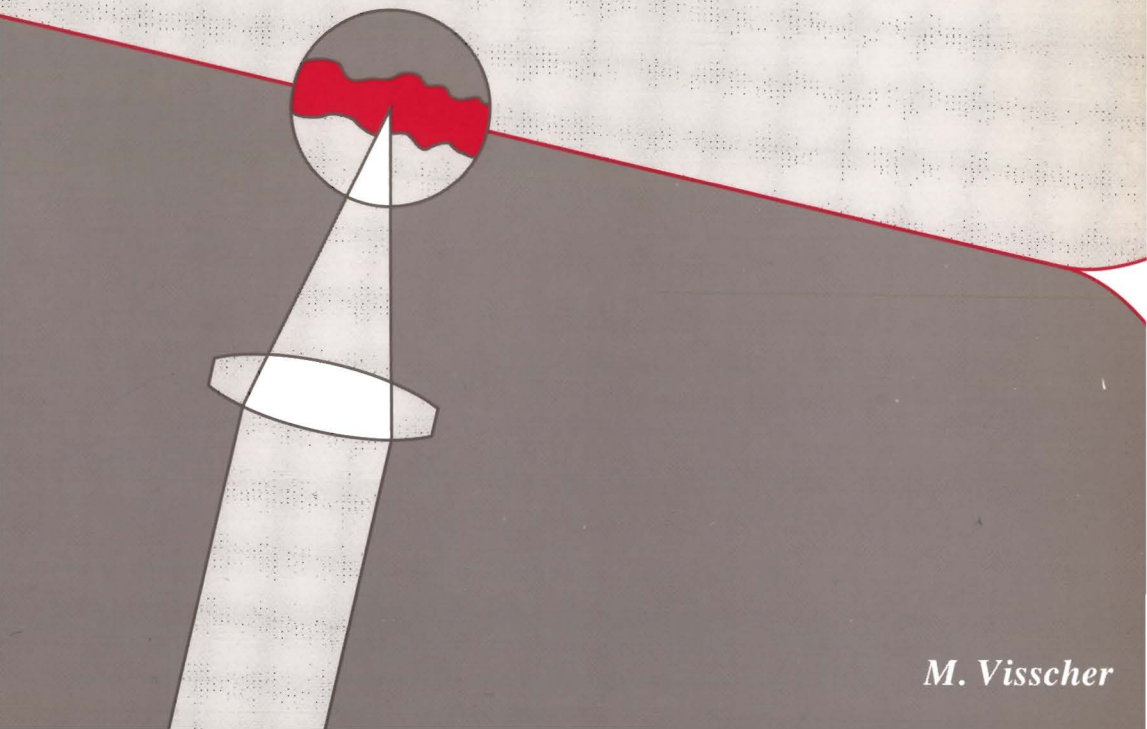
The measurement of the film thickness and the roughness deformation of lubricated elastomers

The measurement of the film thickness and the roughness deformation of lubricated elastomers



M. Visscher

THE MEASUREMENT OF THE FILM THICKNESS AND THE ROUGHNESS DEFORMATION OF LUBRICATED ELASTOMERS

THE MEASUREMENT OF THE FILM THICKNESS AND THE ROUGHNESS DEFORMATION OF LUBRICATED ELASTOMERS

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de Rector Magnificus, prof. dr. J.H. van Lint, voor een commissie aangewezen door het College van Dekanen in het openbaar te verdedigen op dinsdag 29 september 1992 te 16.00 uur

door

MARNIX VISSCHER

geboren te Goes

Dit proefschrift is goedgekeurd door de promotoren:

prof. dr. ir. M.J.W. Schouten
prof. dr. ir. E.A. Muijderman

en de copromotor:

dr. J.J. Baalbergen

voor mijn ouders

SUMMARY

A well known lubricated elastomeric machine element is the contact seal. Such seal types (e.g. for reciprocating motion) commonly operate in the mixed lubrication regime, where the friction is relatively large. In this regime the surface roughness can have a significant influence on the friction. It appears that the roughness asperities can deform, even when the mating surfaces are totally separated by the lubricant film. This deformation (flattening) must be accounted for and has been studied by a number of investigators, both theoretically and experimentally.

Reviewing the literature a large number of numerical results for varying, often simplified, conditions can be found nowadays, but experimental verification appears to be difficult. Therefore it was decided to develop a method for film thickness measurements, which has a sufficient resolving power to detect the (eventual deformed) roughness texture. The method must be applicable to elastomeric seals or similar contacts, in which a rough elastomer is in sliding contact with a smooth rigid body.

Different methods have been investigated on their suitability for the proposed measurements and the focus error detection method has been chosen as the most appropriate.

Further analysis of this method and some preliminary measurements showed its ability for film thickness measurements and the conditions which must be fulfilled.

The method can also be used to measure the deformed roughness texture in a statically loaded glass to elastomer contact. Two series of measurements are presented, one with and the other without liquid in the contact area. It is shown, that a liquid in the contact area can influence the contact situation significantly. The measurements on the dry contact yields the conclusion, that real contact areas can be found at different length scales.

Further investigation and development is needed to obtain the required accuracy and frequency range for the film thickness measurements.

SAMENVATTING

Een bekend voorbeeld van een gesmeerd elastomeer machine-element is de contact-afdichting. Een dergelijk type afdichting functioneert doorgaans in het gemengde smeringsgebied, waar de wrijving verhoudingsgewijs groot is. Hierbij kan de oppervlakteruwheid een belangrijke invloed hebben op de wrijving en de ruwheids-toppen blijken te kunnen vervormen, ook als de loopvlakken geheel door een smeerfilm gescheiden zijn. Deze vervorming (afplatting) moet verdiskonteerd worden in een theoretisch model en verschillende onderzoekers hebben dit feno-meen reeds bestudeerd, zowel theoretisch als experimenteel.

Uit literatuurstudie blijkt, dat een groot aantal numerieke resultaten beschikbaar zijn voor verschillende, vaak vereenvoudigde, kondities. Experimen-tele verifikatie blijkt echter niet eenvoudig te zijn. Daarom is besloten een methode voor filmdikte-metingen te ontwikkelen met een voldoend hoog scheidend vermogen om de (eventueel vervormde) ruwheidsstruktuur te kunnen onderscheiden. De methode moet geschikt zijn voor meting aan elastomere afdichtingen of aan vergelijkbare kontakten tussen een ruw elastomeer en een glad en hard lichaam in glijdende beweging.

De geschiktheid van verschillende methoden is onderzocht en de focus-fout methode bleek de meest geschikte te zijn.

De geschiktheid van de focus-fout methode voor filmdikte-metingen bleek uit verder onderzoek en uit eerste metingen. Hieruit is ook afgeleid aan welke kondities voldaan moet worden.

Naast filmdiktemeting kan de methode ook gebruikt worden voor meting van de vervormde ruwheidsstruktuur in een statisch belast kontakt tussen een glas-plaatje en een elastomeer. Van dit soort metingen worden twee series gepresen-teerd. De ene is uitgevoerd met, de andere zonder vloeistof in het kontakt. Het blijkt, dat de vloeistof grote invloed kan hebben op de kontaktsituatie. Uit de metingen zonder vloeistof volgt de konklusie dat ware kontakten op verschil-lende lengteschaal aanwezig zijn.

Verder onderzoek is nodig om de voor de filmdiktemetingen vereiste nauw-keurigheid en het benodigde frequentiebereik te realiseren.

CONTENTS

NOMENCLATURE

(the page where the symbol definition can be found is given in brackets)

A	**Area of contact**	[m]
	A_a = apparent area of contact	
	A_r = real area of contact	
A,B	**Photodiode signal** (32)	[V]
a,b	**Contact dimension** (211)	[m]
d	**Spot diameter**	[m]
	$d_{0.5}$ = fifty-percent-irradiance width (160)	
f	**Focal length**	[m]
fes	**Focus error signal** (33)	[V]
h	**Lubricant film thickness or gap height**	[m]
I	**Irradiance**	[Wm^{-2}]
	I_0 = irradiance maximum (160)	
n	**Index of refraction**	[-]
NA	**Numerical aperture** (161)	[-]
p	**Pressure**	[Pa]
	p_a = avarage or apparent contact pressure	
	p_0 = Hertzian contact pressure (211)	
R,r	**Radius**	[m]
res	**Radial error signal** (35)	[V]
$\mathcal{R}$	**Reflectance**	[-]
R	**Roughness height**	[m]
	R_a = Centre line avarage roughness height (CLA) (99)	
	R_a = Root mean square avarage roughness height (RMS) (99)	
	R_z = Peak to valley roughness height (avaraged over five adjoining sampling lengths)	
T	**Temperature**	[K]
t	**Window thickness**	[m]
u	**Velocity**	[$m \cdot s^{-1}$]
z	**Height distance**	[m]
η	**Dynamic viscosity**	[Pa·s]
ϑ	**Angle**	[rad]
λ	**Wavelength**	[m]
φ	**Angle**	[rad]

CHAPTER 1 INTRODUCTION

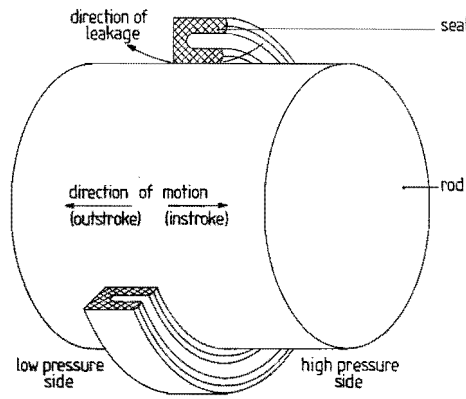
A familiar example of a lubricated elastomeric element is the elastomeric contact seal. Seals are widely used, e.g. in hydraulic cylinders, to prevent oil leakage. The performance of the seal, which is determined by the tribological process in the contact, is often hardly understood. Seal design is therefore merely based on trial and error methods and on the designer's experience. Many investigations have thus been dedicated to the understanding of the tribological behaviour of seals in particular and the lubrication of elastomers in general. At Eindhoven University the leakage and friction of reciprocating seals have been studied (see e.g. Kanters and Visscher, 1989; Kanters, Verest and Visscher, 1990; Kanters, 1990, 1991) and this thesis can be regarded as a follow-up.

In section 1.1 investigations on reciprocating seals (used in e.g. hydraulic cylinders) will be briefly reviewed to find out, what is at present known about the tribological behaviour of such seals. The seal's surface roughness will appear to be an important factor and a further review will therefore focus on that matter. One of the conclusions will be that the seal's surface roughness can be deformed due to local hydrodynamic action at the asperity slopes. However, theoretical calculations on this matter, reviewed in section 1.2.1, are difficult because of the more or less random nature of the roughness and the importance of local effects, while proper experimental methods, reviewed in section 1.2.2, are hardly available to measure the roughness deformation in the lubricated contact. Therefore a method will be develop for film thickness measurements, enabling also the detection of the real roughness texture in the lubricated contact (section 1.3).

1.1 Reciprocating seals

A large number of reciprocating seal types exists, as can be found in the catalogues of manufacturers, varying from a simple O-ring design to seal systems with a complex geometry. Figure 1.1 shows a so-called U-type rod seal as mounted in the housing of e.g. a hydraulic cylinder (not shown), while the rod moves relative to the seal yielding lubrication of the seal-rod contact. This lubrication has both a positive and a negative effect: it reduces the friction (and wear), but it also causes leakage. Consequently, seal designers have to fulfil two contradictory requirements since both a low friction and a low leakage are

Figure 1.1
Reciprocating
U-type seal



wanted. In practice different seal types must be (and are indeed) developed, each being designed to realize a reasonable balance between friction and leakage for a specific application. Some examples are:

- Fluid leaking at a piston seal remains in the system. Piston seals can therefore be designed for low friction.
- Oil leaking at the rod seals of hydraulic cylinders is lost to the environment and should thus be prevented.
- The fluid in pneumatic cylinders is cheap while the driving pressure, and thus the power density, is low. A low friction of the rod seal is then more important than a low leakage rate.

Proper calculation of leakage and friction of seals is important in seal design and the present knowledge of the lubrication of reciprocating seals will therefore be discussed in the following.

1.1.1 The lubrication of reciprocating seals

Regardless of the specific design of the seal, all reciprocating seals have in common that the rod motion is in the same (or opposite) direction as the pressure gradient (see fig. 1.1). As a result, the lubrication problem is characterized by a one-dimensional flow through the seal-rod contact and the leakage is directly correlated to the lubricant film thickness by conservation of mass. (This correlation has frequently been used in reciprocating seal research to overcome the problem of direct film thickness measurement. Leakage measurements are then performed to estimate the lubricant film thickness in the

contact, as we will see later in this section).

Theoretical solutions of the one-dimensional lubrication problem are available to calculate e.g. friction and leakage, provided that the mating surfaces are completely separated by the lubricant film. However, reciprocating seal design appears to be a trial and error process. This is among others caused by the difficult calculation of the contact situation due to large deformations, the nonlinear stress-strain relation, the nonlinear boundary conditions and the (nearly) incompressibility of the elastomers. The possible occurrence of local contact between the mating rod and seal surfaces implies another difficulty. These problems makes theoretical seal research difficult and a lot of experimental work has therefore been performed.

Experimental reciprocating seal research often includes friction measurements. The measured friction curves often resembles (a part of) the well known Stribeck curve (fig. 1.2), in which three lubrication regimes are distinguished:

1. *Full Film Lubrication (FFL):* The mating surfaces are completely separated by a lubricant film and the friction only originates from the viscous shear in the lubricant.
2. *Mixed Lubrication (ML):* Decreasing the velocity the mixed lubrication regime is reached, where the friction increases remarkably at further reduction of the velocity. A general accepted explanation is, that the film thickness reduces at decreasing velocity and is now too thin too prevent asperity contact. The friction is then a result of both viscous shear in the lubricant film and of friction in the asperity contacts.

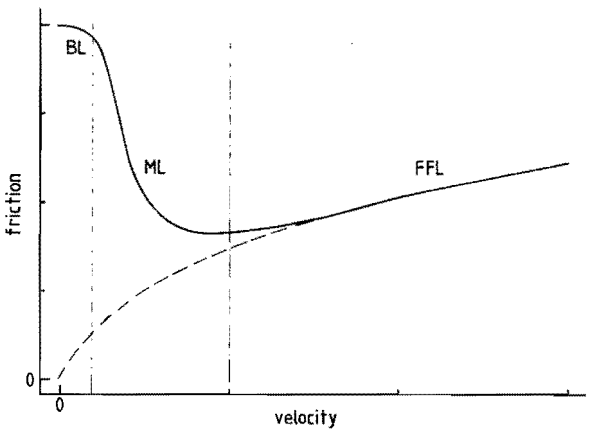


Figure 1.2
The Stribeck curve.
FFL= Full Film Lubrication;
ML = Mixed Lubrication;
BL = Boundary Lubrication;

(The dashed line represents the friction according to full-film theory)

3. *Boundary Lubrication (BL):* At lower values of the velocity the lubricant film is very thin. The number of contacting asperities is then so large, that the friction is fully determined by the shear stresses in the asperity contacts. (However, direct contact between the solids can still be prevented by an eventual one molecule thin layer of an oil component, chemically fixed to the surfaces. This explains the name "boundary lubrication").

Nowadays, reliable theoretical models to calculate the friction are only available for the full film lubrication regime.

1.1.1.1 The lubrication regime in which reciprocating seals function

The state of lubrication of the seal is often derived from the Stribeck curve only. However, we will see in the following that one must be very cautious in doing this.

When full film lubrication occurs, the friction and leakage can directly be derived from the lubricant film profile. Therefore the friction force can be estimated from a calculated or measured film profile or from measured leakage, when full film lubrication is assumed. Whether full film lubrication really occurred can then be evaluated by comparing this estimation of the "full film" friction with the measured friction.

Several investigators reported, that the friction was often much higher than expected from film thickness and/or leakage measurements (see Field, 1973 pp. 108 and 160-161; Field and Nau, 1973[a] pp. 15-16, 1973[b] pp. 14-15, 1973[c], 1976; Kanters and Visscher, 1989) or from full film calculations (e.g. Johannesson, 1989). Kawahara, Muto et al. (1981) concluded from the measured friction curves, which were like a Stribeck-curve, that the seals commonly operate in the mixed lubrication regime. A similar conclusion was drawn by Johannesson (1989).

Field and Nau (1973[c], 1976) reported that their capacitive and interferometric film thickness measurements indicated full film lubrication, while the measured friction curves were typical for mixed lubrication. The origin of this difference remained unclear.

Considering their capacitive film thickness measurements, the simultaneously measured leakage (Field, 1973 pp. 138ff.) was also in quantitative disagreement with the measured film thicknesses, the leakage being about five to

ten times lower than expected[1]. An explanation was not given.

Considering the interferometric film thickness measurements, the presence of undetected asperity contacts was suggested to explain the high friction (Field, 1973 p. 108; Field and Nau, 1976). However, the origin of such asperities is unclear since both the elastomer and the glass surfaces were optically smooth to obtain a reasonable interference pattern (Field 1973 pp. 70 and 74; Field and Nau 1973[a] p. 6).

The study of Kanters and Visscher (1989) involved leakage and friction measurements on three rods with a different surface roughness. The following roughness values were given:

rod A: $R_a < 0.01$ μm; $R_z = 0.06$ μm
rod B: $R_a = 0.05$ μm; $R_z = 0.36$ μm
seal: $R_a = 0.54$ μm; $R_z = 3.80$ μm

The roughness of rod B was typical for the roughness of rods in hydraulic cylinders.

The friction curve measured on rod B for an instroke (see fig. 1.1) looked like a Stribeck curve (solid lines in fig. 1.3a), which could give rise to the conclusion of full film lubrication at higher velocities. However, full film lubrication did not occur anywhere at outstroke, where the friction was significantly lower (see the dashed lines in fig. 1.3a). This was concluded from comparison of the measured friction with the friction calculated from the leakage flow at outstroke assuming full film lubrication (fig. 1.3b). One must therefore be very cautious to conclude full film lubrication from a measured friction curve alone.

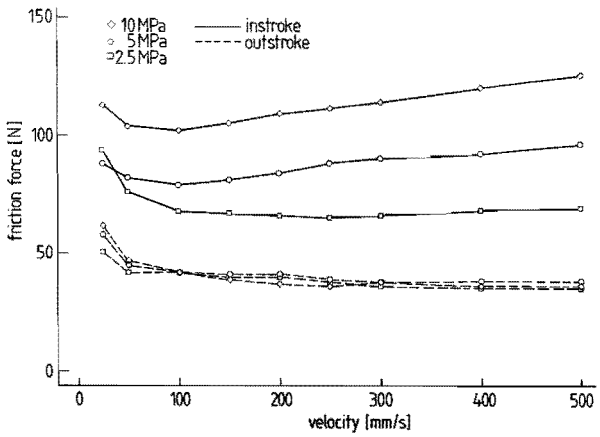
Using the smoothest rod (rod A) the friction and leakage were very similar to those using rod B. This yielded the conclusion that the roughness of the rod, which is significantly lower than that of the seal, is not important.

One may suggest now that the friction can be significantly reduced by reduction of the seal's surface roughness, since the transition from full film to mixed lubrication will then occur at a thinner lubricant film. If the seal's roughness is small enough (e.g. of the same order as the rod roughness), full

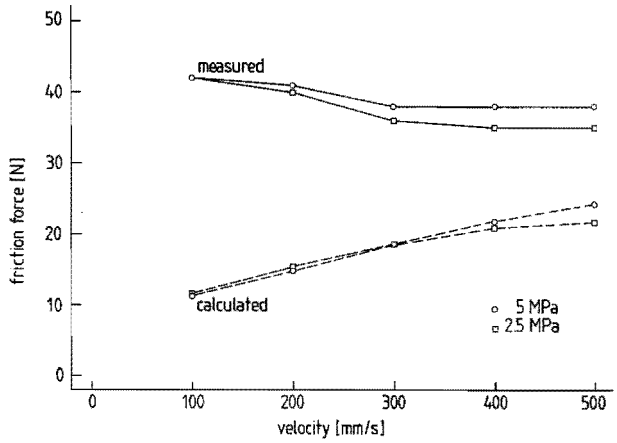
[1] This was concluded from comparing the measured leakage (as e.g. shown in fig. 4.30, p. 167 of Field's thesis) with the leakage expected from the difference in out- and instroke film thickness (shown in e.g. fig. 4.12, p. 146, and fig. 4.21, p. 157, in Field's thesis) for the same measurement series.

Figure 1.3
Friction of a seal with rectangular cross section on rod B.
(Source: Kanters and Visscher, 1989)

a. Measured friction for three values of the pressure in the cylinder.



b. Measured friction at outstroke compared with calculated friction, obtained from outstroke leakage assuming full film lubrication.



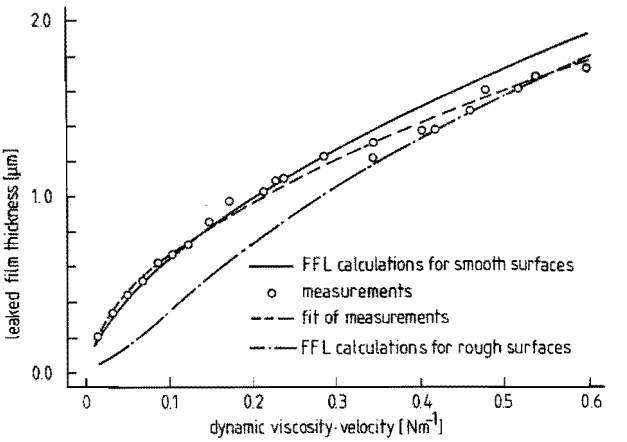
film lubrication might occur in a wide range of working conditions (system pressure, velocity, viscosity) without increase of leakage. However, reduction of the seal's surface roughness will increase the costs of the seals and is only justified if the higher price is at least compensated by the savings in energy and by a longer life cycle. These savings must then be predictable and that means that the effect of the roughness on e.g. the friction must be known. Further study of Kanters therefore considered the influence of the seal's surface roughness.

1.1.1.2 The influence of the seal's surface roughness

Kanters (1990, 1991) and Kanters, Verest and Visscher (1990) calculated the leakage, assuming full film lubrication (FFL) and smooth rod and seal surfaces (solid line in fig. 1.4), and compared it with proper measurements (the points and the dashed line in fig. 1.4). One of their conclusions was that the lubrication is influenced by the seal's surface roughness. To account for this, the average flow model of Patir (1978) and Patir and Cheng (1978, 1979[a], 1979[b]) was applied (the dot and dash line in fig. 1.4). The correspondence of calculated and measured leakages was now very good at large values of the product of viscosity and velocity, but got poorer at lower values (see fig. 1.4). This was attributed to micro-Elasto-Hydrodynamic Lubrication (micro-EHL) at thinner films, i.e. the roughness is flattened because of local pressure generation at the asperity slopes, which was not accounted for by Patir and Cheng.

This roughness flattening influenced the moment, at which transition from full film to mixed lubrication occurred. A traditional idea is that asperity contact occurs at films thinner than about $3R_q$, provided that the roughness height distribution is (nearly) Gaussian (see e.g. Patir, 1978; Patir and Cheng, 1978[a]; Cheng, 1985; Elrod, 1978). However, the transition actually occurred at a film thickness approximately equal to the R_q value (which was about 0.83 μm). This also indicates flattening of the roughness texture. More details are given by Kanters (1990 pp. 99-104; 1991).

Figure 1.4
Outstroke leakage, i.e. the film thickness on the rod after a motion out of the sealed housing.
(Source: Kanters, 1991)



1.1.2 Conclusions

Proper calculation of the leakage and friction is important in the design of (reciprocating) elastomeric seals. Such seals appear to function in the mixed lubrication regime, where the friction is high compared with the friction in the full film regime. Reliable models for prediction of the friction in this regime are not available at the moment and more research on this subject is needed. An important factor is the seal's surface roughness and its deformation due to micro-Elasto-Hydrodynamic Lubrication. Therefore, a literature review on the lubrication of rough surfaces will now be presented.

1.2 The lubrication of rough surfaces, a literature review

As we have seen in the former section, the seal's surface roughness appears to be a factor in the lubrication of seals. This roughness influence must be accounted for and a proper theoretical model is thus needed. The lubrication of rough surfaces has received much attention in literature in general, mostly not specially dealing with elastomers, and a short literature review on the theoretical and the experimental work will be presented below.

1.2.1 Review of theoretical work

Review papers on the subject were written by Elrod (1978), Dyson (1978) and Cheng (1985). Several distinctions can be made considering the theoretical analysis of the lubrication of rough surfaces. Some, given by Elrod and Dyson, are listed below:

- The roughness structure is two-dimensional or three-dimensional. Two-dimensional means, that the roughness is orientated in one direction while the roughness height is constant in the other direction, e.g. in the case of grooves in the surface. A three-dimensional roughness texture has height variations in all directions.
- One surface is rough and the mating smooth (single-sided surface roughness) or both surfaces are rough (two-sided surface roughness).
- The two mating surfaces are fully separated or are locally contacting.
- Elrod also distinguishes between "Reynolds roughness" (having rather small asperity slopes) and "Stokes roughness" (with larger slopes). The

reason for this distinction is that the Reynolds equation neglects the flow in the direction perpendicular to the mating surfaces. This is only justified for small (asperity) slopes. When the asperity slopes are larger, the general Navier-Stokes equations should be applied.

In the case of asperity contact, difficulties are introduced by the necessarily presence of local deformation. The local film thickness is often derived by superposing the roughness profile on the film profile calculated for smooth surfaces. This can lead to overlap of asperities of both surfaces. Commonly, the film thickness is simply taken zero in these overlapping areas, according to Dyson (1978), while the profile outside these areas is left undeformed. Another question is, whether (very) thin films are present in the asperity contact areas or not (see e.g. Jacobson, 1990). Finally, the asperity deformation, due to elastohydrodynamic effects, should be considered.

Different methods are used for the theoretical analysis. The roughness texture is e.g. simplified by a sinusoidal wave or the roughness profile is described by statistical parameters.

