad \text{if} \quad N >> 1. \tag{D.11}$$

Then

$$A_f^{(2)} \cong m^2 \Delta x \sqrt{\frac{\pi}{2}}. \tag{D.12}$$

This yields for the systematic error, when using the second method,

$$\Delta x_s^{(2)} \cong -\frac{7}{8}(N+1)\frac{\sigma^2 x_s}{m^2 \Delta x}\sqrt{\frac{2}{\pi}} \tag{D.13}$$

Comparing Eqs. (D.3) and (D.13), we see that

$$\Delta x_s^{(2)} = \frac{7}{16}\Delta x_s^{(1)}. \tag{D.14}$$

Thus, using $F(x_n)$ instead of $s_{ac}(x_n)$ to calculate the center position, the systematic error is reduced by a factor of 7/16.

Acknowledgments

I am particularly indebted to Prof. Dr. R. Dändliker for his invaluable advice and encouragement during this six-year period toward my Ph.D.. I could not suppose when he hired me in 1983 as an electronic engineer that 14 years later he would act as Director of my thesis. It was not only a great pleasure but also a great chance to work under his direction. Indeed, Prof. Dr. R. Dändliker has, at the same time, the well-recognized scientific competency and also the interest to transmit his inestimable knowledge to his students. He also gave me the opportunity to understand how to approach research in applied sciences. Moreover, he offered me his support and the opportunity to present the scientific results we obtained at the IMT all around the world.

Special thanks go to Dr. J.-F. Willemin from Mettler-Toledo (Greifensee, Switzerland) who designed the electronic circuits of the full-dynamic phase-demodulator and for his interest in my thesis as a member of the jury.

Further, I thank Dr. P. J. de Groot from Zygo Corporation (Middlefield, Connecticut, USA) and Prof. Dr. F. Pellandini from the Institute of Microtechnology of the University of Neuchâtel for kindly agreeing to be the examiners of my thesis.

Significant contributions to this work issued directly from stimulating and helpful collaborations with Prof. Dr. G. Frosio and Dr. J. Morel, who introduced me to the very interesting domains of interferometry and optical fibers. I wish to thank also Mr. Y. Salvadé who shared an office with me during the last three years.

I gratefully acknowledge for their technical assistance Mr. M. Groccia and Mr. A. Jeandupeux who realized the electronics which helped me to obtain important experimental results. I would also like to express my appreciation and gratitude to the many individuals at IMT, research assistants, technical collaborators and secretarial staff, who have contributed to the progress of my work and to the friendly and open atmosphere I encountered at the IMT.

Finally, my greatest thanks go to my family and my friends for their patience, encouragement, support and for many other reasons that do not have dealings with this thesis.