We will now review the theoretical work in more detail, distinguishing between models for two- and models for three-dimensional roughness textures. Where appropriate for the survey, a further distinction will be made between single-sided and two-sided roughnesses. The average flow model, proposed by Patir and Cheng (1978) and already mentioned in section 1.1, does not fit in these classifications and will be reviewed separately, since it can in principle be applied to all kinds of problems mentioned here. Finally, attention will be paid to effects, which may be introduced by non-Newtonian behaviour of the fluid.

1.2.1.1 Lubrication of surfaces with two-dimensional roughness

Single-sided roughness

According to Elrod (1978), the earliest analyses were dedicated to the two-dimensional single-sided model roughness (e.g. sinusoidal waves). Numerical calculations are often performed for a longitudinal or transverse roughness. In the case of rather long wavelengths, an average Reynolds equation (with average pressure and average film thickness) has been used. Also, statistical methods have been applied.

Recent numerical solutions of the film thickness profile and the contact pressure distribution are given by Lubrecht et al. (1988), considering the elastohydrodynamic asperity deformation. The calculations were performed for the point contact of a stationary surface with a restricted number of sinusoidal asperities and a smooth sliding surface. The lubrication problem was thus stationary and two-dimensional. Both a longitudinal and a transverse model roughness were considered. Flattening of the "asperities" is shown in the plots, but not discussed. Similar calculations, but only for a transverse roughness, were performed by Kweh et al. (1989), who found similar results. Venner (1991 section 10.5 pp. 270-283) found, that the flattening was much more pronounced in the case of transverse roughness than in the case of longitudinal roughness. Further studies by Kweh et al. (1992) consisted of calculations on a transverse roughness with two sinusoidal waves, having a different amplitude and wavelength, superimposed. They showed, that the deformation of the smaller scale wave component, with the smaller wavelength and amplitude, was less pronounced than the deformation of the larger scale component. A preliminary calculation using a measured transverse roughness texture (in which small scale components predominates) therefore showed less flattening than the sinusoidal waves.

Venner (1991 section 7.3 pp. 177-186) also calculated the steady state line contact problem of a stationary surface with a transverse roughness and a sliding smooth surface. A measured roughness profile was used in the calculations and Venner found that the surface roughness could be flattened remarkably under sliding motion (p. 183 of his thesis). Chang and Webster (1991) found similar results for a line contact with a stationary surface, having a sinusoidal transverse roughness, and a moving smooth surface, but reported that the roughness deformation was only large at low sliding velocities. Under conditions of high sliding velocities, no significant roughness deformation resulted. These results are in qualitative agreement with the experimental results of Kaneta and Cameron (1980) and of Cusano and Wedeven (1981) (see also section 1.2.2 page 14) and Kanters (1990, 1991) (see also section 1.1.1.2 page 7).

Chang and Webster (1991) found no significant roughness deformation under the transient conditions of pure rolling with a nominal film thickness larger than the roughness amplitude. These results are also in qualitative agreement with the experimental results of Kaneta and Cameron (1980) (see also section 1.2.2 page 13). Venner (1991 section 8.5 pp. 208-211) also calculated the film profile for a sinusoidal transverse roughness, but now with a nominal film thickness which is about equal to the undeformed roughness amplitude. He also found that there was less deformation under conditions of pure rolling, but the

flattening was found to be more pronounced when the asperity has traveled farther in the contact. This was attributed to micro-squeeze effects, which are more pronounced when the nominal film thickness is almost equal to the amplitude of the waviness.

Two-sided roughness

Chang and Webster (1991), already mentioned above, not only considered single sided roughness, but also the transient conditions of a transversal two-sided roughness. One of the conclusions was, that the roughness deformation is now also present under sliding conditions with higher velocities, where the deformation was very small in the stationary single-sided situation.

1.2.1.2 Lubrication of surfaces with three-dimensional roughness

Kweh et al. (1989), already mentioned in section 1.2.1.1 page 10, also calculated the film thickness in the steady state situation of a sliding elliptical contact, formed by a moving smooth surface and a stationary surface with a three-dimensional roughness formed by two sinusoidal waves: one in the transverse and the other in the longitudinal direction. Now they found, that the transverse component of the roughness was almost completely flattened, while the grooves in the longitudinal direction remained.

1.2.1.3 Average flow model

Patir (1978) and Patir and Cheng (1978, 1979[a], 1979[b]) introduced an "Average Flow Model" to calculate the flow through the contact of rough surfaces. All kinds of surface roughness textures, two- and three-dimensional, single- and two-sided, can in principle be treated. They used an average Reynolds equation, introducing flow factors which account for the influence of the asperities superimposed on the (smooth) average film profile. These flow factors are numerically calculated from statistically generated roughness textures. Similarly, the expected friction can be calculated using shear stress factors. Patir and Cheng claimed that their model is valid for both (elasto-) hydrodynamic lubrication and mixed lubrication. Local asperity deformation is, however, not considered.

The average flow model of Patir and Cheng has received much attention in literature and the calculation of the flow and shear stress factors has been widely discussed. The earlier discussions were summarized by Cheng (1985) and later discussions were given by Hu and Zheng (1985, 1989), van Bavel (1987), Lubrecht et al. (1988), Zhu et al. (1990), Kanters (1990, 1991; see also section 1.1.1.2 page 7), Venner (1991) and Chang and Webster (1991). Several methods, both analytical and numerical, have been applied to calculate the flow and shear stress factors, yielding different values. These differences can be large (e.g. up to 100 percent!) when the film is thin relative to the standard deviation of the roughness height distribution.

Also, the negligence of the asperity deformation can cause large errors. Kaneta and Cameron (1980), Kanters (1990, 1991) and Chang and Webster (1991), found that the flattening of the roughness, present on the stationary surface in a sliding contact, was more pronounced at a thinner (nominal) film, while the transition from full film to mixed lubrication appears to occur at a thinner film than expected from the undeformed roughness (Kanters, 1990, 1991).

1.2.1.4 Lubrication of rough surfaces with non-Newtonian fluids

Thus far, the influence of the viscosity model was not mentioned in our review. Chang et al. (1989) found different results applying different viscosity models. They calculated the film thickness and pressure distribution for a line-contact with a transverse simple shaped surface irregularity (representing a roughness asperity) on one surface. Considering the steady state condition of a moving smooth surface in contact with a stationary "rough" surface, they found that the roughness deformation was less pronounced (roughly two times smaller) using the non-Newtonian Eyring model than in the case of a Newtonian model. Under conditions of pure rolling, no significant difference was found. This was explained by the fact that the shear stresses are much lower in the case of pure rolling than in the case of sliding, while the Eyring model approaches the Newtonian model for low shear stresses.

A different subject is the possible occurrence of asperity contact. Jacobson (1990) proposed that asperity contact can not occur with Newtonian fluids, since local decrease in film thickness at an asperity will cause pressure increase and therefore flattening of the roughness. When the pressure would not be large enough to deform the asperity, the film will become thinner, the pressure

increase will be larger (because of the wedge of the asperity slope or because of squeezing) and the asperity will deform at last.

Jacobson proposed a model which can explain the existence of contact at sliding motion with high pressures. The essence is that the lubricant behaviour is non-Newtonian at a high pressure and at a large strain rate. The shear stress is then not proportional to the strain rate and the pressure flow perpendicular to the sliding direction of the mating surfaces can be larger than it would be with a Newtonian fluid. When the side flow is large enough, the initially oppressed asperity can reappear and contact is possible.

1.2.2 Review of experimental work

A significant number of papers deals with experimental study of the lubrication of rough surfaces, e.g. to determine whether asperity contact occurs (see e.g Kawahara, Ohtake and Hirabayashi, 1981; Leather and McPherson, 1978; Ogata et al., 1987; Schmidt et al., 1987[a], 1987[b]), but only a few provide clear inside into the subject. These will be reviewed below.

Kaneta and Cameron (1980) studied the lubrication of a metallic ball, in contact with a flat and smooth glass disc, by means of optical interferometry. A three-dimensional periodic roughness texture, consisting of trapezium shaped asperities, was applied to the ball. The regularity of the roughness texture was necessary to obtain good quality interferograms, since trials with normal, random rough surfaces were not successful (see also Jackson and Cameron (1976) and appendix B4.1.4 of this thesis).

The results of Kaneta and Cameron were especially very interesting. No asperity deformation was observed under pure rolling conditions. Under pure sliding, however, the asperities flattened at low velocities and with thin films, just as Kanters (1990, 1991) concluded from his experiments (see section 1.1.1.2 page 7) and as some investigators concluded from numerical calculations (see section 1.2.1.1 and 1.2.1.2, pages 9-10). At high velocities, they observed that the height of the deformed asperities was larger than the undeformed asperity height of the free surface. Kaneta and Cameron suggested that this might be caused by deepening of the valleys due to hydrodynamic action between the asperities, but this phenomenon is not understood.

Cusano and Wedeven (1981) also used interferometry to study the surface roughness effects. They also measured the film thickness in the contact of a smooth glass plate and a ball with some simple shaped "asperities" on it. One of their conclusions was, that the asperity deformation was more pronounced at sliding than at pure rolling. However, this difference was very small for a longitudinal roughness, where it was significant for e.g. transverse roughness. In the case of sliding motion and transverse roughness, they also found that the deformation was less pronounced when the film was thicker. These results compare qualitatively to the results of Kaneta and Cameron (1980).

Jacobson (1990) measured the oil film resistance between a smooth, metallic ball and a rough, hard metallic, flat surface to verify his theoretical model mentioned in section 1.2.1.4 (page 12). Part of the experiments was under pure squeeze conditions and part was under combined squeeze and sliding conditions. The occurrence of contact was concluded when a resistance decrease was observed during the contacting time. The method was not able to determine the roughness texture in the lubricated contact.

Jacobson found that a sliding contact needs a higher viscosity to prevent roughness contact than a squeezing or rolling contact. This agrees with the theory (see section 1.2.1.4, page 12), since the strain rate is typically higher under conditions of sliding motion than under pure squeeze (and pure rolling) conditions.

1.2.3 Conclusions

Theoretical analyses and experimental methods for the study of roughness effects in lubricated contacts still have their limitations, but some conclusions on the roughness deformation can already be drawn. The limitations of theoretical and experimental methods and the conclusions on the roughness deformation will be summarized separately.

The limitations of the theoretical analyses for the study of roughness effects in lubricated contacts

We have seen that the lubrication of rough surfaces has received much attention, especially during the last decade when the possibilities of numerical methods increased extensively. However, there are still restrictions concerning the

problems which can be treated:

- Most calculations are dedicated to simple two-dimensional roughness textures like sinusoidal waves, some including transient effects, e.g. in the case of pure rolling.
- Some calculations are presented for the more general and more common three-dimensional roughness, but only with a simplified regular texture and for steady state conditions.
- Finally, some calculations using measured roughness profiles are already performed, but they are still limited to the steady state situation with two-dimensional roughnesses.

Further development of numerical methods and computer power will certainly enable the calculation of more complex problems, e.g. incorporating three-dimensional roughness textures.

The average flow model of Patir and Cheng has received much attention and seems to be promising. However, it is only applicable in the present form for conditions in which asperity deformation does not occur. Also, the proper calculation of the flow and shear stress factors, needed in their model, is still subject of discussion, especially in the mixed lubrication regime.

The limitations of the experimental methods for the study of roughness effects in lubricated contacts

Considering experimental work, only a view papers are available. The eventual roughness deformation in lubricated contacts has only been detected by using a simplified regular roughness texture. Using surfaces with a "normal" roughness texture, only the eventual occurrence of contact was detected. Therefore, their is need for better experimental methods, which allows the detection of more realistic roughness textures in lubricated contacts.

Conclusions on the surface roughness effects in lubricated contacts

Both theoretical and experimental work yielded the conclusion that the surface roughness can be deformed due to hydrodynamic action on the asperity slopes. This deformation ("flattening") occurs especially at sliding motion, both in line and in point contacts. Besides, the deformation seems to be less pronounced at a surface waviness with a smaller wavelength.

Furthermore, non-Newtonian fluid behaviour can be an important factor, since there is some theoretical evidence that non-Newtonian fluid behaviour reduces the roughness flattening. Further, consideration of possible non-Newtonian behaviour is essential concerning the eventual occurrence of asperity contact.

1.3 Objective of this thesis

We have seen that the surface roughness can be an important factor in lubrication. As concluded in section 1.1 the lubrication of reciprocating seals is influenced by the seal's surface roughness and this surface roughness is deformed (flattened) due to micro-EHL. Besides, reciprocating seals appear to operate generally in the mixed lubrication regime, where the friction is much higher than in the full film regime and where the lubricant film is thin compared with the seal's surface roughness. The reason is that thin films are desired to reduce the leakage. In this mixed lubrication regime, the seal's roughness is especially important, but theoretical models to predict the leakage and friction are not available. Understanding of the influence of surface roughness on these important seal properties is, however, necessary to enable better seal design.

At Eindhoven University research will focus on the lubrication of rough elastomers, considering the asperity deformation caused by micro-EHL. This research will include theoretical as well as experimental work.

Different theoretical models have been developed in the past, but are still restricted to simplified roughness textures or the steady state situation of a two-dimensional roughness. Also, experimental verification appears to be difficult. The aim of this study is therefore **to develop a method to measure the lubricant film thickness** in the contact of a rough elastomer and a smooth rigid body. These measurements must be performed on a **sufficient small scale to detect the (eventual deformed) surface roughness during motion.**

The distinction between metal to metal and elastomer to metal contacts is made, because of the much higher pressures in metal to metal contacts and the difference in physical properties like the conductance and the reflectance. The latter fact may cause different methods to be preferred for the two different configurations. Considering the contact pressures, the pressure dependence of many physical (fluid) properties is not properly described at high pressures. Well known are the discussions on the description of piezo-viscous effects (see e.g. Witt, 1974; Dyson et al., 1966). The accuracy of calculations is therefore

uncertain, when the pressures are high (e.g. 1 GPa). Similarly, the pressure influence on e.g. the electrical permittivity and the index of refraction is not well known, reducing the accuracy of the measurements at higher pressures, especially when the pressure itself is hardly known. In fact, relatively large pressure variations may occur in the contact because of the presence of the asperities.

The use of a rigid body as counterpart for the elastomer implies, that the counterface does in essence not deform. Therefore the film thickness can be determined by measurement of the distance between a transducer and the elastomeric surface, when the transducer is fixed relative to the rigid body.

Only single-sided roughness, with a rough elastomer and a smooth rigid body, will be considered, because the reciprocating research programme yielded the conclusion that the roughness of the rod (i.e. the rigid body) was of minor importance. A transducer can then be properly attached to the rigid surface without the necessity to account for the possible influence of the surface roughness of the rigid surface nor for the possible influence of the mechanical properties of the elastomer, when a transducer would be attached to it.

The method should be applicable to elastomeric seals. However, the initial experiments on the influence of the surface roughness will not be performed on a seal, but on an elastomeric body of simple geometry. A rig on which the load can easily be varied will be used, but the test conditions (velocities and contact pressures) will be typical for seals.

CHAPTER 2 THE REQUIREMENTS FOR THE FILM THICKNESS TRANSDUCER AND CHOICE OF THE METHOD

As mentioned in section 1.3, the aim of this research is to develop a method for film thickness measurements in the contact of an elastomer and a smooth rigid surface at a sufficient small length scale to detect the eventual deformed roughness during motion. Many methods, e.g. mechanical, electrical and optical, are in principle available and we must therefore firstly investigate them on their suitability before a final choice can be made. The suitability of a particular method depends of course on the requirements which must be coped. These requirements originate partly from the environment, like e.g. the properties of the contacting bodies, the contact pressure and the contact temperature. Other requirements are e.g. the resolution and the dynamics of the measurement. The most important is of course the accuracy.

Before the requirements can be specified, the situation, in which the measurements must be performed, has to be described properly. This description has essentially been given in section 1.3 and will be summarized below in section 2.1 with addition of some more aspects.

When the requirements are specified (section 2.2) the method will be chosen in section 2.3.

2.1 Characteristics of the measurement situation

The situation is characterized by:

1. *The materials of the mating surfaces:* One is elastomeric and the other is rigid (commonly metallic, sometimes the rod of hydraulic cylinders is of ceramics);
2. *The surface finish of the mating surfaces:* The rigid surface is smooth and the elastomeric is rough. The characteristics of common roughness textures are given in appendix A, where is shown that roughness height variations on a micrometre scale (i.e. occurring with a wavelength of the order of a micrometre) are significant;
3. *The kind of motion*, which is sliding motion;
4. *The position of the contact*, which is related to the position of the elastomer, as is the normal situation with seals;

5. *The items to be studied:* Besides the nominal film profile, the possible deformation of the surface roughness is important;
6. *The thickness of the (nominal) film profile*, which is commonly of the order of 0.1 to 1 μm;
7. *The velocities in the contact*, which can e.g. be up to 1 m/s for reciprocating seals.
8. *The pressures in the contact*, which can be of the order of 10 MPa (e.g. up to 50 MPa in hydraulic actuators).
9. *The temperatures in the contact*, which are generally low for reciprocating seals, mainly due to the axial rod motion spreading the generated heat over a large area. In general the temperatures in the contact of elastomeric machine elements can be higher (e.g. 100 °C in the case of radial lip seals, according to Stakenborg (1988 section 5)). Temperatures higher than roughly 200 °C are rare, because most elastomers are not resistant to such temperatures.

2.2 Requirements

The most important characteristic of a method is its accuracy, which is determined by the systematic and by the random deviations. The uncertainty in the measurements caused by random deviations must be determined by testing the reproducibility, i.e. by statistical analysis of a number of measurements under equal conditions. The influence of systematic deviations can be eliminated, when this influence is quantified. This requires analysis of the method, considering its own physical limits as well as parasitic influences from the environment.

Now the required accuracy must be specified. In general, the aim is to achieve the highest accuracy possible. At the moment, however, the accuracy which can be achieved by the different methods is hardly known, but we must consider that it is often more difficult to obtain a high accuracy when the films are thinner. For the moment, the aim is a maximum uncertainty in the film thickness measurement of 1 percent for a film thickness in the range of 1 to 10 μm, and 0.01 μm for thinner films.

Also the spatial resolution (i.e. the smallest wavelengths in the profile which can be determined) is important. It should be of the order of 1 μm to enable the detection of the (eventually deformed) roughness texture in the lubricated contact.

The following **requirements** can now be formulated:

1. *The position of the transducer:* The transducer must be fixed on or in the rigid body to avoid too much influence on the mechanical properties of the elastomer. The film profile in the direction of the sliding motion can thus be measured as a function of time, because the rigid surface moves relative to the contact. Also, the surface roughness of the elastomer can then in principle be determined;
2. *The surface finish of the rigid body* must not be changed by the transducer to avoid disturbance of the lubricant film, just at the measurement spot;
3. *The lubricant film must not be disturbed by local decrease of the stiffness of the rigid body*, which can especially occur at higher pressures.
4. *The vertical resolution*, which is one of the factors determining the final accuracy in the film thickness measurement, should be at least 1 percent to cope with the aimed accuracy mentioned above. Then the resolution should be 0.001 μm (1 nm) for a film thickness of about 0.1 μm. However, this can probably hardly be achieved, since film thicknesses up to about 10 μm should also be measured. Therefore, a resolution of about 0.01 μm will be tolerated for the thinner films.
5. *The dimension of the measurement spot* should be of the order of 1 μm to obtain the required spatial resolution.
6. *The maximum allowable response time* of the transducer is determined by the required spatial resolution and by the maximum sliding velocity. At a required spatial resolution of the order of 1 μm and a maximum velocity of about 1 m/s, the maximum allowable response time is of the order of 1 μs;
7. *The physical properties of the elastomer and the lubricant* which are used to measure the film thickness, like the conductance, the electrical permittivity or the index of refraction, must be specified to obtain a reasonable sensitivity, and thus a sufficient accuracy, of the method. The minimum value of the sensitivity, required to meet the aimed accuracy, depends on the accuracy of the measurement equipment which determines the value of the film thickness related physical parameter.

   When improvement of a particular physical property seems to be necessary to cope with the requirements, e.g. by filling the elastomer with conducting particles to obtain the required conductance for an electrical method, the possible changes of other properties, e.g. the mechanical, must be considered.
8. *The pressure and temperature influence* on the measurement must be negligible small or be predictable. Any physical property used for the measurement is

pressure and temperature dependent, one more than the other. When e.g. the pressures are high during the experiment, a method with a small pressure influence (i.e. the physical property used for the measurement must have a low pressure dependence) should be used, or alternatively, both the pressure and its influence must be known quantitatively.

Finally, we can make some remarks on the possibility, that the wanted accuracy or spatial resolution can hardly or not be achieved. Then one can consider modification of the experimental set up to obtain thicker films, which can probably be measured more accurate. The surface roughness height should then also be enlarged to keep the range of the film thickness to roughness height ratio constant. Also, on can consider to apply a regular "model roughness" texture with a rather long wavelength, as was e.g. done by Kaneta and Cameron (1980) (see section 1.2.2, page 13) when the required spatial resolution of about 1 μm can not be obtained. This is, however, not preferred, since the roughness deformation is possibly larger for longer wavelengths than for shorter wavelengths (see section 1.2.3, page 15).

2.3 Choice of the method

Having specified the requirements, a proper method must be chosen for the film thickness measurements. Before we make a choice, the suitability of different methods must be investigated and compared.

In appendix B different methods, possibly suitable for the film thickness measurements, are reviewed and discussed. These methods are:

- Mechanical methods;
- Electrical methods:
  - Resistive;
  - Capacitive (using the elastomer as electrode);
  - Capacitive (using two electrodes on the rigid body);
- Magnetic induction;
- Optical methods:
  - Interferometry;
  - Moiré;
  - Ellipsometry;
  - Focus error detection;

- Absorption;
- Fluorescence;

- Ultrasonic methods.

These methods were compared on their the suitability for film thickness measurements in an elastomer to metal/glass contact in general and for the detection of the real surface roughness in the lubricated contact in particular. The focus error detection was chosen as the most appropriate. Compared with the other methods, its most significant advantage is that a lateral resolution of the order of 1 μm is easily realized. This method will be described in detail and further analysed in chapter 3. Preliminary measurements using this method will be presented in chapter 4 and 5.

CHAPTER 3 DESIGN CRITERIA FOR THE FOCUS ERROR FILM THICKNESS TRANSDUCER

As discussed in chapter 2 and appendix B, focus error detection has been chosen to measure the film thickness and the roughness deformation in the lubricated contact of a rough elastomer and a smooth rigid body. A brief preliminary discussion on the method's suitability has been given in section B4.4.

A focus error system has been developed by Philips Research Laboratories for the reading of optical discs (Bouwhuis and Braat, 1978; Bouwhuis et al., 1987) and nowadays it is widely applied in e.g. compact disc players.

The same system was modified by Philips Research Laboratories to enable non-contacting displacement measurements and Struik and Chang (1987) applied this device for shape and surface roughness measurements. A commercial version is available from Rodenstock.

In this chapter, the present systems for optical disc reading and for displacement, shape and roughness measurements will be described and it will be shown how a focus error system can be applied for film thickness measurements, including analysis of the accuracy.

3.1 The present focus error systems for optical disc reading and for displacement, shape and roughness measurements

Focus error systems are optoelectronic devices, able to determine the position of a surface with respect to the focal point of a lens. In the most simple configuration, three position levels are distinguished (fig. 3.1):

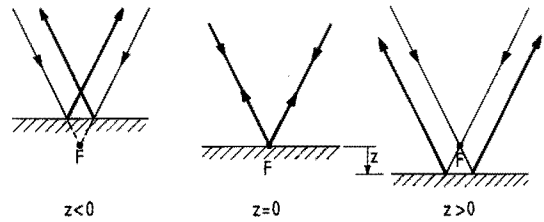
The position of the surface can be:

1. between the lens and the focal point ($z < 0$);
2. in the focal point ($z = 0$);

or 3. beyond the focal point ($z > 0$).

The output of the device is the so-called **focus error signal**, the sign of which changes when the surface moves through the focal point F.

Figure 3.1
Definition of the surface height z relative to the focal point F.



As mentioned in appendix B4.4, several focus error detection systems are available, i.e. the focus error signal can be derived in several different ways. One of these, the so-called pupil obscuration method, is applied in the systems considered here and will therefore be explained below, followed by a description of two types of compact disc transducers and the displacement, shape and roughness sensors derived from it. Finally the performance of these systems will be discussed, considering the dimension of the focus spot, the signals of the system, the measurement range and the dynamic behaviour.

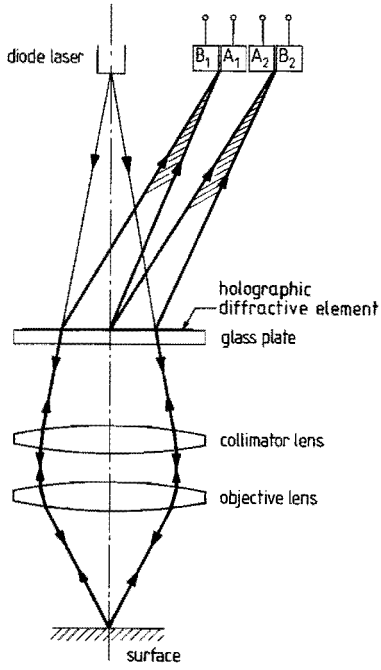
3.1.1 The principle of the pupil obscuration method

The principle of the "pupil obscuration" focusing system is shown in fig. 3.2. The diverging beam from a diode laser is collimated and the collimated beam is focused by the objective lens to a small spot. A more or less specular surface will reflect the beam and the reflected rays are captured by the photodiodes A and B.

When the surface is in the focus of the objective lens (fig. 3.2a), the reflected beam is focused on the boundary of the photodiodes A and B. If we forget the presence of the knife for a moment, both photodiodes will receive the same amount of light and will therefore yield an equal signal. The addition of the knife means that half of the light is blocked and the photodiodes receive only half of the light, but both diodes still receive the same amount of light, as is shown by Bouwhuis et al. (1987 pp. 77-78).

If the surface is out of focus (fig. 3.2b and c) one photodiode is shadowed by the knife and the amount of light received by both photodiodes is therefore different. Which diode receives more light depends on the out of focus direction (i.e. whether the surface is closer to the lens (fig. 3.2b) or further from the lens (fig. 3.2c)). The difference in the signals from both photodiodes can thus be used to determine the position of the surface relative to the focal point.

Figure 3.4
Modified pupil obscuration focusing system: The holographic diffractive system (HDS).



of the two parts of the holographic element yields two first order maxima, one originating from one part and the other originating from the other part. Two pairs of photodiodes are thus used and out-of-focus is detected in a similar way as in the double wedge system.

An optical profilometer, similar to the "double wedge" profilometer of Struik and Chang (1987), presented in section 3.1.2 above, was modified by replacing the double wedge element, the diode laser and the photodiodes by the holographic diffractive element which includes the diode laser and the photodiodes. The wavelength of this new device[3] is slightly different, but the lenses are the same. The optical characteristics are thus:

Diode laser:	λ	=	780 nm
Collimator lens:	f	=	22.5 mm; NA = 0.1
Objective lens:	f	=	4.5 mm; NA = 0.45

(λ being the wavelength; f the focal distance and NA the numerical aperture, see eq. (3.2) on page 30 below).

[3] Referred to as **HDS** (Holographic Diffractive Sensor) throughout this thesis.

An advantage of this system, with two pairs of photodiodes, is the possibility to measure the surface slope (see section 3.1.4.2 page 32), which is important for the film thickness measurement as will be discussed in section 3.2.3 (page 40).

Such a device[2] is used at our institute for shape and surface roughness measurements (Struik and Chang, 1987) and the preliminary experiments in this thesis were performed with the same device. The optical characteristics are:

Diode laser:	λ	=	820 nm
Collimator lens:	f	=	22.5 mm; NA = 0.1
Objective lens:	f	=	4.5 mm; NA = 0.45

(λ being the wavelength; f the focal distance and NA the numerical aperture, see eq. (3.2) on page 30 below).

The window near the objective lens is necessary to obtain a so-called "diffraction limited" focus spot (see section 3.1.4.1 below) in combination with the objective lens. This lens is similar to the one applied in the compact disc transducer, which was specially designed to compensate for the spherical aberration introduced by the 1.2 mm thick protective layer of the compact disc (see section 3.1.4.1 for more details).

3.1.3 The diffractive focus error detection system

The film thickness transducer will be based on a newer focus error detection system, using a holographic diffractive element developed by Philips and Sharp. Compared with the double wedge system, this newer device has the advantage that the positioning of the photodiodes relative to laser diode is rather critical, since both the photodiodes and the laser diode must be positioned in the focal plane of the collimator lens. In the DWS, they are geometrically spoken not in the same plane, making accurate positioning difficult.

In the new device the diode laser and the photodiodes are in the same plane, making the positioning much easier. It uses a holographic diffractive element (in its working comparable to a diffractive grating), which is split into two parts with a slightly different average pitch (fig. 3.4). Part of the returning light will be diffracted by the element and the light of the first order maximum is captured by a pair of photodiodes. The different average pitch

[2] Referred to as **DWS** (Double Wedge Sensor) throughout this thesis.

3.1.2 The double wedge focus error detection system

A focus error system, in which the knife is replaced by a double wedge, is shown in fig. 3.3 (Bouwhuis and Braat, 1978; Bouwhuis et al., 1987 pp. 75-78; Struik and Chang, 1987). Now, two images of the focus spot are produced when the surface is in focus: Both on the boundary of one pair of photodiodes. When the surface is not in focus, both the outer diodes (B_1 and B_2) or both the inner diodes (A_1 and A_2) are shadowed and the out-of-focus can be determined by the signal difference $(A_1 + A_2) - (B_1 + B_2)$ from the inner and outer diodes[1]. A servo controller can be used to move the objective lens to a position where the signal difference is zero, focusing the lens onto the surface (see fig. 3.3). Measurement of the lens displacement then yields the vertical surface displacement. The shape and surface roughness can be determined by moving the surface in a horizontal direction.

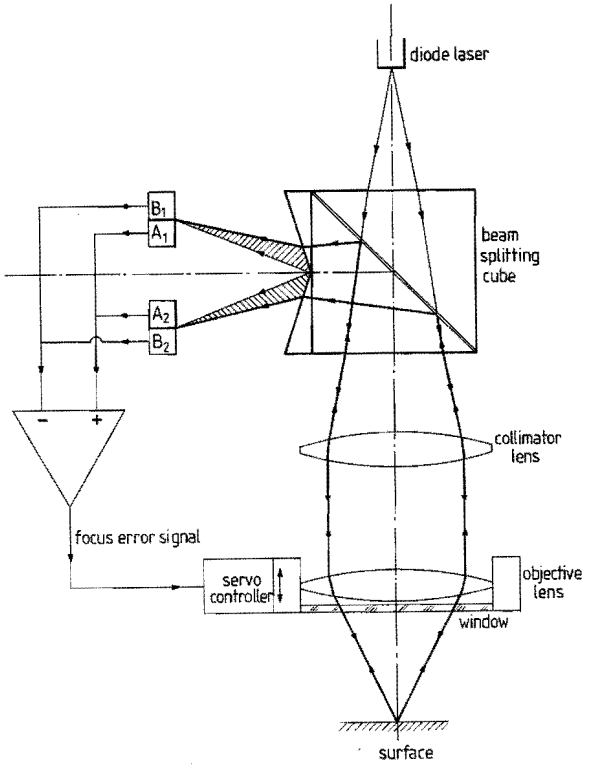


Figure 3.3
Modified pupil obscuration method: The double wedge system (DWS).

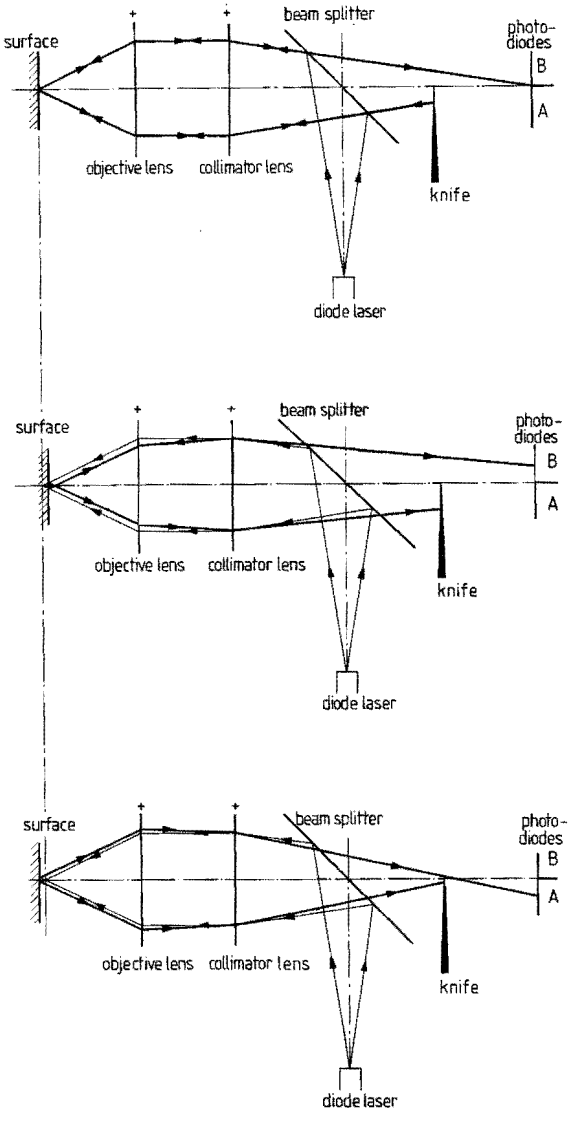
[1] A_i and B_i are the signals from photodiode A_i and B_i respectively.

Figure 3.2
Principles of the pupil obscuration focus error system:

a. Surface in focus.
(A and B receive the same amount of light)

b. Surface in front of the focus.
(A receives less light than B)

c. Surface beyond the focus.
(A receives more light than B)



Focus error systems, derived from this principle, have been constructed and will be presented below. First an earlier design, at present in use as roughness sensor and also used for some tests in this thesis, will be described, followed by a newer design which will be used for the film thickness transducer. Finally, the performance of the focus error systems will be discussed.

In this thesis some tests are performed with this sensor and the final film thickness transducer (described in section 4.2) is based on this type of sensor

3.1.4 The performance of the focus error systems

In this section, the performance of the focus error detection system will be described, considering the following items:

1. The focus spot dimension.
2. The signals;
3. The measurement range;
4. The dynamic range;

Measurements with the focus error device can be performed in two ways:

- In the "**closed loop**" mode: The objective lens is focused onto the surface and continuous measurement of the lens position yields the height variations, as mentioned in section 3.1.2 (page 26);
- In the "**open loop**" mode: The lens is fixed and the height variations are directly derived from measurement of the focus error signal.

The performance of the focus error device is different for the two different modes and both will be considered in the performance discussion presented below.

3.1.4.1 The focus spot dimension

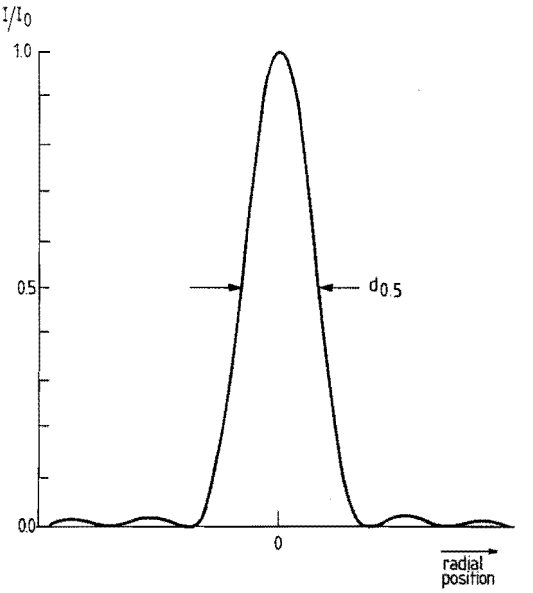
The focus spot dimension is an important factor, since it determines the spatial resolution (see point 5 in section 2.2 page 20). The smallest spot which can be obtained is the so-called **diffraction limited spot**, the size of which is determined by diffraction only. The irradiance[4] distribution or "Airy pattern" of such a spot is shown in fig. 3.5 and its width can be characterized by the **fifty-percent-irradiance width** $d_{0.5}$, which is derived in appendix F1 and reads

$$d_{0.5} \approx \frac{\lambda}{2\,NA} \tag{3.1}$$

in which λ is the wavelength [m] and NA the numerical aperture [–] of the objective lens:

[4] *Irradiance* is the amount of light energy per unit area per unit time (in the past often called *intensity*)

Figure 3.5
Relative irradiance distribution I/I_0 of the diffraction limited focus spot (Airy pattern).



(I_0 = Irradiance maximum;
$d_{0.5}$ = Fifty-percent-irradiance width)

$$NA = n \sin\varphi \qquad (3.2)$$

(n is the index of refraction [–] at the right hand side of the lens, fig. 3.6, and φ is the half top angle of the light cone).

To give an example, the fifty-percent-irradiance width $d_{0.5}$ of the DWS (λ = 0.82 μm; NA = 0.45: see section 3.1.2 page 26) is 0.82 μm.

Two requirements must be fulfilled to obtain the diffraction limited spot:
1. The irradiance distribution must be uniform over the beam width;
2. The optical system must be free of aberrations.

The first requirement is in principle not fulfilled using diode lasers. However, the irradiance distribution appears to be sufficiently uniform when the numerical aperture of the collimator lens is 0.1 or smaller, using the sensors described in section 3.1.2 and 3.1.3 above (Bouwhuis et al., 1987 chapter 2; see also appendix F2).

The second requirement means, that high quality optics ("diffraction limited lenses") must be used, yielding a focused beam with a perfectly spherical wavefront (fig. 3.6). However, when the surface is scanned through a window (fig.

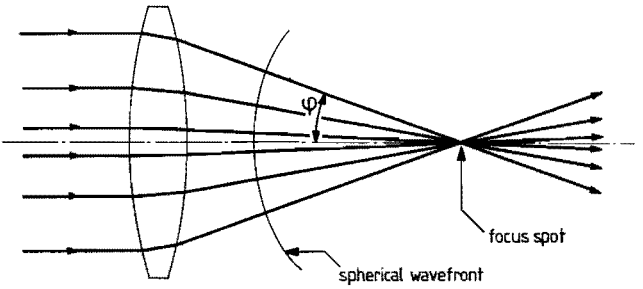


Figure 3.6
Aberration free optical system with perfect spherical wavefront.

3.7) the light is refracted on the window surface and the wavefront is not spherical anymore. Consequently, rays of different angle of incidence have a different focal point. The light is thus spread over a larger area and the spot diameter is increased.

Such a window is e.g. present in the compact disc configuration, where the disc surface is scanned through a 1.2 mm thick protective layer. Measuring the lubricant film thickness, the elastomeric surface must also be scanned through a window (see section 4.2).

Reduction of the spherical aberration is of the greatest importance to obtain the smallest (diffraction limited) spot for the required high spatial resolution. The spherical aberration caused by the window is given by Bouwhuis et al. (1987 p. 30ff.) and it is derived in appendix F3 that the following condition must be fulfilled to obtain a diffraction limited spot

$$\frac{n^2 - 1}{8\ n^3}\ t\ (NA)^4 \ \leq\ 0.95\ \lambda \tag{3.3}$$

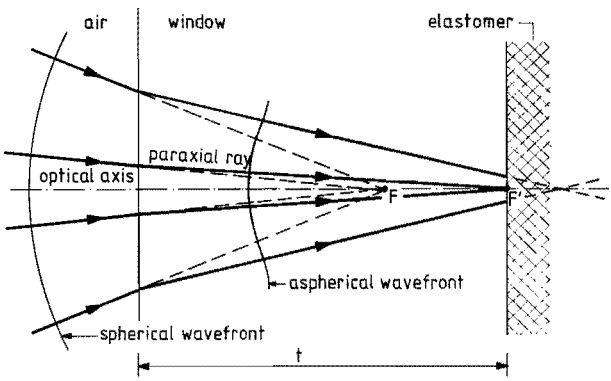


Figure 3.7
Spherical aberration caused by refraction on the window surface.

in which:

n	=	Index of refraction of the window	[–]
t	=	Window thickness (fig. 3.7)	[m]
NA	=	Numerical aperture (eq. 3.2) of the objective lens	[–]
λ	=	Wavelength of the radiation	[m]

In the compact disc configuration, the aberration is characterized by

$$\frac{n^2 - 1}{8\, n^3}\, t\, (NA)^4 = 2.3\ \mu m$$

since the protective layer of the disc has an index of refraction n = 1.56 and a thickness t = 1.2 mm, while the numerical aperture NA = 0.45 (Bouwhuis et al., 1987 p. 44). This is clearly too large, since the wavelength λ of currently available diode lasers is in the range of 0.75 to 0.83 μm. Therefore the objective lens was designed to have the same spherical aberration, but with opposite sign.

The same objective lens is used in the shape and roughness sensor DWS (see section 3.1.2 page 26) and the protective layer is not present in these measurements. Therefore a 1.2 mm thick window is attached near the objective lens to keep the focus spot diffraction limited.

In this thesis a number of experiments are performed with the DWS, using a 1.2 mm glass plate on the scanned surface (see e.g. chapter 5 and appendix G1.3 and G2.3). In these experiments the window near the objective lens was removed.

3.1.4.2 The signals

Two signals, which are important for our purpose, are derived from the **photodiode signals** A_1, A_2, B_1 and B_2 (originating from the photodiodes A_1, A_2, B_1 and B_2 respectively, see fig. 3.3 and 3.4 page 26 and 28). One is the **focus error signal**, which can be used to position the lens, and the other the so-called **radial error signal**, which can be used to eliminate influence of the surface slopes on the focus error signal, as will be discussed below.

The signals were measured as a function of the axial surface position, which is defined as the height z of the focal point above the surface (see fig. 3.1 page

24). These measurements will be shown below, where the signals will be discussed in more detail.

The photodiode signals

The photodiode signals A_1, A_2, B_1 and B_2 are the first to be considered, since the others are derived from them. When the scanned surface is perpendicular to the optical axis, A_1 is (in theory) equal to A_2 and B_1 is equal to B_2. Their sums A (= A_1 + A_2) and B (= B_1 + B_2) are shown in fig. 3.8.

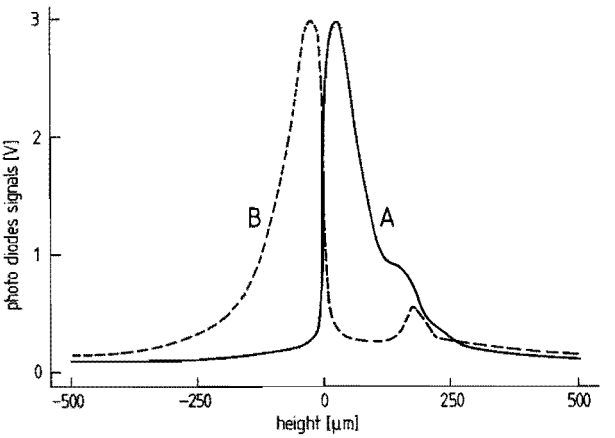


Figure 3.8
Photodiode signals A and B versus surface height z relative to the focal point.
(Measured with the DWS on a silicon surface)

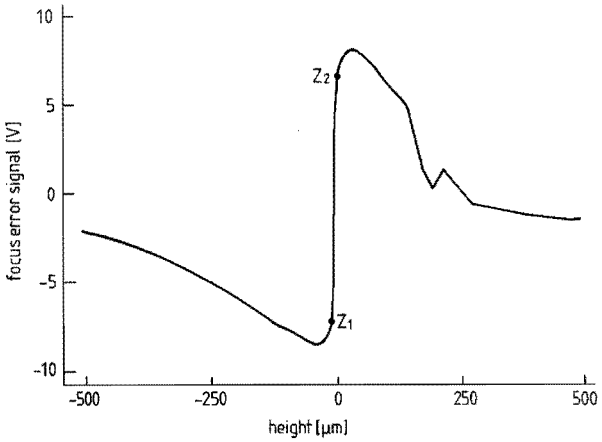
The focus error signal

The position of the surface, relative to the focal point of the objective lens, is determined by the difference $(A_1 - B_1) + (A_2 - B_2)$. Division by the sum $(A_1 + B_1) + (A_2 + B_2)$ yields the focus error signal *fes*

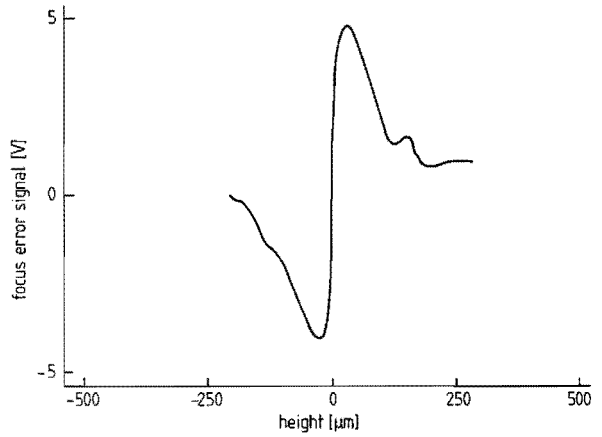
$$fes = \frac{(A_1 - B_1) + (A_2 - B_2)}{(A_1 + B_1) + (A_2 + B_2)} \tag{3.4}$$

The division with the sum of the signals is done to make the signal independent of the surface slopes in one direction (see appendix E1) and more or less independent of the surface reflectance. Fig. 3.9 e.g. shows that the maximum in the focus error signal, measured on the glass, is only a factor 2 lower than on the silicon surface, in spite of the large difference in reflectance (about 4 percent for the glass and 70 percent or more for the metal).

Figure 3.9
Focus error signal versus surface height z relative to the focal point.
(Measured with the DWS)



a. measured on a silicon surface



b. measured on glass

The focus error signal is influenced by the slopes in the surface (see appendix E). This influence originates from the fact that the beam reflected on a skew surface is not symmetrical with respect to the optical axis, as shown in fig. 3.10.

In appendix E1 is shown that this influence is negligible for slopes in one direction, but the influence is significant for slopes in the other direction (fig. 3.11). The different influence of the slopes in the two different directions is caused by the asymmetric positioning of the photodiodes.

In the *closed loop mode* (see page 29) the measurements are not influenced by the slope influence, since the focus error signal is 0 when the surface is in focus, regardless of the slope. If, however, the measurements are performed in the *open*

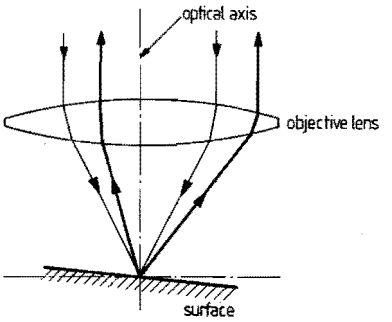


Figure 3.10
Reflected light beam at a surface slope.

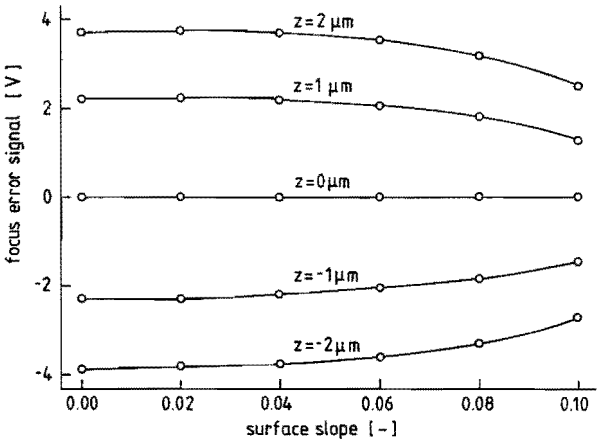


Figure 3.11
Influence of the surface slopes on the focus error signal.
(curves derived from measurements with the HDS (page 27), presented in appendix E1)

loop mode, the slope influence is significant, since the surface height is derived directly from the focus error signal. This slope influence can be eliminated by simultaneous measurement of the radial error signal, which will now be discussed.

The radial error signal

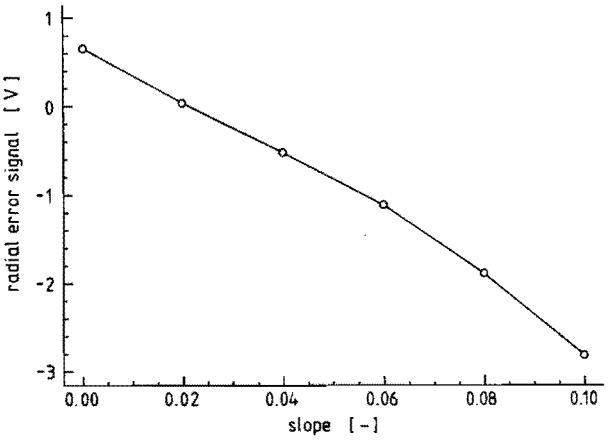
In the compact disc reading system, a radial error signal is constructed to detect whether the focus spot is in the centre of a track or at a radial distance from the track centre (see Bouwhuis et al., 1987 p. 70-75 and p. 85). The essence is that the two diode pairs (A_1B_1 and A_2B_2) only receive the same amount of light, when the spot is in the track centre. The radial error signal *res* is thus

$$res = \frac{(A_1 + B_1) - (A_2 + B_2)}{(A_1 + B_1) + (A_2 + B_2)} \tag{3.5}$$

A servo controller, which positions the focus error device in the radial direction, is used to keep the radial error signal zero, which means, that the focus spot is kept in the centre of a track. In this way, the transducer is able to follow a track and will not jump to a neighbouring track.

The radial error signal can also be used to measure the surface slopes. It appears to be hardly influenced by the distance between the surface and the focal point (see appendix E2). Therefore the local surface slope can be directly derived from measurement of the radial error signal (shown in fig. 3.12), enabling direct compensation for the slope influence on the focus error signal when measurements are performed in the open loop mode.

Figure 3.12
The radial error signal.
(curve derived from measurements with the HDS (page 27), presented in appendix E2)



3.1.4.3 The measurement range of focus error systems

Considering the measurement range, we must distinguish between the closed loop mode and the open loop mode.

In the *closed loop mode* the objective lens is continuously focused onto the surface. The measurement range then depends on the maximum displacement of the objective lens, which is e.g. 1 mm for the DWS and the HDS (page 26 and 27).

In the *open loop mode* the measurement range is determined by the focus error signal (shown in fig. 3.9 on page 34). Accurate measurements can only be performed around the in-focus position (i.e. around zero height between the points Z_1 and Z_2 shown in fig. 3.9), where the curve is very steep. Using the DWS (page 26) the measurement range is about -5 µm to +5 µm.

3.1.4.4 The dynamic range of focus error systems

Considering the dynamic range, we must also distinguish between the closed loop mode and the open loop mode.

In the *closed loop mode*, the dynamic range is limited by the eigenfrequency of the objective lens and its suspension. The maximum measuring frequency of the DWS and the HDS (page 26 and 27) is e.g. 600 Hz. Higher frequencies will hardly be possible in the closed loop mode, because eigenfrequencies higher than some kHz can hardly be realized because of the mechanical inertia.

Measuring in the *open loop mode*, with stationary objective lens, enables extension of the dynamic range, since the mechanical inertia is not a factor. Then the limiting factor is the electronic amplifier and frequencies of 1 MHz can be reached, using e.g. the HDS (page 27) (see Sharp, 1988).

3.2 Film thickness measurent by means of focus error detection

Up till now, the principles of focus error detection and its performance as displacement, shape and surface roughness sensor have been discussed. In the following we will pay attention to its use for measurement of the film thickness and the surface roughness deformation in lubricated contacts.

Film thickness measurement by means of focus error detection requires the contact area to be optically accessible. Therefore, a window is needed in the rigid body (fig. 3.13), enabling the measurement of the film profile as well as the the seal's surface roughness (requirement no. 1, page 20).

Measurement of the film profile means in essence detection of the local height position of the elastomeric surface and seems therefore quite similar to a normal surface roughness measurement with the focus error detection system presented in section 3.1. However, there are differences concerning the dynamic range and the addition of the window and a lubricant film. The influence of these differences and other factors will be discussed below, after the requirements of section 2.2 (page 20) have been repeated and, where possible, reformulated for this specific method.

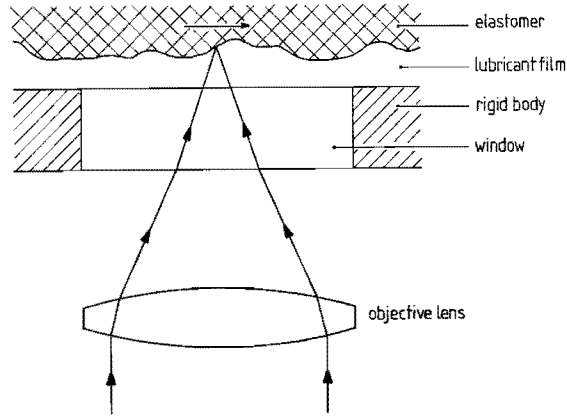


Figure 3.13 Arrangement for measurement of the lubricant film thickness.

3.2.1 The requirements for the focus error detection system

The following requirements are derived directly from section 2.2:

1. *The transducer position* is in the rigid body.
2. *The surface finish of the rigid body* must not be changed significantly by the transducer.
3. *The lubricant film must not be disturbed* by local decrease of the stiffness of the rigid body.
4. *The accuracy (and thus the vertical resolution)* must be of the order of 0.01 μm at a film thickness in the range of 0.1 to 1 μm and about 1 percent at thicker films (up to about 10 μm).
5. *The diameter of the measurement spot* must be of the order of 1 μm.
6. *The response time* must be of the order of 1 μs or less.
7. *The reflection on the elastomer to lubricant interface*, which is necessary to use the focus error detection system, must be high enough.
8. *The pressure influence* on the measurement must be small or predictable.
9. *The temperature influence* on the measurement must also be small or predictable.

The requirements 1, 2 and 3 are automatically or rather easily fulfilled:

- Requirement no. 1 means that a window must be present in the rigid body, as already mentioned above and shown in fig. 3.13.
- Concerning requirement no. 2, special attention must be given to the fabrication of the window in the rigid body. After grounding of the surface, the window surface can be in a lower position than the surrounding surface when

the glass window is mounted in a steel body, due to difference in stiffness and hardness of glass and steel. This eventual height difference must be very small (less than e.g. 0.01 μm) to avoid significant influence on the lubrication of the elastomeric body. An alternative, which will avoid this problem, is to make the whole rigid body, including the window, of the same transparent material.

- Requirement no. 3 is also fulfilled, since the transducer, which is mounted beyond a window, causes no sudden change in stiffness at the measurement spot as can e.g. be caused by an electrode mounted in the rigid body for use of an electrical method (see appendix B2.1).

The requirements no. 4 and 5 also seem to be fulfilled already:

- The accuracy of the DWS (described in section 3.1.2 page 26) is 0.01 μm or better according to Struik and Chang (1987).
- In section 3.1.4.1 (page 30) was derived that the dimension of the (diffraction limited) focus spot is of the order of 1 μm.

However, we must consider that the introduction of the window and the lubricant film possibly affects the accuracy and the lateral resolution. *Reflection on the window surfaces*, e.g., may introduce errors (see section 3.2.5 below). Also, *the slopes of the surface* (inherent to rough surfaces) may cause improper working of the system (see section 3.2.3).

We must also account for influence of *spherical aberration* on the focus spot dimension due to the light refraction on the window surfaces (see section 3.2.4 below).

These points need more discussion, as well as the other requirements (6 to 9), to make clear whether and when these requirements are fulfilled.

This discussion will start with the dynamics of the system (requirement no. 6), followed by the influence of the surface slopes, the spherical aberration caused by the window, the influence of reflection on the window surfaces, the minimum required reflectance on the elastomer to lubricant interface (requirement no. 7) and the pressure and temperature influence on the measurement (requirements no. 8 and 9).

3.2.2 The dynamics of the system

The maximum response time must be of the order of 1 μs, as mentioned in section 3.2.1 above (requirement no. 6). The required dynamic range is therefore 1 MHz. This implies, that the measurements must be performed in the "open loop mode" (see section 3.1.4.4 page 37) and the objective lens will therefore be fixed in the rigid body.

3.2.3 Influence of the surface slopes

The use of a fixed objective lens (see section 3.2.2 above) has the disadvantage that the focus error signal, from which the film thickness will be derived, is influenced by the slopes in the surface (see section 3.1.4.2 page 34). These slopes are inherent to rough surfaces and can not be avoided since the roughness behaviour in the lubricated contact is to be investigated. Therefore, simultaneous measurement of the radial error signal is necessary to eliminate the slope influence (section 3.1.4.2 page 36).

3.2.4 Spherical aberration caused by the window

As already mentioned in section 3.1.4.1 (page 29), a diffraction limited focus spot (which is free of aberrations) has the smallest possible dimension and is therefore preferred, since the lateral resolution is then the highest possible. However, the addition of the window introduces spherical aberration.

Using a 1.2 mm thick window, the aberration can be eliminated by use of the compact disc objective lens, which is also used in the DWS sensor. However, a thicker window is preferred because of the contact load and consequently the compact disc lens can not be used. Otherwise, development of a new special lens, which compensates the aberration of a thicker window, is not considered because of the complexity of the design process and because of the high production costs of such lenses (having a difficult geometry) for small series. It is therefore decided to use a standard lens, which does not compensate for the spherical aberration caused by the window. As a consequence, we must pay special attention to the reduction of the spherical aberration.

The possibilities to reduce the spherical aberration can be derived from eq. (3.3) on page 31. Considering that the window thickness t is prescribed by the contact pressures, the influence of the aberration can be reduced by use of:

- light with a long wavelength λ
- a glass with a suitable index of refraction n (i.e. such a value of n, that $(n^2 - 1)/(8n^3)$ is small);
- an objective lens with a low numerical aperture NA.

Influence of the wavelength

As expressed by eq. (3.3) a larger aberration can be tolerated when the wavelength is longer. Otherwise, the diameter of a diffraction limited spot is proportional to the wavelength and this limits the wavelength to obtain the required high spatial resolution. Further, use of light with a significantly larger wavelength has the practical objection that a totally new focus error device should be specially designed for our application. This is not considered, because of the complexity of the system.

Influence of the index of refraction

The influence of the index of refraction is shown in fig. 3.14. It appears that the index of refraction can not be used to reduce the aberration significantly, since all glasses and other transparent solid materials have an index of refraction of more than about 1.4.

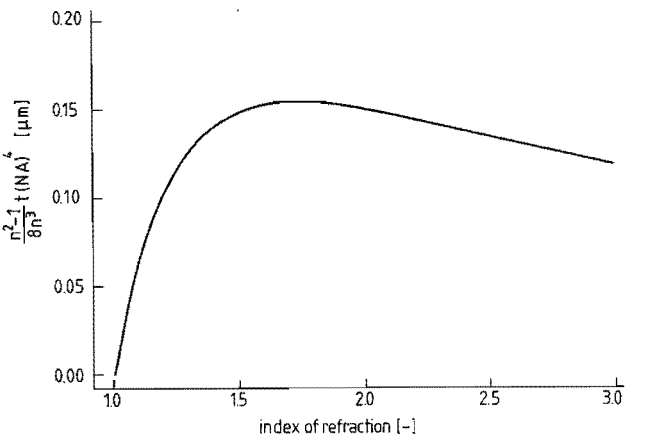


Figure 3.14
Influence of the index of refraction on the spherical aberration given by eq. (3.3), page 31.
(t = 2 mm; NA = 0.2)

Influence of the numerical aperture

Finally, reduction of the numerical aperture appears to be the only practical way to reduce the influence of spherical aberration. In fact, this is a very effective option since the spherical aberration is proportional to the fourth power of the numerical aperture and a small reduction of the numerical aperture yields thus a large reduction of the aberration, while the dimension of the diffraction limited spot is hardly inceased. Halving the numerical aperture, e.g., reduces the influence of the spherical aberration 16 times, while the diffraction limited spot size only increases by a factor 2.

For an index of refraction in the range of 1.4 to 1.7 we find

$$0.35 \leq \frac{n^2 - 1}{n^3} \leq 0.38$$

and considering that the wavelength of diode lasers is between 0.75 and 0.83 μm, eq. (3.3) yields the criterion for the maximum allowable numerical aperture of the objective lens

$$NA \leq 2t^{-1/4} \qquad (t \text{ in } \mu m) \tag{3.6}$$

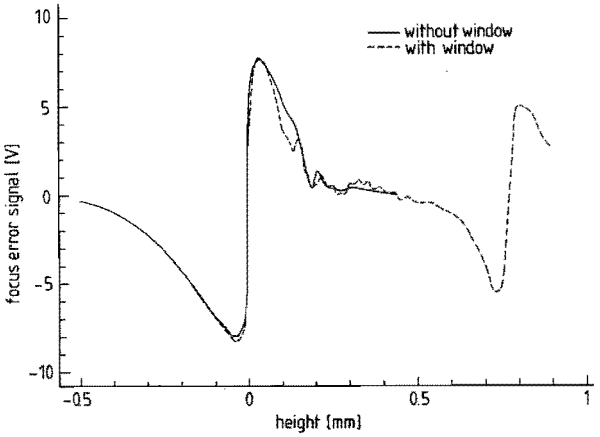
3.2.5 Influence of reflection on the window surfaces

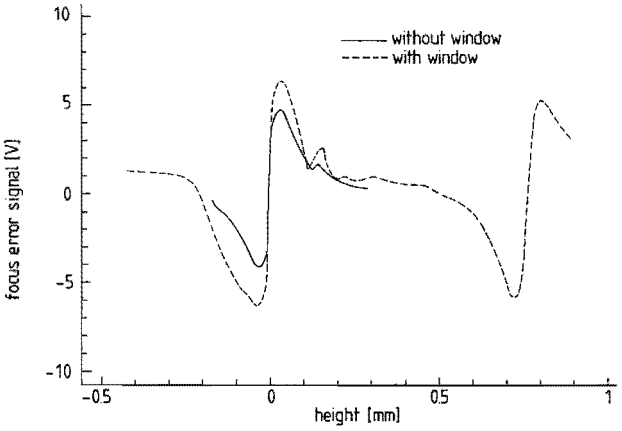
Reflection generally occurs on both sides of the window and contributes to the focus error signal, possibly leading to serious errors in the measurements. Therefore, its influence must be investigated. This was firstly performed by measuring the focus error signal as a function of the focus height above the surface on a metallic test surface and on a glass test surface (having essentially equal optical properties as elastomers), both with and without a 1.2 mm thick window (glass plate) on it. The set up for these tests is described in appendix C1 and the curves of these measurements are shown in fig. 3.15, in which the definition of the surface height is as given in fig. 3.1 (page 24), i.e. the test surface is in focus at height $z = 0$ (see also fig. 3.16, page 44).

In fig. 3.15, the lower window surface is very close to the test surface (which is at height $z = 0$), while we find the upper surface in the measurement at a height of nearly 0.8 mm, as is expected for a 1.2 mm thick glass plate with an index of refraction of about 1.5.

Figure 3.15
Influence on the focus error signal of reflection on the window surfaces. (Measurements performed with the DWS, section 3.1.2).

a. On a silicon surface





b. On a glass surface

The influence of the reflection on both window surfaces will be discussed in the following. First, we will consider the influence of the reflections on the focus error signal and next, the influence on the surface roughness measurement will be studied.

Upper surface influence

Comparison of the measurement with the window and the measurement without the window (fig. 3.15) shows that the reflection on the upper glass surface has no influence on the focus error signal around the in-focus position of the surface (height $z \approx 0$), since both curves (with and without window) almost coincide at a height in the range of roughly 0.2 to 0.4 mm. The upper window surface,

therefore, does not influence the roughness measurement, as is proved in appendix G2.3.

Lower surface influence

The lower surface of the window appears to influence the focus error signal in the neighbourhood of the in-focus position of the test surface (i.e. at zero height, see fig. 3.15). This may be expected, since the lower window surface is close to the test surface.

In the case of the **silicon test surface** (fig. 3.15a), the difference between the two curves is small around $z = 0$. This can be explained by the high reflectance of the silicon test surface, compared with the low reflectance of the window surface.

In the case of the **glass test surface** (fig. 3.15b), the influence of reflection on the lower window surface is much larger around $z = 0$, due to a nearly equal reflectance of both the test and the window surface. In appendix G the influence of the lower window surface is studied in more detail and some results are presented for the case of a test surface with an equal reflectance. This case is of special interest, in comparison with the case of a highly reflecting test surface, since elastomers have in general a reflectance in the same range as glass[5].

In appendix G1 is derived that the focus error signal is zero for a position of the focal point somewhere between the lower window surface and the test surface (see fig. 3.16). Therefore, the measured height variations are is expected to be

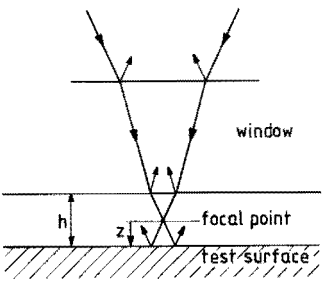


Figure 3.16
Gap between window and test surface.
(*h* is the gap height; *z* is the distance between the focal point and the test surface)

[5] The indices of refraction are almost equal and the reflectance on the glass to air and on the elastomer to air interface are therefore nearly equal (see appendix D).

smaller than the real height variations. However, some experiments, presented in appendix G2.3, yielded contradictory results: When the profile or roughness height is scanned through the glass plate, the measured profile (or roughness) height appears to be larger than the real height and the shape of the profile is disturbed. The origin of these effects is not understood and need thus more investigation.

To prevent such disturbances, it is recommended to use a liquid in the contact between the test and the window surface. Then the accuracy of the measurement can be significantly better, since reflection does not occur on the window surface when the indices of refraction of the liquid and of the glass are equal (see appendix D).

3.2.6 The minimum required reflectance on the elastomer to lubricant interface

An important factor for the accuracy is the reflectance on the elastomer to lubricant interface, since the working of the system is based on this reflectance. If there is no reflection on the elastomeric surface, the film thickness and the roughness deformation can not be measured. Therefore a minimum reflectance on the lubricant to elastomer interface is required. Consideration of the value of the minimum required reflectance is important, since this reflectance on the lubricant to elastomer interface is typically low (see appendix D2).

The value of the minimum required reflectance mainly depends on the laser power and on the signal to noise ratio (see appendix H). If the reflection is low, the the photodiode signals are low and the noise is then relatively large. The measurement can then in principle be improved by increasing the laser power, but this is limited by the consequently shorter laser life. Another solution is to apply the elastomeric surface with a thin metallic coating. Such a coating must of coarse not influence the mechanical properties of the elastomeric specimen, nor the roughness texture.

Considering the present electronics, improvement of the signal to noise ratio is possible and therefore required (appendix H).

In this thesis these factors have not been studied elaborately yet, since the value of the reflectance on the lubricant to elastomer interface is not easily determined (see appendix D2). Nevertheless some measurements have been performed to test the working of the system (see section 4.3 and appendix L).

3.2.7 Influence of the contact pressure

The contact pressure influences the measurement in three ways (see appendix I3):
- By change in the index of refraction of the lubricant;
- By change in the index of refraction of the window;
- By bending and impression of the window.

For pressures up to 50 MPa, negligence of the pressure influence leads to the following errors in the film thickness measurement:
- The change in the lubricant's index of refraction yields an underestimation of up to 1 percent;
- The change in the window's index of refraction may cause an underestimation of the order of 0.1 μm;
- The bending and the impression of the window may cause an underestimation of the order of 1 μm.

The pressure influence on the film thickness measurement is far from negligible, since the expected film thickness is in the range from 0.1 to 10 μm and should be measured with an accuracy of 1 percent (or 0.01 μm for films thinner than 1 μm). Therefore the pressure influence must be accounted for. Especially the influence caused by the change in the window's index of refraction and by the bending and impression of the window are severe.

We must consider the window dimensions here, since the influence of the bending and the impression is the smallest when the window radius is small relative to its thickness, i.e. the radius to thickness ratio R_w/t (see fig. 3.17) must be small. At larger values of R_w/t, the impression is increased and the bending becomes significant. Besides, the stresses in the window are larger when the

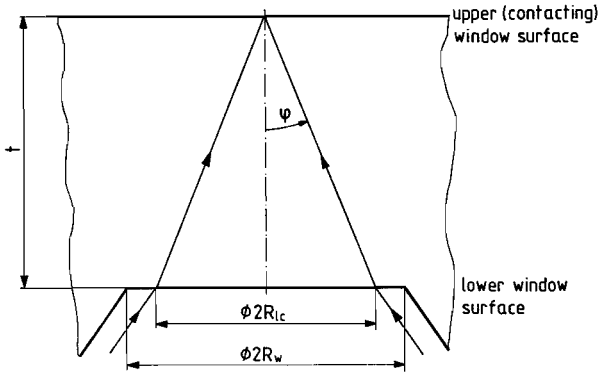


Figure 3.17
Restriction of the top angle φ of the light cone (and thus of the numerical aperture) by the window dimensions.

radius to thickness ratio R_w/t is larger. R_w/t should therefore be as small as possible.

In appendix I3.2 is derived that the bending is negligible for a value of R_w/t smaller than about 0.2, when the maximum contact pressure is about 50 MPa. In the preliminary measurements presented in chapter 4, the maximum pressure will be 5 MPa. Then R_w/t should be smaller than 0.25.

When the bending is negligible, the impression is the smallest possible (of the order of 1 μm) for a given pressure distribution and almost independent of the window's radius and thickness. Then the impression can in principle be calculated using the elastic half space approach as e.g. presented by Johnson (1985 chapter 3).

A practical consequence of this limit in the window's radius to thickness ratio is that the numerical aperture of the objective lens is limited, since the light cone of the transducer may not exceed the window boundaries. Considering fig. 3.17 the maximum value of the numerical aperture can be derived. The criterion is that the radius R_{lc} of the light cone on the lower window surface must be smaller than the window radius R_w.

Using the definition of the numerical aperture NA

$$NA = n \sin\varphi$$

we find for small φ

$$\sin\varphi \approx \tan\varphi = \frac{R_{lc}}{t}$$

The maximum value for the numerical aperture is therefore given by

$$NA < n \frac{R_w}{t}$$

which yields

$$NA < 0.3$$

for $R_w/t \approx 0.2$ and $n \approx 1.5$

In general, the total pressure influence on the film thickness measurement is not easily quantified theoretically. Then it must be determined experimentally as discussed in appendix I3.3.

3.2.8 Influence of the temperature

In this section the temperature influence on the film thickness measurement will be discussed. We will only consider the temperature rise of the construction due to frictional heat dissipation in the contact area. Eventual variation in the environmental temperature will have a small influence on the measurement, since these variations are small. They can be compensated by measuring the environmental temperature and by calibration at different environmental temperatures, if necessary. The heat dissipation in the contact area, however, may cause relatively large temperature differences in the neighbourhood of the contact area, which influence the film thickness measurements significantly, but which are not easily determined.

The temperature influences the measurement in four ways (see appendix I4):
- By change in the index of refraction of the lubricant;
- By change in the index of refraction of the window;
- By thermal expansion of the construction;
- By change in the focal distance of the objective lens.

For temperatures up to 200 °C, negligence of the temperature influence yields the following errors in the film thickness measurement:
- The change in the lubricant's index of refraction yields an overestimation of up to 5 percent;
- The change in the window's index of refraction may cause an overestimation of the order of 1 μm;
- The thermal expansion of the construction may cause an overestimation of the order of 0.1 μm;
- The change in the focal distance of the objective lens may cause an underestimation of the order of 0.1 μm.

The temperature influence on the film thickness measurement is in general far from negligible when contact temperatures up to 200 °C occur, since the expected film thickness is in the range from 0.1 to 10 μm and should be measured with an accuracy of 1 percent (or 0.01 μm for films thinner than 1 μm). Therefore the temperature influence must be accounted for.

Calculation of the temperature influence is complicated by the fact that the temperature gradients are relatively large in the neighbourhood of the contact. Measurement of the temperature at some points in the rigid body near the contact

area and near the transducer is therefore recommended to enable this compensation, but a more elaborate study of this matter is necessary.

3.3 Conclusions

Focus error detection can be used to measure the film thickness and the (eventual deformed) surface roughness in the contact of lubricated elastomers, provided that the following conditions are satisfied:

1. *The objective lens must be fixed* in the rigid body to obtain the required measurement speed (response time about 1 μs). The film thickness will then be derived from the measured focus error signal (section 3.2.2 page 40);
2. Besides the focus error signal, *simultaneous measurement of the radial error signal* is needed to eliminate influence of the surface slopes on the film thickness measurement (section 3.2.3 page 40);
3. *The numerical aperture of the objective lens must be limited* to prevent significant influence of the spherical aberration introduced by the window through which the elastomeric surface is scanned (section 3.2.4 page 40ff.).

   The following criterion for the maximum value of the numerical aperture *NA* was derived

   $$NA \leq 2\, t^{-1/4}$$

   (t is the window thickness in μm)
4. *The indices of refraction of the window and of the lubricant are preferably equal* to avoid influence of the reflection on the window to lubricant interface (section 3.2.5 page 42ff.);
5. *The indices of refraction of the lubricant and of the elastomer must be different* to obtain reflection on the lubricant to elastomer interface. This difference must be large enough to assure a minimum reflectance at which the required accuracy can be obtained. A complication is, however, that the index of refraction of elastomers, and especially of opaque looking materials like polyurethane, is not easily determined. Whether the reflectance on the lubricant to elastomer interface is high enough must therefore be tested directly. This matter needs more investigation and improvement of the signal to noise ratio of the system will be necessary (section 3.2.6 page 45);
6. *The window thickness must be at least 5 times its radius* to prevent bending of the window. (This limits the numerical aperture *NA* of the objective lens to 0.3). Further *the pressure influence on the measurement must be determined theoretically or experimentally* (section 3.2.7 page 46);

7. *Measurement of the temperature at some points in the neighbourhood of the contact and the transducer is in general recommended* to enable compensation for the temperature influence on the film thickness measurement, but needs more investigation (section 3.2.8 page 48).

Concerning point 6 and point 7, more investigation is needed to quantify the contact pressure and temperature influence on the measurements. In the measurements presented in chapter 4, the contact pressure and temperature will be restricted by appropriate choice of the running conditions (contact load, velocity and lubricant viscosity) to keep the pressure and temperature influence small.

Further we must consider the general requirement for focus error devices using diode lasers:

8. *The numerical aperture of the collimator lens must be 0.1* (section 3.1.4.1, page 30, and appendix F2). A larger value yields a focus spot which is not diffraction limited and consequently the lateral resolution will be lower. Otherwise, a smaller numerical aperture would yield an unnecessarily large loss of light.

Finally the general requirement for measurements in the contact area (like the film thickness measurements) that the surface finish of the rigid body must not influenced by the transducer, must be satisfied, i.e.

9. *The manufacturing of the window in the rigid body may not lead to an irregularity ("step") in the surface finish of the rigid body* (point 2 in section 3.2.1, page 38).

CHAPTER 4 FILM THICKNESS MEASUREMENTS

This chapter deals with the lubricant film thickness measurement in an elastomer to glass contact, using the focus error detection method described in chapter 3. As discussed in section 1.3 (page 16) these measurements will not be performed on reciprocating seals, because of the complex contact problem of such seals. Instead we will use a simplified set up in which the tribological process is in essence similar to that of reciprocating seals, but which avoids the extra complexity.

First this set up and the test rig used for the measurements will be presented. Next the specimen used for the film thickness measurements will be introduced and the expected film thickness will be derived for different operating conditions. Then the film thickness transducer (designed according to the specifications mentioned in section 3.3) will be described in more detail, including its tests. Finally some preliminary film thickness measurements will be presented.

4.1 The test rig and the elastomeric specimen

4.1.1 Test rig

The test rig on which the preliminary film thickness measurements are performed is shown in fig. 4.1. It uses the computer controlled hydraulic power unit and the linear motion roller bearing of the reciprocating seal test rig, presented by Kanters (1990 section 2.2 pp. 21ff.; 1991), from which the rod and the seal

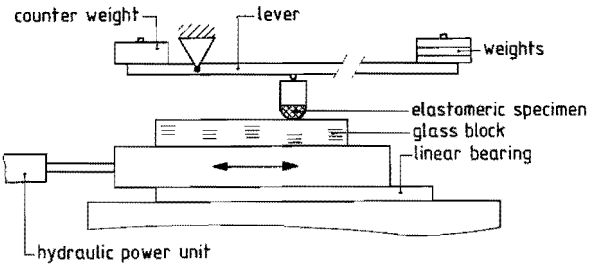


Figure 4.1
Modified test rig for the preliminary measurement of the film thickness and the roughness deformation of lubricated elastomers under sliding motion.

The film thickness transducer is mounted in the glass block as shown in fig. 4.6 (page 55).

housing are removed. In their place a glass block, with the film thickness transducer therein, is mounted on the linear motion roller bearing, while a lever construction is used to load the elastomeric specimen.

In this set up the load can be simply controlled, which is different from a real seal configuration where the contact load is determined by the geometry of the seal and its housing in which the seal is compressed. Variation of the load would then require a number of housings with different dimensions, which is obviously not straightforward.

4.1.2 The elastomeric specimen

The elastomeric specimen used in the experiments is cut from an O-ring seal (Parker-Prädifa code V1 E235 P5008). Its geometry is then characterized by three parameters (fig. 4.2): The two radii r and R and the length l. The following values are given:

$$r = 5 \quad \text{mm}$$

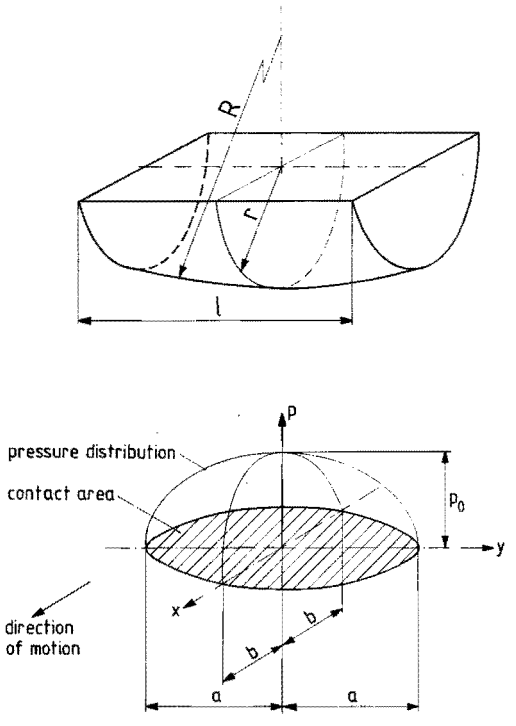


Figure 4.2
The elastomeric specimen for the film thickness measurements, the contact area and the pressure distribution (p).

R = 82.5 mm

l = 10 mm

The surface roughness of the elastomeric specimen is of the order of 1 μm and compares to the roughness of the rod seal shown in appendix A2.

The contact area and the contact pressure distribution of such a specimen, pressed onto a flat rigid body, are elliptical. The direction of motion will be in the x-direction (indicated in fig. 4.2) and the film thickness will be measured through the centre of the contact. The contact pressure and the film profile will only change slightly in the y-direction, since the contact is a long ellipse ($a \approx 6b$ as is shown below). Consequently a small uncertainty in the y-position of the scanned line through the contact area will hardly affect the reliability of the measurement.

A specimen of this shape is used, mainly because it yields a so-called elliptical Hertzian contact area. This enables the use of analytical formulas to calculate the dimensions of the contact area, the contact pressure distribution and the lubricant film thickness for the idealized condition of smooth surfaces.

The advantage of using analytical formulas for calculation of the contact pressure and the film thickness (for smooth surfaces) is, that the contact situation can be controlled by adjusting the running parameters like the contact load and the contact velocity, i.e. we can choose the load and the velocity on purpose to realize a certain combination of contact pressure and film thickness, which is regarded as interesting for the study of the surface roughness effects in the lubricated contact.

Of course the same can be done with specimen of more complex geometry (e.g. resembling a U-lip type seal as shown in fig. 1.1. on page 2), but this would require numerical calculations to find the contact pressure distribution and the film thickness for different loads and velocities. This would cost much more time without any significant benefit for our roughness investigation.

The dimensions of the contact area and the contact pressure distribution are derived in appendix I3.4 and shown in fig. 4.3 and 4.4. The major semi-axis a is 5 mm at a contact load of about 40 N. The contact ellipse is then extended over the full specimen length (= l = 10 mm) and a load of 40 N will therefore be the maximum in the experiments.

Figure 4.3
Dimension of the contact area versus the contact load.

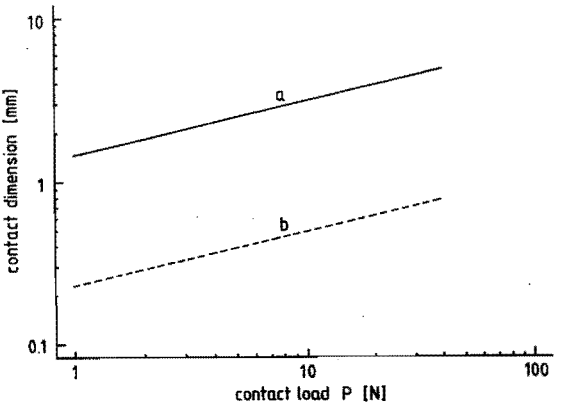
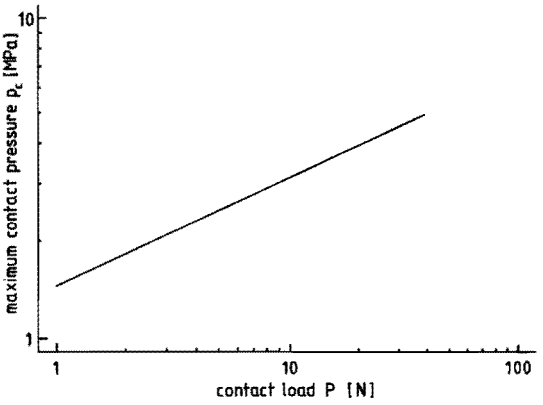


Figure 4.4
Maximum contact pressure of the elastomeric specimen versus the contact load.



The film thickness h_c in the centre of the contact area is derived in appendix K and depends on the contact load P, on the dynamic viscosity η of the lubricant and on the velocity u of the rigid body. It reads

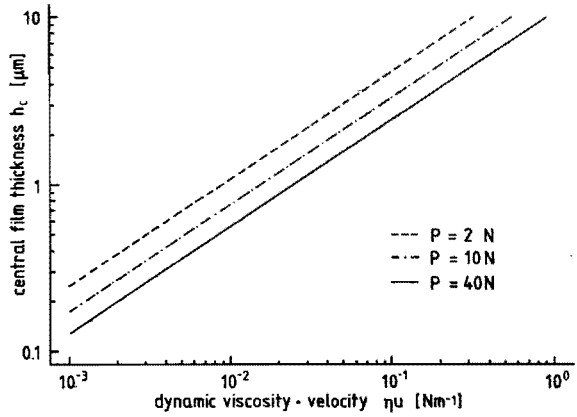
$$h_c = 23.9 \cdot 10^{-6} \, (\eta u)^{0.64} \, P^{-0.22} \qquad (4.1)$$

(h_c in [m]; (ηu) in [N·m^{-1}]; P in [N])

and is shown in fig. 4.5.

Figure 4.5
Film thickness h_c in the contact of the elastomeric specimen and the rigid body.

(P is the contact load).

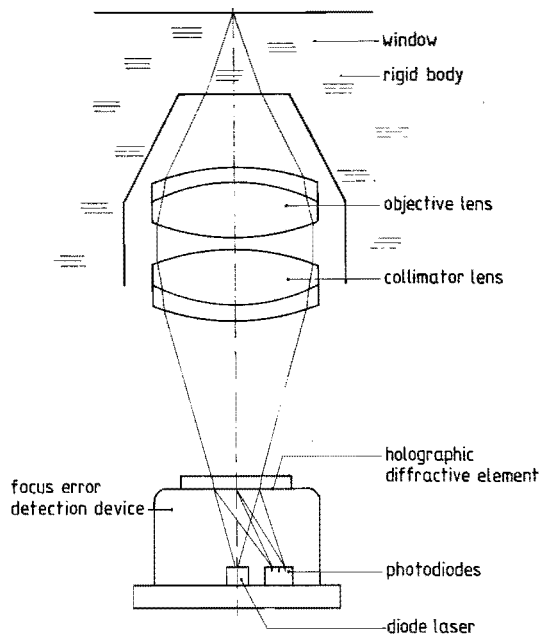


4.2 The film thickness transducer

4.2.1 Design

The film thickness transducer, which is mounted in the in the glass block of fig. 4.1, is shown in fig. 4.6. Its design satisfies the specifications derived in chapter 3 (see section 3.3 page 49). In essence it consists of 4 elements:

Figure 4.6
The film thickness transducer.



- The window;
- the objective lens;
- the collimator lens;
- the focus error detection device, which includes the diode laser, the photo-diodes and the holographic diffractive element.

The first three will be described below, the fourth (the focus error detection device) has been introduced in section 3.1.3 (page 27).

The window

The dimensions of the window were chosen to satisfy the requirements derived in chapter 3 (see point 6 in section 3.3 page 49):

The thickness is: $t = 2$ mm;
and the radius: $R_w = 0.5$ mm.

The index of refraction of the glass is:

$$n = 1.47$$

The lenses

The chosen lenses are two diffraction limited achromatic doublets delivered by Spindler & Hoyer. The main specifications are:

	focal distance f:	Product number:
Objective lens:	10 mm	32 2260
Collimator lens:	20 mm	32 2201

The aperture between the lenses is 4 mm in diameter, yielding a numerical aperture NA of

0.2 for the objective lens;
0.1 for the collimator lens.

fulfilling the requirements derived in chapter 3 (see point 3, 6 and 8 in section 3.3 page 49-50).

The light cone diameter R_{lc} at the lower window surface (see section 3.2.7 page 47) is now

$$R_{lc} = \frac{t\ NA}{n} = 0.27 \text{ mm}$$

The resulting radial tolerance for the mounting of the transducer in the rigid body is then

$$R_w - R_{lc} \approx 0.23 \text{ mm}$$

i.e. an eccentricity of up to 0.23 mm between the optical axis of the transducer and the centre axis of the window can be tolerated in the fabrication.

4.2.2 The spot dimension

The spot dimension is characterized by the fifty-percent-irradiance width (see appendix F1. The spot can be regarded as diffraction limited, as discussed in section 4.2.1 above, and its fifty-percent-irradiance width $d_{0.5}$ is then

$$d_{0.5} \approx \frac{\lambda}{2\,NA} = 2\ \mu\text{m}$$

since the wavelength l = 0.78 μm and the numerical aperture of the objective lens NA = 0.2.

4.2.3 Signal measurements and slope influence

The film thickness transducer, presented in section 4.2.1 above, is different from the displacement, shape and roughness sensor presented in section 3.1 (pages 26ff.), especially the focal distance of the objective lens is different (10 mm instead of 4.5 mm). Then the signals, measured as a function of the surface height relative to the focal point, will be different and they are therefore measured again for the film thickness transducer.

In this section, these measurements will be presented and the following items will be discussed:

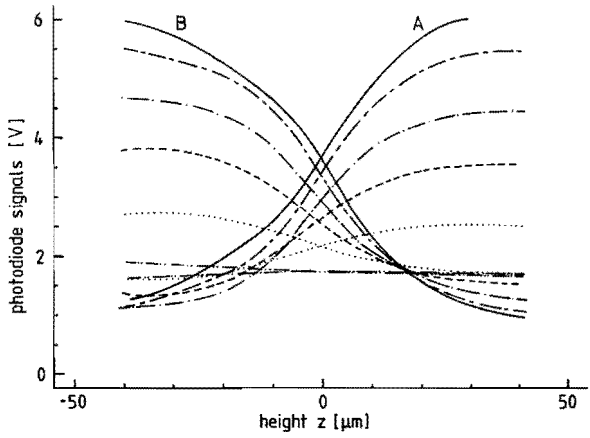
- The consequence of the different design for the photodiode and the focus error signals;
- The focus error signal, measured on the elastomer in contact with the lubricant.

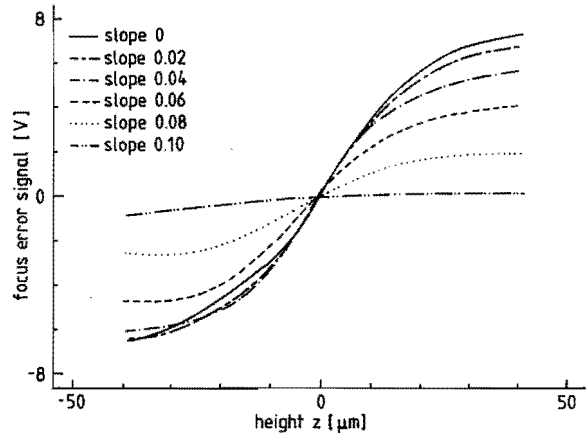
The consequence of the new design for the signals

The photodiode and focus error signals were measured in the same way as the signals of the displacement, shape and roughness sensors (described in appendix C1) and on the same silicon surface. The measurements were performed for different values of the surface slopes and the results are shown in fig. 4.7.

Figure 4.7
Signals of the film thickness transducer versus surface height z relative to the focal point (see fig. 3.1 page 24).
(Measured on the silicon surface).

a. photo diode signals.





b. focus error signal.

Two items are of special interest:
- The measurement range (i.e. the range in surface height z where the slope in the focus error signal is large;
- The influence of the surface slopes.

Considering the **measurement range** (the steep part of the focus error curve) we can compare the signals for zero slope of this film thickness transducer (fig. 4.7) with the signals of the displacement, shape and roughness sensor with the same holographic diffractive element, shown in fig. E6 (page 158):
- In fig. E6, the slope in the focus error signal is large for a surface height in the range from roughly –2.5 to 2.5 μm.

- In fig. 4.7, the slope in the focus error signal is large for a surface height in the range from roughly –15 to 15 μm.

We can conclude that the measurement range is increased by use of an objective lens with a larger focal distance The total measurement range of the film thickness transducer is 30 μm (from z = –15 μm to z = +15 μm), which is sufficient to measure film thicknesses up to 10 μm (the maximum film thickness expected) even when the focal point is at a relatively large distance (e.g. 5 μm) from the upper window surface.

Fig. 4.7 also shows the **influence of the surface slopes** on the signals of the film thickness transducer. Fig. 4.8, which shows the focus error signal versus the surface slope, is derived from fig. 4.7 and can be compared with fig. 3.11 on page 35. The slope influence appears to be larger than for the displacement, shape and roughness transducer having an objective lens with a smaller focal distance. This means that simultaneous measurement of the radial error signal (see section 3.1.4.2, page 36) is necessary to eliminate the slope influence.

Besides, the focus error signal of the film thickness transducer is hardly sensitive to the surface height (and thus to the film thickness) when the slopes are larger than about 0.08. The measurements will therefore not be reliable at slopes larger than 0.08.

It must be noted here, that the slopes in the undeformed roughness texture can be significantly larger (up to about 0.5 as shown in appendix A2), yielding erroneous measurements. However, this is probably not very serious, since the roughness will be deformed ("flattened") in the lubricated contact due to micro-EHL (see section 1.1.1.2, page 7, and section 1.2 page 8ff.). Therefore the

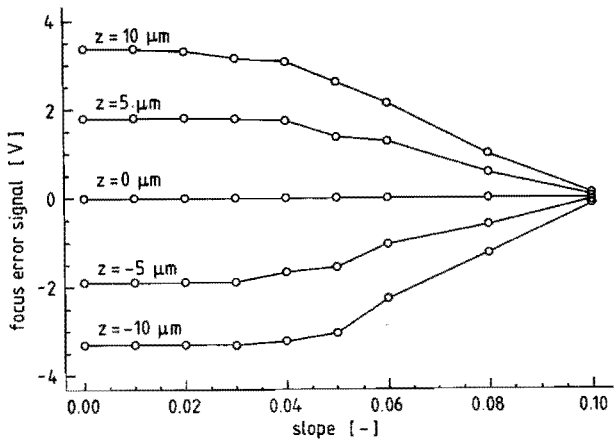


Figure 4.8
Influence of the surface slopes on the focus error signal.

roughness height and, consequently, the surface slopes will probably be smaller. Besides, measured points where the slope is nevertheless apparently too large (to be determined from low photodiode signals A_1, A_2, B_1 and B_2) can possibly be rejected without essentially influencing the measurement results, since such points are not present in the neighbourhood of the extremes (the summits and the valleys).

Simultaneous measurement of both the focus error signal and the radial error signal requires a new electronic set up for the signal amplification and manipulation, since the present set up (developed for the displacement, shape and roughness measurements) can only handle the focus error signal. This new electronic set up, which must also incorporate a better signal to noise ratio as discussed in appendix H, is not available yet. Therefore only some preliminary measurements will be presented in this thesis to show the working of the method qualitatively.

The focus error signal measured on the elastomer in contact with the lubricant

Fig. 4.9 shows the set up for the measurement of the focus error signal on the elastomer (of the same polyurethane as the specimen; Parker-Prädifa code P5008) in contact with the lubricant and fig. 4.10 shows the resultant focus error

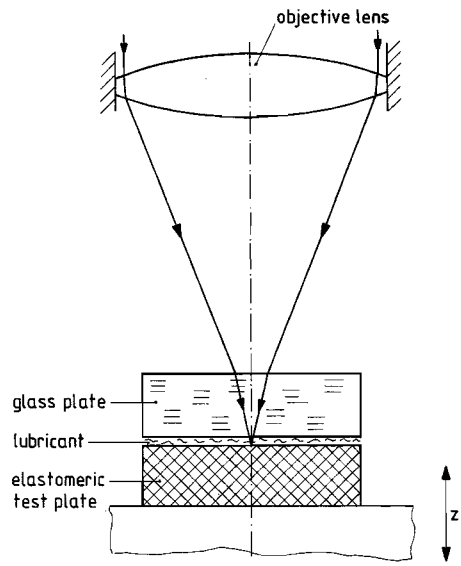
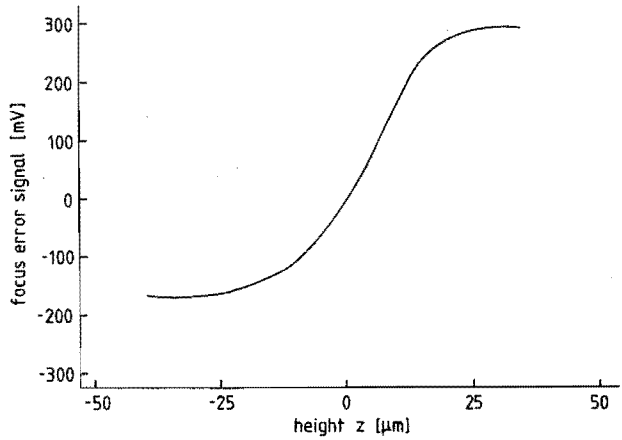


Figure 4.9
Set up for the measurement of the focus error signal on the elastomer in contact with the lubricant.

Figure 4.10
Focus error signal of the film thickness transducer versus surface height z relative to the focal point (see fig. 3.1 page 24).
(Measured on the elastomer, Parker-Prädifa code P5008, in contact with the lubricant)



curve. The focus error signal appears to be linear in the range $0 \leq z \leq 15$ µm where

$$\frac{d(fes)}{dz} \approx 17 \text{ mV}\cdot\mu\text{m}^{-1}$$

This curve can be used to derive the film thickness from the measured focus error signal. However, in this measurement the gap height ("film thickness") between the elastomer and the glass plate was constant. When deriving the film thickness from the measured focus error signal, we must account for the lubricant's index of refraction n (see appendix I2). Then we find for the slope in the focus error signal

$$\frac{d(ffs)}{dz} \approx \frac{17}{n} \text{ mV}\cdot\mu\text{m}^{-1} = 11.5 \text{ mV}\cdot\mu\text{m}^{-1} \qquad (4.2)$$

since the index of refraction $n = 1.47$.

4.2.4 Influence of the contact pressure and temperature

In section 3.2.7 and 3.2.8 (page 46ff.) was discussed that the pressure and temperature influence on the film thickness measurement can be large for contact pressures up to 50 MPa and for contact temperatures up to 200 °C. Therefore the running conditions for the measurements in this chapter will be chosen appropriately to keep the contact pressure and temperature low.

Influence of the contact pressure

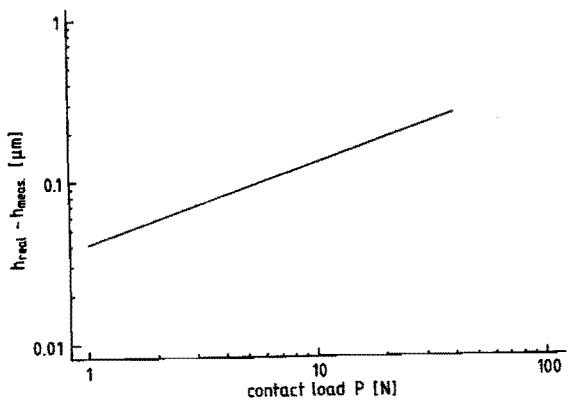
The influence of the pressure in the contact of the elastomeric specimen and the glass window used in this chapter is estimated in appendix I3.4. Using the relation between the pressure and the load shown in fig. 4.4. (page 54) the resulting error in the film thickness measurement $\Delta_p h$ can be derived from the adjusted contact load according to fig. 4.11. This error is defined as

$$\Delta_p h = h_{real} - h_{meas}$$

in which: h_{real} = Real film thickness in the contact;
h_{meas} = Film thickness derived from the measurement, when the pressure influence is not accounted for.

At a contact load of 2 N negligence of the pressure influence yields an error in the film thickness measurement of 0.06 μm. This is not negligible according to the requirements mentioned in section 2.2 (page 19), but it is already reasonable.

Figure 4.11
Influence of the contact load on the film thickness measure-ment.



Influence of the contact temperature

The influence of the temperature rise in the contact area and of the consequent temperature increase of the whole rigid body is discussed in appendix I4.5. The total temperature influence on the film thickness measurement is roughly estimated, yielding

$$\Delta_T h = h_{real} - h_{meas} = c\ \Delta T$$

in which: h_{real} = Real film thickness in the contact;
h_{meas} = Film thickness, derived from the measurement, when the temperature influence is not accounted for;
ΔT = Temperature increase
c = Constant $\approx$ 0.07 $\mu m \cdot K^{-1}$

The temperature rise in the contact area will not be larger than about 0.5 K and the resulting influence on the film thickness measurement will therefore be limited to roughly 0.035 μm. This is not negligible according to the requirements mentioned in section 2.2 (page 19), but it is also reasonable.

4.3 Preliminary measurements

Two series of preliminary measurements were performed:
- The macroscopic shape of the elastomeric specimen was measured at different loads and very low velocities, using a lubricant with a low viscosity. (Then a lubricant film is hardly formed);
- Film thickness measurement at a small load and different velocities using a high viscosity lubricant. (Then a lubricant film will be formed with a velocity dependent thickness).

The results will only be discussed qualitatively, i.e. the shape and/or the lubricant film thickness will not be derived from the measured focus error signal, mainly because of the influence of the elastomer's surface roughness on the focus error signal (see section 4.2.3 page 59). This influence can not be eliminated, because simultaneous measurement of the radial error signal is not possible at the moment.

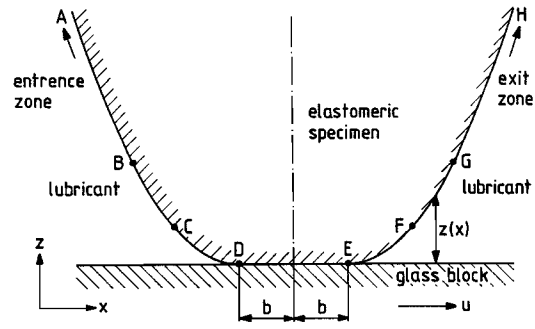
4.3.1 Measurement of the shape of the loaded specimen

In this section, it will be tested whether the transducer is able to measure the macroscopic shape of the elastomeric specimen, when the specimen is surrounded by the lubricant and when the reflectance on the specimen's surface is consequently low (see appendix D2). These measurements will be performed using the test rig presented in section 4.1.1 (fig. 4.1 page 51). The glass block will slide relative to the specimen and the transducer will experience a narrowing

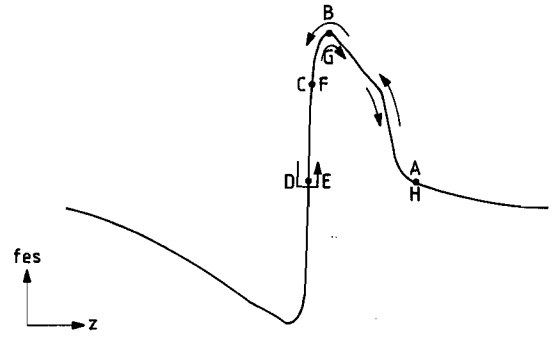
gap in the entrance zone (at the left hand side in fig. 4.12a). In the contact area (between D and C) the gap height will be almost zero, since the film thickness will be very thin in these experiments, as shown below. The width of the contact area (2*b*) can be derived from figure 4.3 (page 54). Finally the transducer will arrive in the exit zone, where the gap height increases.

Figure 4.12
The shape of the loaded specimen and the expected focus error signal.

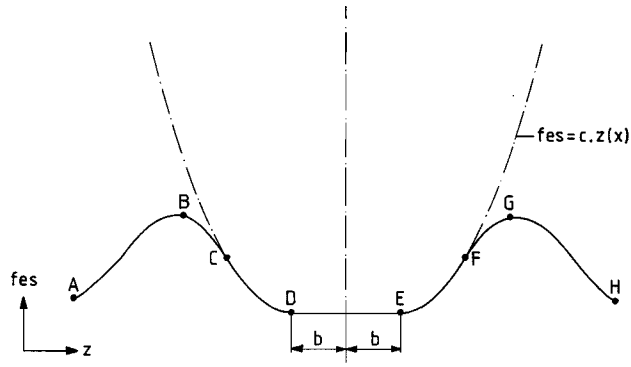
a. The shape of the specimen: *z(x)*.
(The x- and y-axis are not equally scaled)



b. The focus error signal versus surface height: *fes(z)*



c. The expected focus error signal versus the position: *fes(x)*.



The expected shape of the focus error signal is shown in fig. 4.12c and can be divided in different parts:

AB: When the gap height z is small enough (of the order of 100 μm, in point A) the focus error signal starts to increase, until the maximum is reached in point B (z is roughly 50 μm).

BC: The signal now decreases at decreasing gap height z, but is not proportional to z.

CD: In point C the linear part of the focus error curve is reached and the focus error signal follows the line $fes = c\ z(x)$ (c = 11.5 mV·μm^{-1} as shown by eq. (4.2) on page 61).

DE: The gap height is zero and the focus error signal remains constant. (The value depends on the focal point position relative to the window surface, see appendix I1).

DH: Similar to AD, but in the opposite direction.

In the measurements the following conditions were applied:

- The lubricant is the mix of 75 percent Shell Ondina 15 and 25 percent Shell Ondina 68, used to eliminate the reflection on the lubricant to glass interface (see appendix G2.3).

  The dynamic viscosity η at 20 °C is between 0.03 Pa·s (the viscosity of Ondina 15) and 0.21 Pa·s (the viscosity of Ondina 68);
- The velocity u is 0.5 mm/s;
- The contact load is in the range of 2 to 40 N.

Consequently, the film thickness will be smaller than about 0.06 μm, according to eq. (4.1) on page 54. Then eventual film thickness variations will also be very small (apart from the roughness influence) and the focus error signal will therefore be constant between the points D and E in fig. 4.12a.

The results of the measurements are presented in fig. 4.13, as well as a fit of the expected focus error signal between the points C and F, which is derived theoretically as follows:

- The width $2b$ of the contact area (between the points D and E in fig. 4.12a) is derived from fig. 4.3 (page 54);
- The curve outside the contact (but within the linear part of the focus error signal, between point C and D and between point E and F in fig. 4.12a) is calculated using eq. (4.2) on page 61

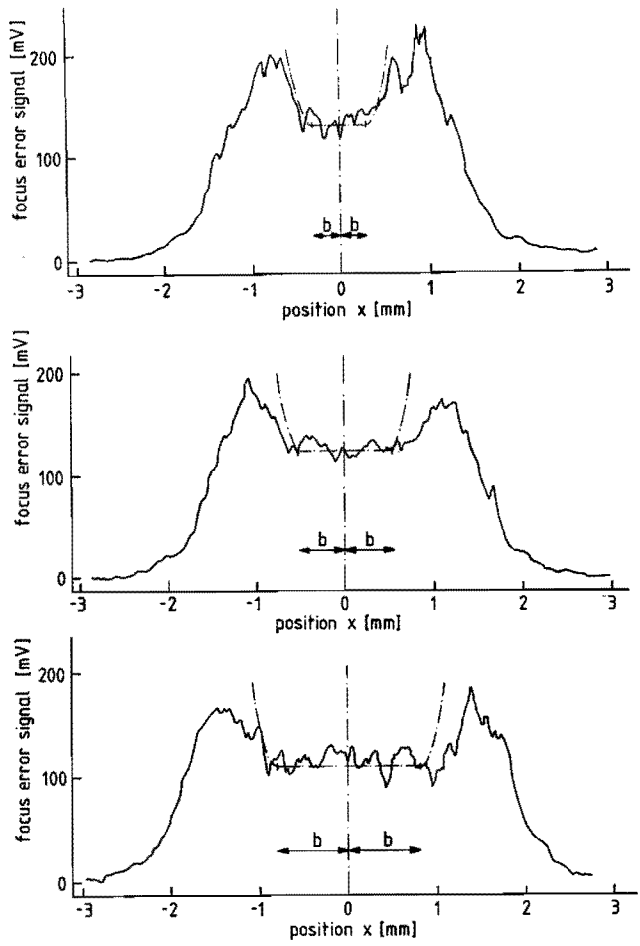
$$fes(z) = c\ z(x) \qquad \text{for } |x| \geq b$$

Figure 4.13
The measured and expected focus error signal versus the position: *fes(x)*.
——— measured curve
–·–·– expected curve

a. At a contact load *P* of 2 N.

b. At a contact load *P* of 10 N.

c. At a contact load *P* of 40 N.



in which

$$c = 11.5 \text{ mV}\cdot\mu\text{m}^{-1}$$

and

$$z(x) = r\left[1 - \sqrt{1 - \left(\frac{x-b}{r}\right)^2}\right]$$

r being the radius of the (undeformed) specimen (r = 5 mm)

(This formula expresses the undeformed curvature of the specimen at zero load. The real curvature will be different when the specimen is loaded, but this expression will due for the qualitative comparison).

As shown by these three figures (4.13a to 4.13c) the measured and expected curves compare qualitatively. The increase in contact width is clearly seen in the measured curves.

Another feature of the measured curves is the relatively large ripple. These ripples have an amplitude of roughly 10 mV (which compares to a height variation of roughly 1 μm) and are perhaps caused by the specimen's surface irregularities (e.g. roughness). This would then suggest that the asperities are not oppressed by the contact load, but we must emphasize here that the roughness and the resulting height variations can be much smaller than the height variations derived from these variations in the focus error signal, due the the influence of the surface slopes. Further, the asperities are perhaps not totally oppressed, e.g. due to local entrapments (see section 5.4). However, whether the ripple is caused by surface irregularities can not be decided. Other origins, like a relatively large noise due to the very low reflectance on the lubricant to elastomer surface, are also possible. Further investigation is therefore needed.

4.3.2 Film thickness measurement

Preliminary qualitative film thickness measurements were performed on the same elastomeric specimen, described in section 4.1.2 on page 52, using Shell Tellus C320 oil with a dynamic viscosity η of 1.1 Pa·s at a temperature of 20 °C. The contact load P in these experiments is 2 N. The consequent Hertzian contact width $2b$ is then 0.6 μm (see fig. 4.3 page 54) and the following values for the film thickness can be derived from eq. (4.1) on page 54:

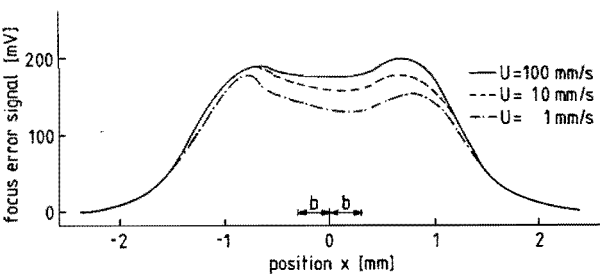
h_c = h(x=0) = 0.26 μm for: u = 1 mm/s
h_c = h(x=0) = 1.14 μm for: u = 10 mm/s
h_c = h(x=0) = 5.00 μm for: u = 100 mm/s

The calibration curve, shown in fig. 4.10 (page 61), is not valid for these measurements, because of the reflection on the window to lubricant interface which is caused by a different index of refraction of the Tellus C320 oil and of the duran glass (see appendix D1). This reflectance is, however, not important at the moment, since the measurements are only qualitatively. Quantitative film thickness values will not be derived.

The results of these measurements are presented in fig. 4.14, where the nominal focus error curve is shown. The oscillations present in the measurements, and probably caused by the elastomer's surface roughness, are eliminated to obtain a clear comparison of the curves.

The shape of the curves is as derived in the former section (see fig. 4.12c page 64). The focus error signal in the contact centre (x = 0) is clearly smaller for a lower velocity, as is expected, since the film thickness will be smaller at a lower velocity and the focus error signal is smaller at a thinner film. We can therefore conclude that the film thickness is properly measured qualitatively and we may expect that quantitative measurements will be possible when the pressure and temperature influence are further investigated and when new electronic instrumentation is available.

Figure 4.14
Measured focus error signal at different velocities of the elastomeric specimen.



4.4 Conclusions

Preliminary film thickness measurements were presented. These measurements showed, that the method is in principle suitable for film thickness measurements, despite the low reflectance on the lubricant to elastomer interface.

The test conditions were chosen appropriately to keep the contact pressure and temperature influence on the measurements small (uncertainty within 0.1 μm). However, when an accuracy of 0.01 μm in the measurements is wanted, more investigation on the pressure and temperature influence is needed.

Finally, new electronic equipment is needed to fulfil the required maximum response time, to increase the signal to noise ratio and to enable simultaneous measurement of the radial error signal.

CHAPTER 5 MEASUREMENT OF THE ROUGHNESS DEFORMATION OF ELASTOMERS UNDER STATIC LOAD

Up till now we discussed the development of a method for measurement of the lubricant film thickness and the roughness deformation in the sliding contact of a rough elastomer and a smooth metallic body. As discussed in chapter 1 the importance of such measurements is to obtain more insight in the roughness influence on the lubrication of e.g. reciprocating seals for hydraulic cylinders, especially in the mixed lubrication regime where the roughness (and its deformation) of the elastomer (the seal) has a significant influence on the friction.

During the investigation of the focus error film thickness transducer it was considered that the method could also be used to measure the deformed roughness texture in a statically loaded contact (i.e. without motion of the surfaces relative to the contact area). Measurement of the roughness deformation in the statically loaded contact can e.g. be helpful in studying the frictional behaviour of pneumatic seals (which are generally not or hardly lubricated) and in investigating the tightness of static seals (what contact load is needed to prevent leakage, i.e. to close all leakage paths between the contacting roughness asperities in the contact area).

Such preliminary measurements will be presented in this chapter, but first the literature on the roughness deformation in statically loaded contacts will be briefly reviewed.

5.1 Literature review on the contact of rough surfaces

In literature most attention is paid to the real area of contact. Both theoretical and experimental work can be found and will be reviewed separately, based on recent reviews by Visscher and Struik (1992 section 2.2 and 2.3) and Hendriks (1992[b]). An earlier review was presented by Thomas and King (1977).

5.1.1 Theoretical work

Most theoretical work on the contact of rough surfaces is dedicated to the calculation of the real area of contact and thus considers the deformation of the asperity summits and their immediate surroundings. We can distinguish between models based on random process theory and models based on fractals. Besides, numerical methods were recently used to calculate the contact pressure and deformations, using a real measured profile instead of parameters characterizing the profile.

It is essential for the calculations to decide properly, whether the deformation is elastic or plastic and possible interaction between asperities can be important. These items will therefore also be discussed.

Models based on random process theory

McCool (1986) reviewed and compared the different models based on random process theory and the most important models are reviewed here.

Greenwood and Williamson (1966) based their model on the assumption that all asperities were spherical with equal radius, while the summit height distribution was assumed to be Gaussian. They also assumed that the contact load on one asperity did not influence the height of neighbouring asperities. The real area of contact was calculated for the contact of a rough and a smooth surface, assuming pure elastic deformation and using Hertzian theory, and they found that the real area of contact A_r was almost proportional to the contact load P:

$$A_r \sim P^{0.96} \tag{5.1}$$

Their theory was generalized by Greenwood and Tripp (1971), describing the contact of two rough surfaces and involving eventual plastic asperity deformation. The asperities are now modeled by paraboloids instead of spheres. Conical asperities were also considered. All had the same size (top radius or slope), while the distribution of the summit height was Gaussian, and they also found an almost proportional relation between the real area of contact and the contact load.

Onions and Archard (1973) presented a model similar to the Greenwood and Williamson model. The asperities were also modeled by spheres, but with varying radii of curvature. They found that the real area of contact was exactly proportional to the contact load.

Models based on fractals

Majumdar and Bhushan (1990 and 1991) and Majumdar and Tien (1990) used fractals to characterize the surface roughness texture. In this description, it is assumed that the same roughness height variations can be found on different length scales. Only the amplitude of the height variations is smaller at smaller length scale, but the profile is similar.

Majumdar and Bhushan (1991) derived new formulas for the real area of contact and the surface load. The individual contact area is calculated describing the summits as spheres and using hertzian theory, just as Greenwood and Williamson (1966) did, but now accounting for the different radii of curvature at different length scales. They also neglected the eventual deformation of the bulk material. According to their fractal model, the real area of contact is in general less than proportional to the contact load:

$$A_w \sim P^c \tag{5.2}$$

in which c is a constant, depending on the fractal dimensions of the surface ($1/2 \leq c \leq 1$).

Calculations, using the real, measured profile directly

Lubrecht and Ioannides (1991) numerically calculated the contact pressure of a surface with unidirectional roughness and a smooth surface. A measured roughness profile was directly used in the calculations, instead of derived parameters. This obviously increases the number of points in the calculations extremely and multilevel techniques were used to reduce the computing time. Conclusions concerning the deformed roughness texture, e.g. on the real area of contact, were not derived.

Xian and Zheng (1991) calculated the real area of contact for a three-dimensional roughness, also using the measured texture directly. They proposed a simplification by fitting the shape of the contacting asperities with a quadratic function. This enabled the use of Hertzian theory for the calculation of the individual contact areas and reduced thus the computing time. These calculations again yielded a proportional relation between the real area of contact and the the contact load.

They also calculated the real area of contact for a two-dimensional roughness fully numerically, without approximation of the asperity summits with a quadratic function. Now the real area of contact was proportional to the contact

load at smaller roughness heights (roughly $R_a < 1$ μm), but a non-linear relation was obtained for higher roughnesses.

Elastic and plastic asperity deformation

In early times, it was considered that the stresses in the real contacts were high, because of the small real area of contact. The asperities would therefore deform plastically (e.g. Bowdon and Tabor, 1954, pp. 10-14)). However, Archard (1957, 1974) suggested, that, although the deformations may be initially plastic, they will be elastic after running in, due to increase of the real area of contact (by the initial plastic deformation) and thus decrease in the real contact pressures.

Greenwood and Williamson (1966) introduced a plasticity index, which accounts for the mechanical properties and the surface texture. The value of this plasticity index indicates, whether the deformation is elastic or plastic. They found that the load was hardly a factor: the asperities of most surfaces deform plastically, even at the lightest loads.

Bhushan (1984) investigated the contact of a magnetic tape and a hard material and derived a plasticity index for polymers. He concluded, that the deformations were elastic for most magnetic tapes and similar results were found for magnetic discs (Bhushan and Doerner, 1989).

Majumdar and Bhushan (1991) derived a criterion for plastic deformation for fractal roughness characterization, and showed that the smaller contacts were plastic, while the larger contacts were elastic. However, Greenwood and Williamson (1966) found the contrary: Small elastic and large plastic spots. Majumdar and Bhushan explained the difference, considering that the small scale asperities, which lead to small contacts, have also a small radius of curvature, while Greenwood and Williamson assumed equal radii for all asperities.

Asperity interaction

In general no attention is paid to eventual asperity interaction. Such an interaction can exist due to deformation of the bulk material, which then causes a vertical displacement of a non-contacting asperity when a neighbouring higher asperity is in contact and therefore oppressed (see Podbevsek, 1992 section 5.3). The lower asperity will then first come into contact at a higher load than expected from calculations in which the deformation of the bulk material is neglected. The real area of contact will therefore be smaller. Also the valley

between two asperities is lowered and can remain even at high loads (see e.g. Vergne et al., 1985) . The final consequence of asperity interaction is that the real area of contact is smaller than without interaction.

5.1.2 Experimental work

A lot of methods has been applied to determine the real area of contact experimentally and most were reviewed by Woo and Thomas (1980) in general, while Bhushan (1985a) reviewed and discussed techniques to study the contact of polymeric magnetic media. Some methods only determine the real area of contact and other methods yield the surface texture under load, from which the real area of contact can be derived, if desired. The most important methods are the electrical and the optical, which will be recalled here.

Electrical methods

The electrical resistance in the contact of rough surfaces is higher than the bulk resistance because of the constrictions at the small contact spots. The contact resistance can thus be a measure for the real area of contact. However, Bhushan (1985a) reported that the constriction resistance does not depend on the area of the contact, but on the radius of the constriction. He therefore called the method semi-quantitative. Further, the method is obviously only applicable to study the contact of conducting surfaces and the presence of insulating layers, e.g. an oxide film, may cause significant underestimation of the real area of contact. The magnetic tapes, used by Bhushan (1985a) were uncoated, since the conductance was reasonable and trials to coat the tapes with a metal layer were not successful.

In his second paper Bhushan (1985b) concluded the method to be unreliable, mainly because of the influence of insulating films. Later work included therefore only optical methods.

Optical methods

Optical methods need a transparent surface to enable observation of the contact and most of them require a reasonable reflectance on the other surface. They generally overestimates the real area of contact, according to Bhushan (1985a). He also concluded, that most methods are not applicable to study the magnetic

tape contact for distinctive reasons, which also apply for the contact of (non-transparent) elastomers in general. He therefore chose interferometry at last to measure the real area of contact (see also Bhushan, 1985b; Bhushan and Dugger, 1990).

Interferometry yields a pattern of dark and light fringes, each representing a contour of equal gap height (see appendix B4.1). A problem reported by Bhushan (1985a), is the relatively low reflectance of the magnetic tape, which reduces the sharpness of the fringes. The consequence is overestimation of the real area of contact with possibly 80 or more percent. Ohara (1976), investigating the contact of two transparent polymers, coated the surfaces with silver and obtained thus narrow sharp dark fringes and broad light fringes. However, he was also unable to determine the real area of contact accurately, for the first dark fringe near a real contact represented a gap height of a quarter wavelength. The real contact is therefore only an invisible part of the bright spot within this dark fringe.

5.1.3 Conclusions

Several theoretical models concerning the contact of rough surfaces have been developed in the past. They generally differ in both the assumptions of the mechanical asperity behaviour (elastic or plastic deformation) and in the method to describe the roughness texture. Experimental verification of these models appears to be troublesome. The real area of contact, e.g., can be significantly under or overestimated, using electrical or optical methods respectively.

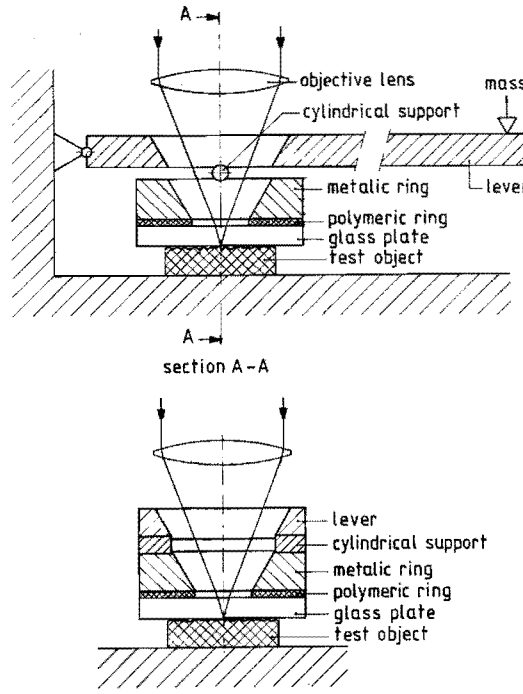
Therefore we can conclude that there is an apparent need for a new and reliable method to measure the roughness deformation (and the real area of contact) in a statically loaded contact.

In this chapter the optical focus error detection system will be used for such measurements on a rough elastomer in contact with a smooth glass plate.

5.2 Test rig

The measurements are performed with the device shown in fig. 5.1 in which the load is adjusted using a lever. The contact is accessible for the optical head through a gap in the lever and the load is transmitted to the glass plate via a thin cylinder and a metallic ring. This enables the load to be spread uniformly

Figure 5.1
Set up for measurement of the deformed roughness texture under static load.



over the contact. The relatively soft polymeric ring is used to ensure a large contact area between the ring and the glass plate, which is needed to avoid high local stresses in the glass plate.

This device is mounted on the test rig for optical shape and roughness measurements described in appendix C2 and more elaborately by Struik and Chang (1987). The used scanning device is the double wedge focus error sensor described in section 3.1.2 (page 26). The window near the lens in the original transducer (fig. 3.3 page 26) is removed for these measurements to avoid significant aberration from the glass plate on the test surface (see section 3.1.4.1 page 32).

The irradiance distribution of the measurement spot is as depicted in fig. 3.5 (page 30), having a fifty-percent-irradiance width $d_{0.5}$ of about 0.82 μm. This spot dimension is the limit in the spatial resolution of the measurements.

5.3 Accuracy in the height measurement

In the next section (5.4) the height profile of elastomeric surfaces will be measured through a loaded glass plate and the dimension of flattened areas ("real areas of contact"), found in these measurements will be derived. In this section the accuracy of the height measurements will be discussed, using the analysis presented in section 3.2. The accuracy in the derived size of the flattened areas will be discussed later in section 5.5.

The accuracy in the measured surface height is determined by the following factors:

- The influence of the reflection on the lower glass surface;
- The accuracy of the scanning device;
- The contact pressure and the temperature influence.

The slope influence, discussed in section 3.2.3, is not a factor in these measurements, because these are performed in the closed loop mode, focusing the objective lens continuously on the surface.

The influence of the lower glass surface

As discussed in section 3.2.5 (and more elaborately in appendix G) the reflection on the lower, contacting, glass surface can influence the measurement significantly. Therefore a liquid is needed in the contact to prevent reflection on the lower glass surface.

The indices of refraction of the liquid and the glass must be equal. Chosen is an oil mixture of 75 percent Shell Ondina 15 and 25 percent of Shell Ondina 68 for the liquid, in combination with Duran glass, both having an index of refraction n of approximately 1.47 (see appendix D1).

The accuracy of the scanning device

It was reported by Struik and Chang (1987) that the uncertainty in the measurements is 10 nm at the maximum, when scanning free surfaces, due to noise in the system. In this chapter, measurements will be performed on an elastomer with a liquid and a glass plate on it. Then the reflectance will be very low (see appendix D2) and consequently the signal to noise ratio will also be low. However, tests showed that the reflectance is still large enough to obtain proper

measurements, using the glass plate and the liquid (see appendix L). In these tests, roughness values derived from measurements with glass plate and liquid on the surface were about 10 percent higher than the values derived from measurements without glass plate and liquid, but this is smaller than the differences (up to 20 percent) which are found in one measurement series. Therefore, the 10 percent difference between the two measurement series is not significant.

The pressure and temperature influence

In the measurements presented in this chapter, the objective lens is focused onto the elastomeric surface. Then the lubricant's index of refraction, and its pressure and temperature dependence, is the only factor to account for. The pressure will be smaller than 0.1 MPa and the temperature variation will be smaller than 1 K. Therefore the pressure and temperature influence on the measurements will be negligible (smaller than 0.03 percent, as derived in appendix I3.1 and I4.1).

5.4 Measurements

Measurements were not only performed with a liquid in the contact area, but also without a liquid. The reasons were:

- The liquid in the contact area appears to influence the roughness deformation significantly, while contact models, calculating the real area of contact, do not account for the presence of a liquid in the contact area;
- Measurement of the real area of contact is reliable, even if there is no liquid in the contact area, since there is no gap between the test surface and the glass plate in a real contact, i.e. there is only one reflecting surface: the interface of the elastomer and the glass.

  Only the roughness profile outside the real contacts is measured inaccurately, but this is of no importance deriving the real area of contact from the measurements.

5.4.1 Measurements with a liquid in the contact

Measurements were performed on both a rough and on a smoother piece of polyurethane (Parker-Prädifa material code P5008) with an E-modulus of 45 MPa at small strains (see appendix I3.4). Height contour plots will be used to show flattened areas and the height distribution curve will be shown to provide an indication on the roughness deformation caused by the contact load. Also some roughness parameters will be derived. These are:

R_a = Centre Line Average (CLA) roughness height (or arithmetic average roughness height); [m]

R_q = Root Mean Square (RMS) roughness height (or the standard variation of the roughness height distribution); [m]

Sk = Skewness [-]

Kt = Kurtosis [-]

(See appendix A1 for the definition and e.g. Halling (1978 pp. 22-39) or Thomas (1982 chapter 5) for a more elaborate outline of surface roughness characterization).

The refraction in the measurements with glass plate and liquid on the polyurethane plate is accounted for in deriving the height distribution curve and the roughness parameters. This is done by multiplying the measured roughness height by the index of refraction n, since the apparent (measured) roughness height variation is $1/n$ times the real height variation (see e.g. appendix I2).

Measurement performed on the rough polyurethane plate

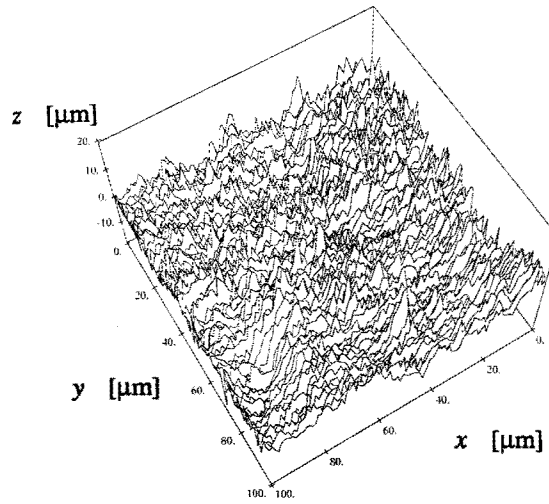
The first measurement of the deformed roughness texture under static load was performed on a rough piece of polyurethane. A plot of the undeformed roughness texture is shown in fig. 5.2 (including the derived roughness height distribution and some roughness values), while fig. 5.3 shows the measured roughness texture and the height distribution at a contact load of 36 N.[1]

The following parameters apply to both measurements:

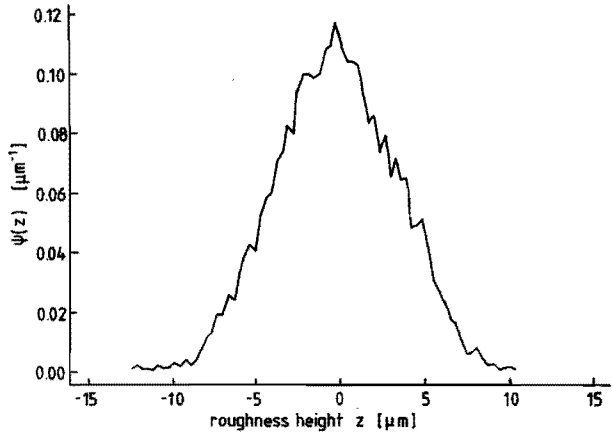
- Diameter measurement spot = 1 μm;
- Sample distance in X-direction = 1 μm;

[1] These two measurements were not performed on precisely the same part of the surface, because this is hardly possible with the present test rig.

Figure 5.2
Roughness texture of the relatively rough polyurethane specimen.



a. Surface plot.



b. Height distribution

R_a = 2.87 µm
R_q = 3.57 µm
Sk = -0.04
Kt = 2.80

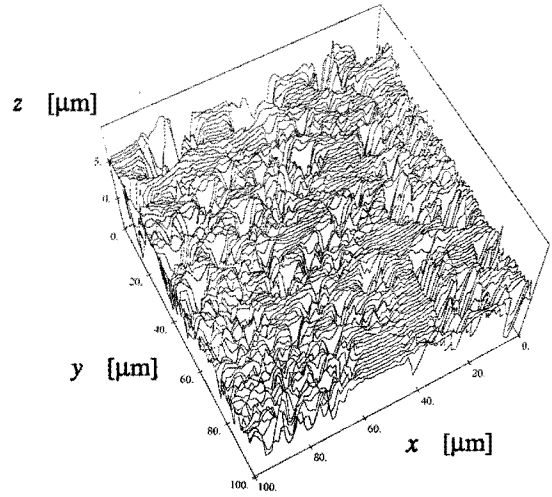
- Sample distance in Y-direction = 1 µm;
- Apparent contact area ≈ 620 mm²;

No additional filtering of the measured data was applied.

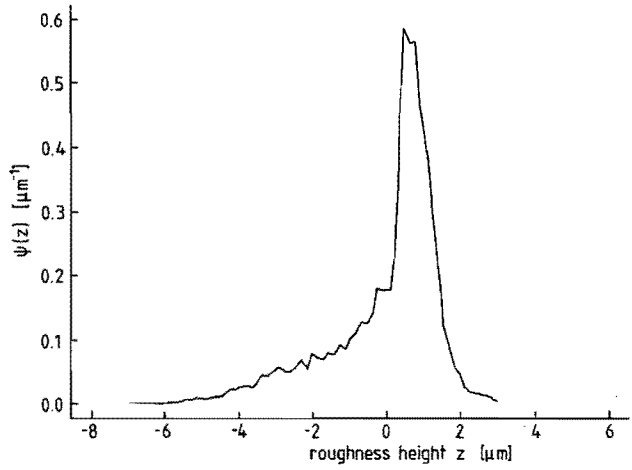
The undeformed roughness height distribution accurately resembles the Gaussian distribution, for which the skewness is 0 and the kurtosis is 3. The roughness deformation in the loaded contact is clearly indicated by the height distribution. Not only the width of the distribution curve is decreased in the loaded situation, but also the skewness is decreased, while the kurtosis is increased.

The decrease in the skewness indicates, that especially the peaks are

Figure 5.3
Roughness texture of the relatively rough polyurethane specimen at a contact load of 36 N (Avarage contact pressure p_a = 0.058 MPa).
(with liquid in the contact)



a. Surface plot.



b. Height distribution

R_a = 1.14 μm
R_q = 1.49 μm
Sk = -1.33
Kt = 4.62

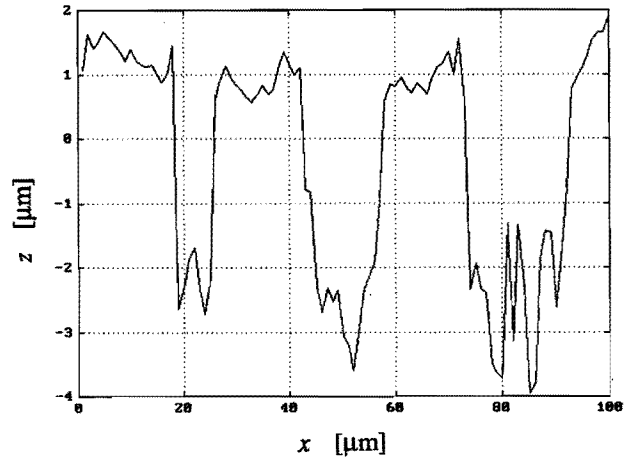
deformed while the valleys are more or less unaffected, as is also indicated by the shape of the distribution curves. The increase in the kurtosis indicates, that a larger part of the surface has a height around the mean level (z = 0). This is because the deformed asperities (with $z > 0$) are lower than in the undeformed situation, which means that the number of points with a height around the mean level is increased. We can also conclude that the deformation (i.e. the decrease in height position) of the points which were around the mean level in the undeformed state, is smaller than the deformation of the asperities. Otherwise the kurtosis would then be smaller, because there would be more points with a height below the mean level (i.e. with $z < 0$) in the deformed state.

As shown in fig. 5.3a, the flat areas can be seen very well and they are significantly larger than the length scale of the height variations in the undeformed texture. We may therefore conclude that the flatness of these areas is a result of deformation due to the contact load.

One might expect, that these flattened areas are all real contact areas. However, fig. 5.4, which is a cross section of fig. 5.3a, shows that the flattened areas have not an equal height, but height differences between these areas of up to 1 μm occurs. This indicates that apparently a significant part of the load is supported by the liquid and not by real contacts.

Also the shape of a lower deformed asperity (e.g. the asperity around x = 35 μm in fig. 5.4) looks like a so-called "entrapment" , i.e the liquid film between the asperity and the glass is thicker in the middle of the asperity than at the boundary. Such an entrapment is typical for squeezing contacts and is caused by the fact that the liquid near the boundary can relatively easy flow out of the contact, reducing the film thickness there, while the liquid in the middle of the contact experiences more resistance (from viscous shear) against outflow, preliminary because of the larger distance to the contact boundary and secondly because of the diminishing film thickness near the boundary. Perhaps the shape of the deformed asperity is indeed caused by squeezing, but more investigation is needed to verify this idea.

Figure 5.4
Cross section of fig. 5.3a at y = 10 μm.



Measurements performed on the smoother polyurethane plate

Further measurements were performed on a smoother piece of polyurethane. The undeformed surface texture is shown in fig. 5.5a, including the height distribution and some roughness parameters. The results of the measurements performed at different loads are presented in fig. 5.6 to 5.10, showing the highest height contour lines and the roughness height distribution. For the lowest load, the roughness texture itself is also shown (fig. 5.6a) for comparison of the deformed and the undeformed situation.

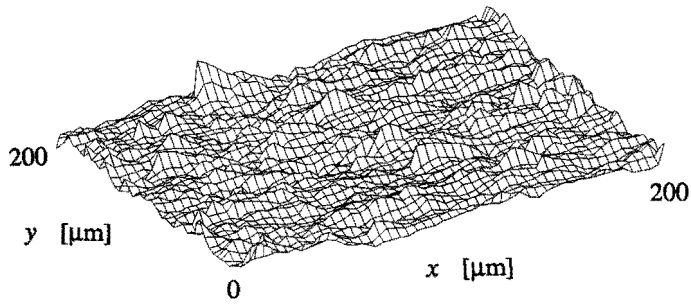
The following parameters apply to all measurements:

- Diameter measurement spot ≈ 1 μm;
- Sample distance in X-direction = 2 μm;
- Sample distance in X-direction = 2 μm;
- Apparent contact area ≈ 600 mm²;

No additional filtering of the measured data was applied.

Figure 5.5
Roughness texture of the smoother polyurethane specimen.

a. Surface plot.



b. Height distribution

R_a = 0.49 μm
R_q = 0.66 μm
Sk = 0.59
Kt = 6.36

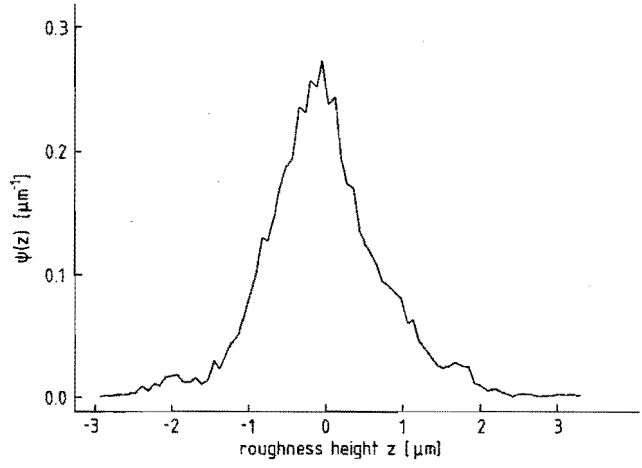
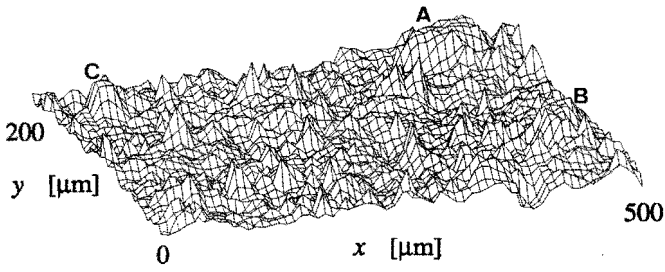
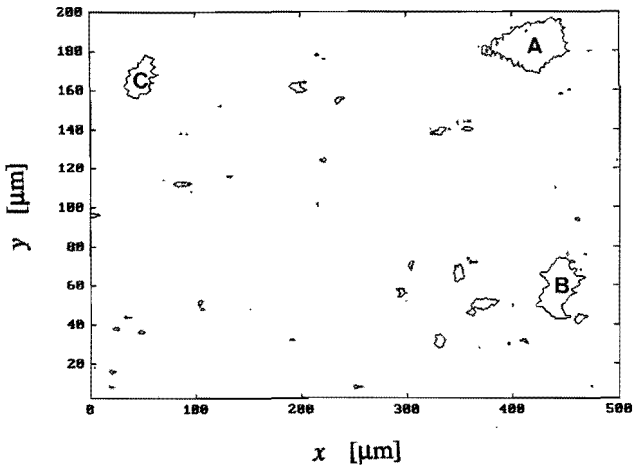


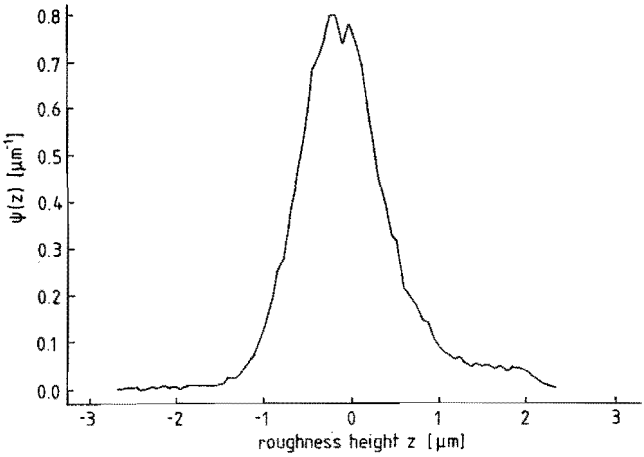
Figure 5.6
Roughness texture of the smoother polyurethane specimen at a contact load of 19 N (p_a = 0.032 MPa).
(with liquid in the contact)



a. Surface plot.



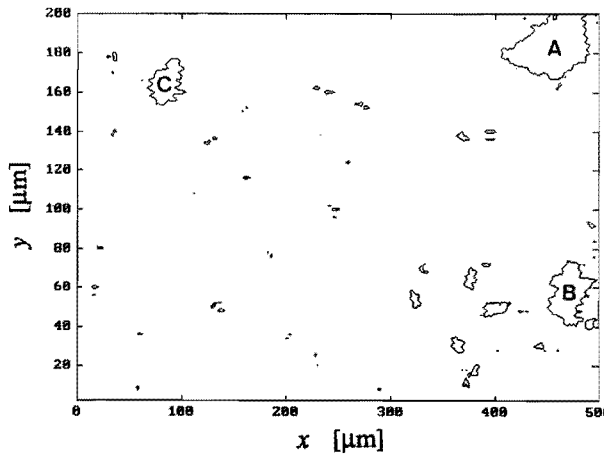
b. Highest height contour lines



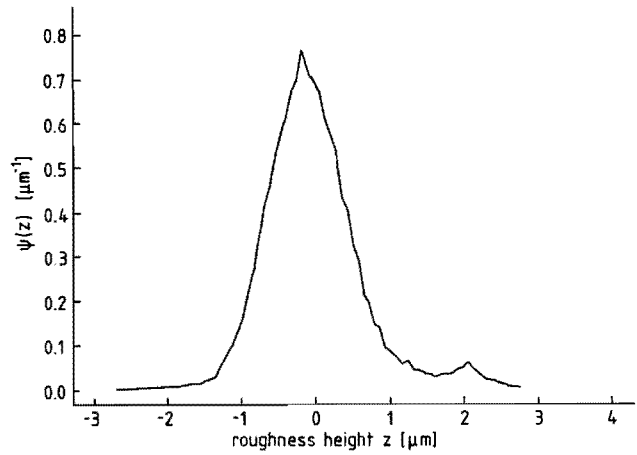
c. Height distribution
R_a = 0.47 µm
R_q = 0.63 µm
Sk = 0.63
Kt = 4.74

Figure 5.7
Roughness texture of the smoother polyurethane specimen at a contact load of 36 N (p_a = 0.060 MPa) (with liquid in the contact)

a. Highest height contour lines



b. Height distribution
R_a = 0.51 µm
R_q = 0.70 µm
Sk = 0.74
Kt = 4.95



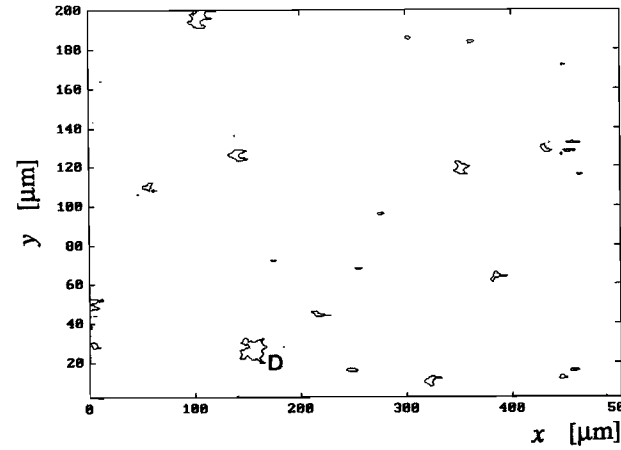
Within the highest contour lines (e.g. the areas A to C shown in fig. 5.6b), the area is flat. Such flat areas were not found in the undeformed texture and therefore, the flatness must be a result of local deformation.

A flattened area can be a real contact, but this is not sure because there can still be liquid between the flattened area and the glass, as discussed for the first measurement above.

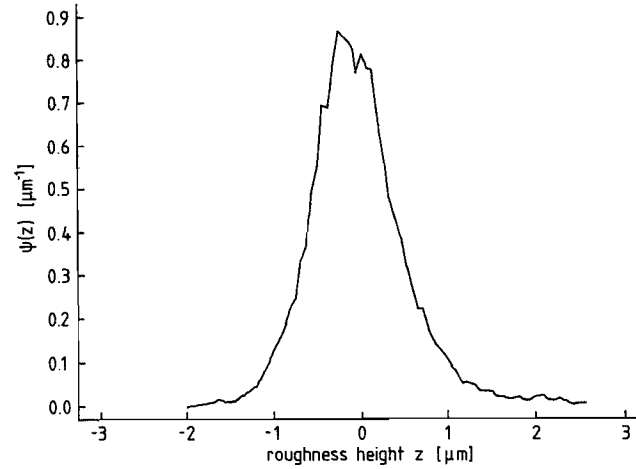
Again, the different succeeding measurements were not performed on exactly the same part of the elastomeric surface, but the corresponding flattened areas in the different measurements can be recognized (some larger corresponding areas in the different height contour plots are indicated by an equal letter).

Figure 5.8
Roughness texture of the smoother polyurethane specimen at a contact load of 53 N (p_a = 0.088 MPa).
(with liquid in the contact)

a. Highest height contour lines



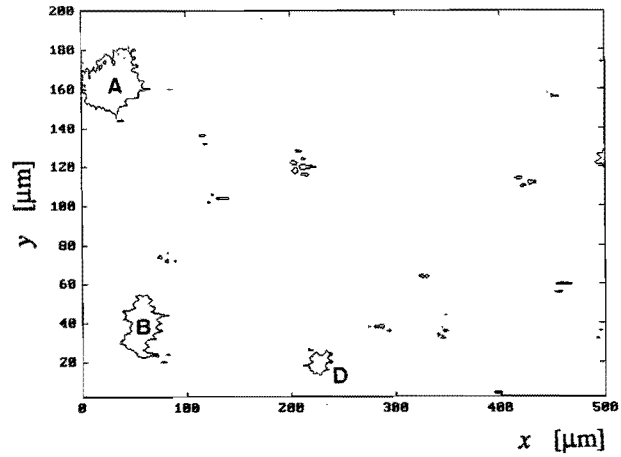
b. Height distribution
R_a = 0.42 µm
R_q = 0.56 µm
Sk = 0.73
Kt = 5.01



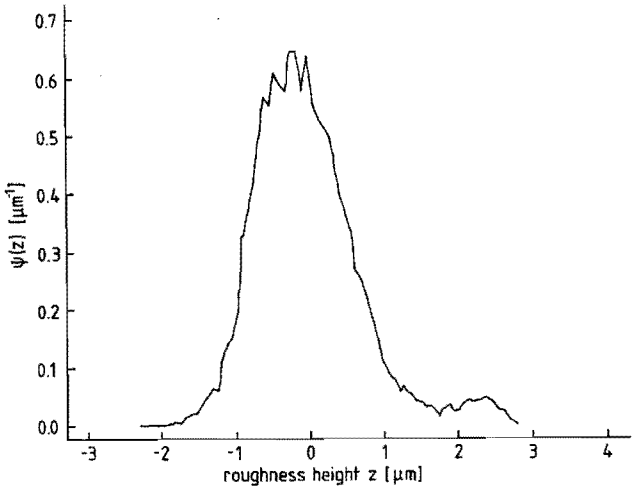
Another point is the "ripple" in the height contour lines, which appears in some measurements (e.g. fig. 5.10a). This is caused by some kind of clearance in the X-direction, yielding a different X-position of the adjoining tracks as explained in appendix C2, but this has in principle no consequence for the determination of the real area of contact.

Figure 5.9
Roughness texture of the smoother polyurethane specimen at a contact load of 70 N (p_a = 0.12 MPa).
(with liquid in the contact)

a. Highest height contour lines



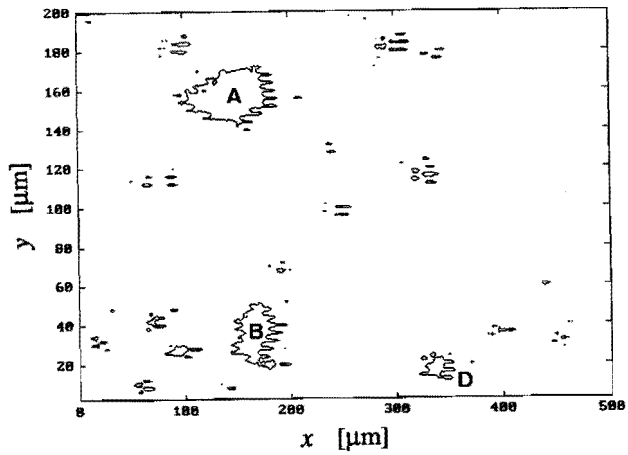
b. Height distribution
R_a = 0.55 µm
R_q = 0.73 µm
Sk = 0.99
Kt = 4.79



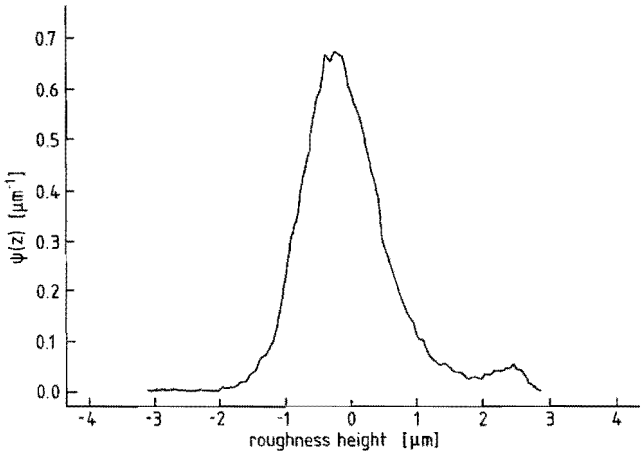
A striking result is, that the size of these corresponding flattened areas is not significantly influenced by the contact load (see fig. 5.11). Also, the roughness height is hardly affected by the contact load: Only the bulk material appears to be deformed and not the roughness summits. This is possible when (the larger part of) the load is supported by the liquid and not by the summits. A further discussion on this matter will be presented by Hendriks (1993).

Figure 5.10
Roughness texture of the smoother polyurethane specimen at a contact load of 87 N (p_a = 0.15 MPa).
(with liquid in the contact)

a. Highest height contour lines



b. Height distribution
R_a = 0.57 µm
R_q = 0.77 µm
Sk = 0.89
Kt = 5.02

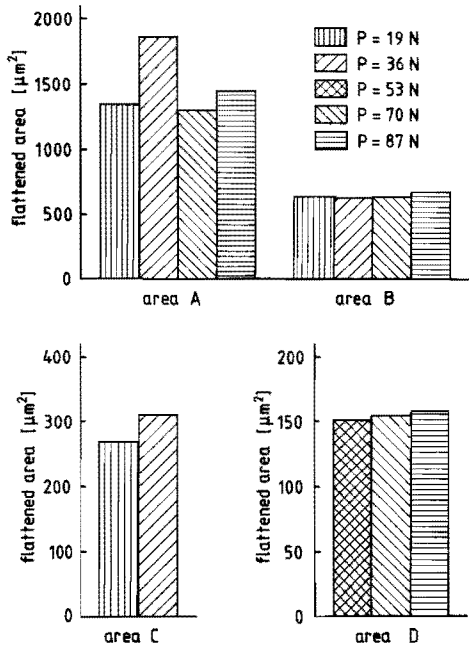


5.4.2 Measurement without a liquid in the contact

The measurements without liquid in the contact area were performed on the smoother polyurethane plate of fig. 5.5 (page 82). The deformed roughness texture is shown in fig. 5.12 to 5.17 for increasing contact loads. The height distribution is not shown and roughness parameters were not derived, because the roughness height variations are not measured properly due to the presence of reflection on the lower glass surface (see section 3.2.5 page 42ff.).

Again, the measurements at the different loads were not performed on precisely the same part of the surface. Consequently, a scratch visible in fig. 5.17 (at position $x \approx 140$ µm) does not appear in the other figures.

Figure 5.11
The area of some contacts shown in fig. 5.6 to 5.10.



These measurements, especially the latter three (fig. 5.15 to 5.17), clearly show increase in the surface flattening when the contact load is increased. The linear dimension of these areas is of the order of 10 µm. Besides, roughness height variations of about 0.5 µm were found within the flattened areas (see e.g. fig. 5.18). This is significantly larger than the roughness height of the glass plate, which is within 0.02 µm. This gives evidence to the idea that the flattened areas, at a 10 µm length scale, are not real "real areas of contact". Instead, smaller scale "real areas of contact" (at a length scale of about 1 µm) can be found in these flattened areas. Analogous to the idea of fractal roughness characterization (section 5.1.1 page 71) more contacts of smaller length scales are perhaps present within the contacts on a 1 µm length scale. Consequently, a derived real area of contact is always related to a particular length scale.

The simultaneous existence of both clearly flattened "long scale asperities" and remaining "small scale asperities" on the flattened surfaces can possibly be explained by difference in asperity interaction at different length scales. Majumdar and Bhushan (1990) found e.g. that the roughness of machined surfaces is fractal at small length scales, but non-fractal at the larger length scales,

Figure 5.12
Roughness texture of the smoother polyurethane specimen at a contact load of 2 N (p_a = 0.003 MPa).
(without liquid in the contact)

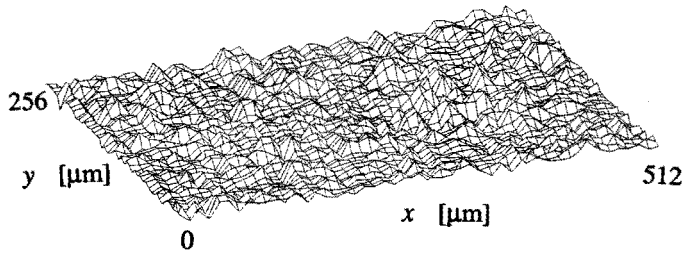


Figure 5.13
Roughness texture of the smoother polyurethane specimen at a contact load of 19 N (p_a = 0.032 MPa).
(without liquid in the contact)

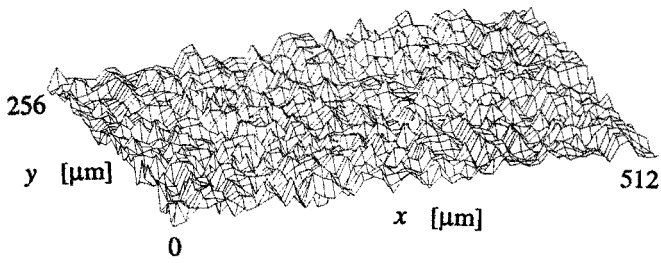


Figure 5.14
Roughness texture of the smoother polyurethane specimen at a contact load of 36 N (p_a = 0.060 MPa).
(without liquid in the contact)

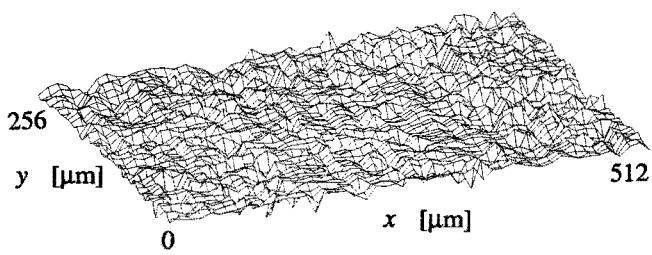


Figure 5.15
Roughness texture of the smoother polyurethane specimen at a contact load of 53 N (p_a = 0.088 MPa).
(without liquid in the contact)

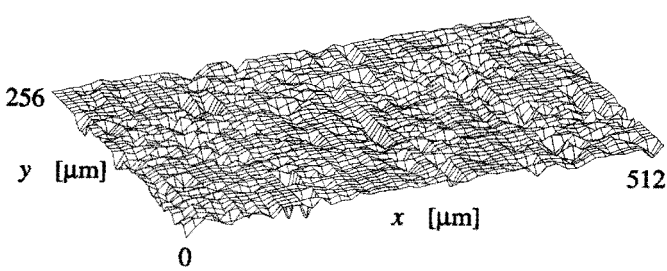


Figure 5.16
Roughness texture of the smoother polyurethane specimen at a contact load of 70 N (p_a = 0.12 MPa). (without liquid in the contact)

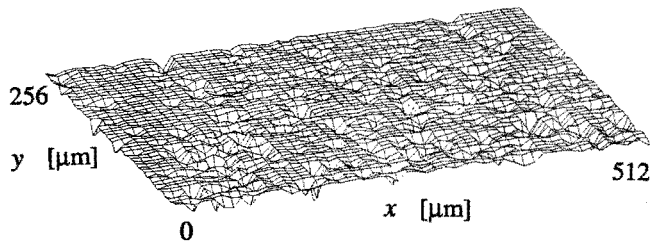


Figure 5.17
Roughness texture of the smoother polyurethane specimen at a contact load of 87 N (p_a = 0.15 MPa). (without liquid in the contact)

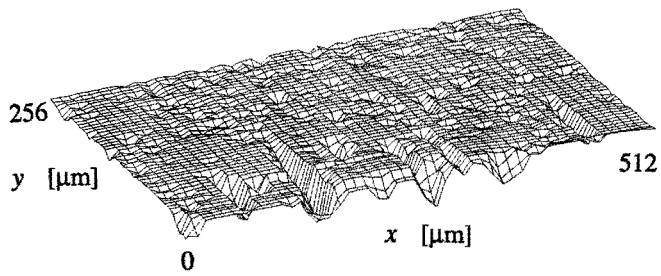
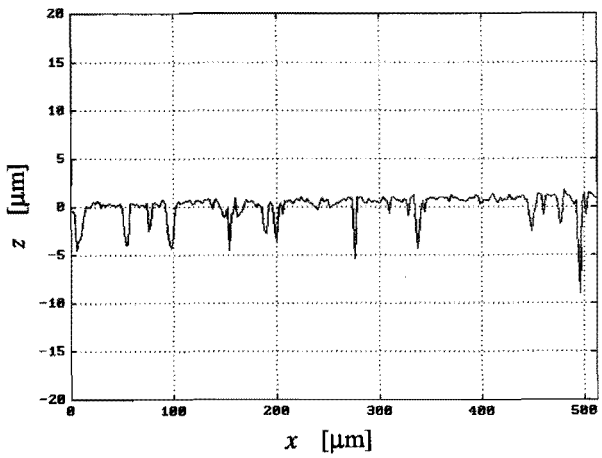


Figure 5.18
Cross section of fig. 5.16 at y = 246 μm.



where the asperity height to wavelength ratio is smaller. The same can be true for an injection moulded elastomeric surface. Consequently, the distance between the larger scale asperities is perhaps large compared with their height and in that case, the asperity interaction (section 5.1.1 page 72) can be small, yielding large flattened area. Otherwise, the distance between the smaller scale

asperities can be smaller compared with their height, yielding a large asperity interaction.

The number and size of the small scale contacts is not easily determined, because of the limited spatial resolution (about 1 μm) of the transducer, but a preliminary estimation of the real area of contact on a 10 μm length scale is given in fig. 7.19. Although the curve is not a straight line, the real area of contact A_r is more or less proportional to the load P. However, it must be noted here that the proper evaluation of the real area of contact is still subjected to investigation. A brief discussion on this matter is given in section 5.5 below.

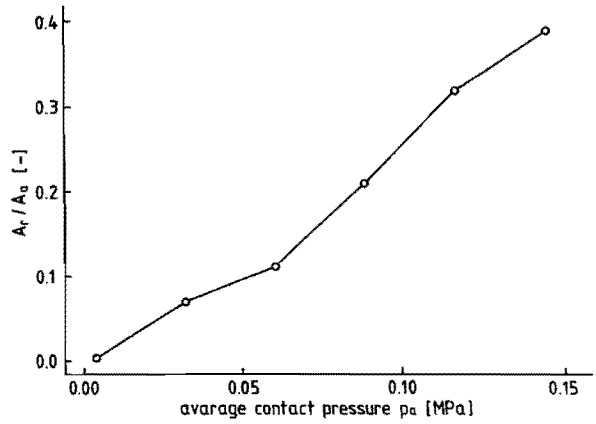


Figure 5.19
Estimation of the ratio of the real area of contact A_r and the apparent area of contact A_a.

5.5 Discussion on the measurement of the real area of contact

Now we will briefly discuss some matters which are important in measuring the real area of contact. This discussion will provide a first idea of the method's capability for such measurements, but a more elaborate investigation is still in progress and will be reported later by Hendriks (1993).

If one is interested in measurement of the real area of contact, e.g. for experimental verification of contact models reviewed in section 5.1.1, it must be considered that "areas of contact" can be present at different length scales, i.e. within a flattened "contact area" at one length scale, smaller scale contacts can be present. Before starting a measurement, the length scale(s)

important for the problem under study must be determined and the measurement parameters should be chosen accordingly, analogous to the idea of functional filtering, proposed by Thomas and Sayles (1978) and by Thomas (1982) for surface roughness measurements (see also appendix A1).

Considering the optical focus error method presented in this thesis, the smallest length scale which can be measured depends on the focus spot size and is about 1 μm. This would implies that only real contact areas larger than some micrometers can be detected. Whether smaller scale contacts can be distinguished within these flattened areas can then be derived from the roughness height variations within these flattened areas, as indicated in section 5.4 above, but the size of these smaller scale contacts is not easily determined.

Considering the measurements presented in section 5.4, the reproducibility appears to be still questionable. One factor is the apparently large differences in the results of measurements on a different part of the surface, as shown by comparison of fig. 5.7 and 5.8: The large flattened areas A and B are found in one measurement and not in the other. On the other hand, the estimated dimension of one and the same flattened area can also vary for different measurements. The estimated size of area A is e.g. significantly larger in the second measurement in the series of fig. 5.6 to 5.10 (at a load P of 36 N) than in the others (see fig. 5.11). This can hardly be caused by the different loads, because the estimated area is almost equal for the other measurements, both at lower and at higher loads.

5.6 Conclusions

The optical surface roughness transducer is helpful to study the deformed roughness texture of soft materials under static load, both with and without a liquid in the contact area. The preliminary conclusions from the measurements will be listed below for the wet and the dry situation separately.

The wet contact

A liquid, present in the contact area, can influence the contact problem. There is evidence from the measurements that a significant part of the applied load is supported by the liquid, e.g. due to micro-squeeze effects, since flattened but

non-contacting asperities were present. It was also found for a relatively smooth elastomer, that increase of the contact load did not yield significant deformation. This also indicates that the liquid supports a part of the load, but more investigation is needed to get better insight into the influence of the liquid.

The dry contact

In the dry contact situation, only the real contact areas can be derived from the measurements. However, the measurement of the real area of contact is still a problem, since contact areas of smaller length scale seem to be present within the flattened areas (at a length scale of 10 μm). The size of these smaller areas can not be determined accurately because of the limited spatial resolution of the transducer.

A more elaborate investigation of the method's capability to determine the real area of contact is under progress and the results will be presented in the near future.

CHAPTER 6 CONCLUSIONS

The main scope of this thesis is the development of a film thickness transducer which has a sufficient resolving power to determine the eventual deformed roughness texture during lubrication. The method must be applicable to a rough elastomer in sliding contact with a smooth rigid body.

Discussion of different methods yielded the conclusion that focus error detection, as e.g. used for compact disc reading and for contactless surface roughness measurement, is in principle an appropriate method for these measurements. Further investigation of this method yielded the main conditions which must be fulfilled for these measurements.

Not all conditions are easily fulfilled and more investigation is therefore needed. The most important items are the contact pressure influence and the temperature influence on the measurement. Further, a new electronic device for the signal handling must be developed for the following reasons:

- The frequency range must be increased to 1 MHz to obtain the required response time.
- The simultaneous measurement of both the focus error and the radial error signal must be possible.
- The laser power and/or the signal to noise ratio must be increased because of the low reflectance on the lubricant to elastomer interface.

The latter item, considering the low reflectance on the lubricant to elastomer interface, also needs further investigation.

Preliminary measurements showed that the method performs well, at least qualitatively. Quantitative measurements have not been performed yet, mainly because of the slope influence on the focus error signal which can not be eliminated using the present electronic device.

Focus error detection can also be used to determine the surface roughness deformation in a statically loaded contact of a glass plate and an elastomer. Such experiments were performed using the existing rig for contactless surface roughness measurements. Two series of measurements are presented:

- One with a liquid in the contact;
- the other with a dry contact.

The former series, with liquid between the glass plate and the elastomer, yielded proper measurement of the deformed roughness texture, because the reflection on the glass surface was eliminated by the liquid, and the following features were observed:

- The presence of a liquid can influence the contact situation significantly.
- Flattened but non-contacting asperities can exist, probably due to micro-squeeze effects.

The other series, without a liquid in the contact, enables the study of the real contact areas. It was found that "real areas of contact" can be found at different length scales, i.e. small scale contacts are present within a flattened area of roughly 10 μm in size. Then the total real area of contact, derived from a measurement, depends on the length scales included in the measurements. Further investigation in in progress to determine the real area of contact properly, considering the length scales important for a particular application.

APPENDIX A SURFACE ROUGHNESS CHARACTERISTICS

In this appendix the characterization of the surface roughness will be briefly discussed and some measurements on a reciprocating seal will be presented to obtain an indication of a roughness texture typical for seals.

The chosen roughness characterization is based on random process theory. Other methods, like the use of standardized roughness parameters, can also be used in principle, but they do in essence not provide more information.

A1 Surface roughness characterization

An elaborate outline of two-dimensional[1] surface roughness characterization is provided by Halling (1978 pp. 22-39) and by Thomas (1982 chapter 5). Both two- and three-dimensional characterization are discussed by Hendriks (1992[a]). Often, only roughness values (e.g. an average roughness height and a peak to valley height) are derived from a measurement, but this is not always sufficient.

In general, the **height distribution** curve gives more information about the roughness, since it clearly shows, e.g., whether the peaks are high or low (this is often important in contact problems, since the peaks are the first contacting parts of a surface).

Besides, information on the characteristic **length scales** of roughness height variations can be important. In this thesis the importance is e.g. that the spatial resolution of the film thickness transducer should be as small as the smallest characteristic length scale in the roughness texture, in order to obtain proper measurements of the (eventual deformed) surface roughness of the elastomer under running conditions. The characteristic length scale(s) can be obtained from the **autocorrelation** function and from the **autopower spectrum**.

Another point is, whether the roughness texture is isotropic or not. A roughness texture is isotropic, when the roughness characteristics are equal in all directions. This can be derived from the two-dimensional autocorrelation or autopower spectrum, derived from a three-dimensional measurement. The texture is

[1] Two-dimensional characterization is based on a roughness measurement along a single line. Three-dimensional characterization needs a scanning of the roughness texture over an area, e.g. by measurement along a number of adjoining lines.

isotropic when the plots of the autocorrelation and of the autopower spectrum are symmetrical with respect to the origin. However, two-dimensional measurements are more common than three-dimensional. Then the texture can be regarded as isotropic when the autocorrelation curve and the autopower spectrum are similar for measurements in different directions.

Finally, it must be considered which length scale should be regarded as "characteristic" and which not, since it is practicably impossible to measure all length scales present in a roughness texture[2]. Thomas and Sayles (1978) and Thomas (1982) therefore introduced the idea of **functional filtering**, i.e. the measurement parameters (spot size, sample distance and measurement length) are appropriately chosen to include the most important length scales in the measurement, while the less important length scales are rejected. If one is e.g. interested in the full film lubrication of rough surfaces, it can be considered that the roughness height variations which are of equal order as the nominal film thickness are the most important and the right measurement parameters will be chosen accordingly, to include the length scales of these variations in the measurement. Otherwise, studying the dry friction between two surfaces, the smaller roughness height variations at smaller length scales (within the real areas of contact) are more important and the measurement parameters for the roughness measurement should be chosen accordingly.

Now the height distribution, some characteristic values derived from it, the autocorrelation and the autopower spectrum will be discussed in some more detail.

---

[2] The smallest length scale measured is determined by the spatial resolution of the scanning device, i.e. by the size of the measurement spot, and by the sample distance of the scanning, while the largest length scale is determined by the measurement length. When all length scales should be included in one measurement, the sample distance should be about 1 nm, while the measurement length must be equal to the size of the surface considered (e.g. 10 mm), i.e. the measurement should include e.g. 10^7 points for two-dimensional characterization and 10^{14} for three-dimensional characterization. Processing of these data is very time consuming and therefore impossible.

The height distribution and derived roughness values

An example of the height distribution curve (or: probability density function $\psi(z)$ of the roughness height z)[3] is shown in fig. A1. Some values can be derived from the height distribution, giving an indication of the shape of the distribution curve:

- The Centre Line Average (CLA) roughness height (or arithmetic roughness height)

$$R_a = \int_{-\infty}^{\infty} |z| \, \psi(z) \, dz \tag{A1}$$

- The Root Mean Square (RMS) roughness height (or the standard variation of the roughness height distribution)

$$R_q = \left[\int_{-\infty}^{\infty} z^2 \, \psi(z) \, dz \right]^{1/2} \tag{A2}$$

- The Skewness

$$Sk = \frac{1}{R_q^3} \int_{-\infty}^{\infty} z^3 \, \psi(z) \, dz \tag{A3}$$

- The Kurtosis

$$Kt = \frac{1}{R_q^4} \int_{-\infty}^{\infty} z^4 \, \psi(z) \, dz \tag{A4}$$

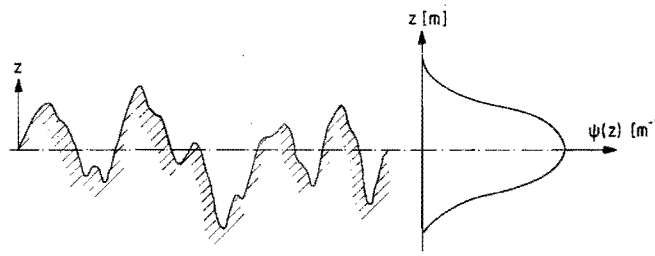


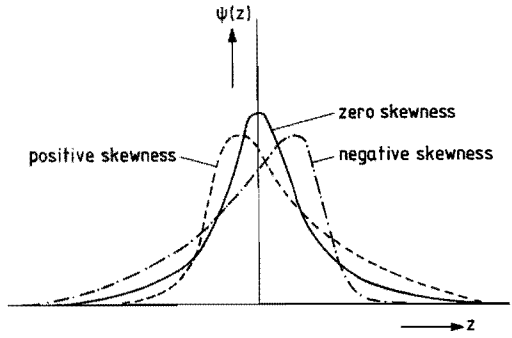
Figure A1
Example of a height distribution curve.

[3] The roughness height z is defined as the local height distance between the surface and the mean line (or the mean plane for a three-dimensional measurement). Consequently: $\int_{-\infty}^{\infty} z \, \psi(z) \, dz = 0$

The R_a and the R_q value indicates the width of the distribution curve, i.e. whether the roughness height variations are large or small. The peak to valley roughness height (z_{max} - z_{min} in fig. A1) could also be used to obtain this information, but is very sensitive to one single extreme in the roughness height (e.g. one single scratch) which is not representative for the whole surface.

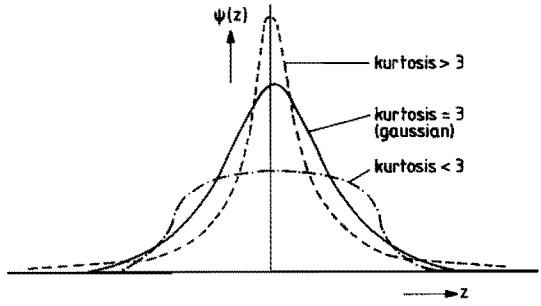
The skewness indicates whether the height distribution is symmetrical or not (fig. A2). A positive skewness means, that the peaks are relatively high compared with the depth of the valleys, while otherwise a negative skewness indicates that the peaks are relatively low. For a symmetric distribution curve (e.g. a Gaussian) the skewness is 0.

Figure A2
The skewness.



The kurtosis indicates whether high peaks and/or deep valleys are present with respect to the average roughness height (fig. A3). A large kurtosis means, that there are peaks which are relatively high and/or valleys which are relatively deep. The kurtosis of a Gaussian height distribution is 3.

Figure A3
The kurtosis.



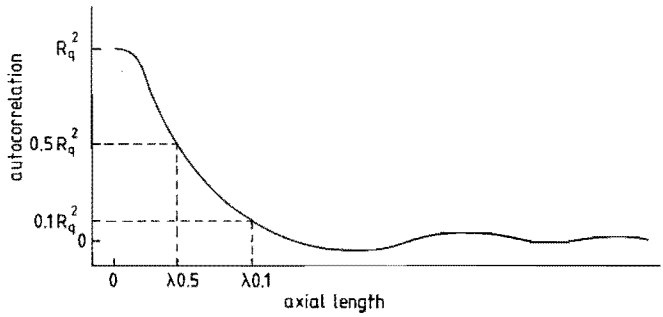
The autocorrelation and autopower spectrum

As already mentioned above, the characteristic length scales of a roughness profile can be derived from the autocorrelation curve and from the autopower spectrum.

When the roughness profile resembles a sinusoidal shape with wavelength λ (i.e. the characteristic length scale), the autocorrelation curve is also sinusoidal with the same wavelength. Its Fourier transformed function, the autopower spectrum, then has a clear peak at the spatial frequency $f = 1/\lambda$.

If the roughness texture is more or less random, the autocorrelation curve rapidly falls to zero (fig. A4) and the autopower spectrum shows a wide range of frequencies. Then the shortest length scale of importance can be derived from the autocorrelation function, e.g. using the 50 percent correlation length $\lambda_{0.5}$ or the 10 percent correlation length $\lambda_{0.1}$. These correlation lengths are defined by the horizontal distance over which the correlation between the points is 50 respectively 10 percent of the maximum value of the autocorrelation (fig. A4). This maximum value is equal to the variance of the roughness height distribution ($= R_q^2$).

Figure A4
The autocorrelation curve for a more or less random roughness texture and the 50 and 10 percent correlation length ($\lambda_{0.5}$ and $\lambda_{0.1}$ respectively)



A2 Surface roughness characteristics of seals

The measurements are performed on a polyurethane rod seal manufactured by Parker-Prädifa (seal code B3 U28 5004) which is made from the same material as the seals used in the studies of Kanters and Visscher (1989); Kanters, Verest and Visscher (1991) and Kanters (1990, 1991). Other seal types, especially seals made from a different material, may have a different roughness texture, but the

measurements presented here are meant to provide a first idea of the roughness height distribution and the characteristic length scales.

Only two-dimensional measurements are performed, both in the axial direction (i.e. the direction of motion of the rod relative to the seal) and in the tangential (circumferential) direction. The plots derived from these measurements are presented in fig. A5 and A6 respectively, while the derived roughness values (averaged over 5 measurements) are shown in table A1.

The measurements were performed with the optical device of Struik and Chang (1987) (see also section 3.1.2 page 26). The following parameters were used:

- Diameter measurement spot: $\approx$ 1 μm;
- Sample distance: 1 μm;
- Measurement length: 2 mm.

Additional filtering of the measured data was not applied.

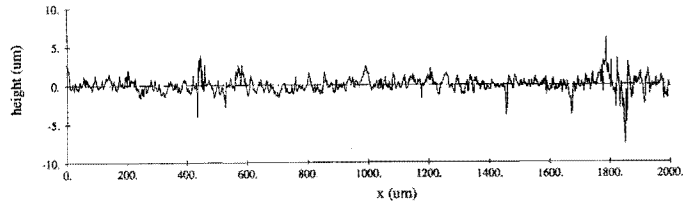
The surface slope distribution is also derived from the measurement and shown in fig. 1.5 and 1.6. In general the derived surface slope largely depend on the measurement parameters, as e.g. discussed by Thomas and Sayles (1978), but the significance of their mentioning here is to obtain an idea how large the slopes can in principle be. This is because application of some methods, considered for the film thickness measurements and discussed in appendix B, is limited to surfaces with small slopes.

The figures 1.5 and 1.6 show, that the largest part of the roughness height is between -2 and +2 μm, but there are clearly some high peaks (up to about 6 μm high) and deep valleys (up to about 8 μm deep). The kurtosis is therefore high, compared with the kurtosis of a Gaussian distribution.

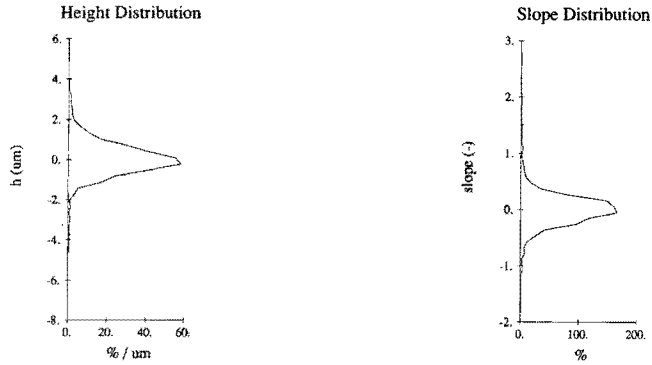
Considering the length scales, the autopower spectrum shows that spatial frequencies up to about 100 mm^{-1} are significantly present in the axial direction, i.e. the smallest characteristic wavelength is roughly 10 μm. This corresponds well with the correlation lengths derived from the autocorrelation. In the tangential direction, the maximum significant frequency in the autopower spectrum is roughly 50 mm^{-1}, i.e. the smallest characteristic length scale is about 20 μm. This also corresponds with the derived correlation lengths.

The difference in the frequency range, shown in the autopower spectra for the axial and the tangential direction, and the difference in correlation lengths indicate that the roughness texture is not isotropic.

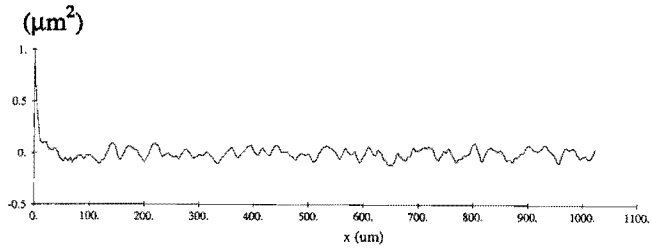
Figure A5
The surface roughness characteristics of the rod seal.
(in axial direction)



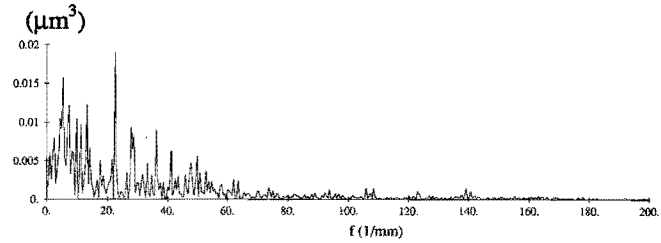
a. The profile



b. Height and slope distribution.



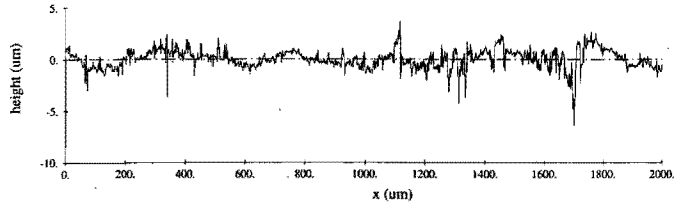
c. Autocorrelation curve.



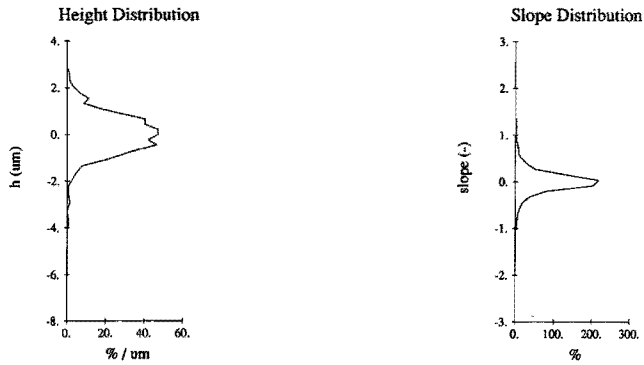
d. Autopower spectrum.

Figure A6
The surface roughness characteristics of the rod seal. (in tangent. direction)

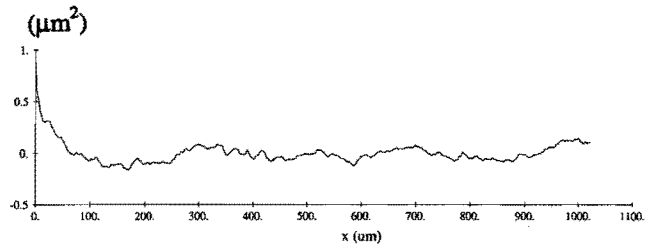
a. The profile



b. Height and slope distribution.



c. Autocorrelation curve.



d. Autopower spectrum.

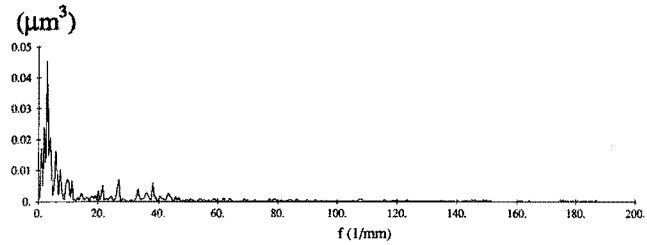


Table A1 Roughness values of the rod seal.

	R_a [μm]	R_q [μm]	Sk [-]	Kt [-]	$\lambda_{0.5}$ [μm]	$\lambda_{0.1}$ [μm]
axial	0.619	0.884	-0.84	12.93	4.8	10.8
tangential	0.689	0.983	-1.4	7.83	8.1	60

APPENDIX B REVIEW AND DISCUSSION ON METHODS FOR FILM THICKNESS MEASUREMENT ON ELASTOMERS

Reviewing the literature, it is shown that various methods have been applied for film thickness measurements (see e.g. Visscher and Kanters (1990) and the following of this appendix). The choice of an other method may be expected to be the result of different requirements. However, these matters are often obscure. There is mostly no indication of differences in the requirements and generally, the motivation of the choice of a particular method is very poor. Sometimes, the motivation is lacking and sometimes, the only motivation is e.g. that "many methods, mechanical, electrical and optical, were tried and only the chosen one was successful" (Jagger, 1957) or "the chosen one was considered to be the most promising" (Poll and Gabelli, 1992[a]). Thus, information about the development of the method is not provided and it is therefore very difficult to compare the tests of the different investigators and to find out subsequently, why a method, successful to one investigator, was poor to the other.

Secondly, the method used is often not analysed and the accuracy not determined. When some considerations concerning the accuracy are made, they are mainly based on assumptions which are often not verified. Some examples are given by Visscher and Kanters (1990) and in the following of this appendix (e.g. in section B2.1.2 page 122).

In spite of more than 30 years of research, it still appears to be a question, which method should be applied and what accuracy can be achieved. The literature shows no real progress and does not offer a kind of a guideline which can be applied by new investigators to get the best film thickness transducer possible for their application. Instead, everybody must start at the same "zero point", where Jagger (probably the first who measured the film thickness of elastomeric elements) had to start.

Bearing these facts in mind, we will now review and discuss the various methods. Practical information on the application of the methods, as found in literature, will be given. It is, however, not the main objective to judge the value of the measurements given in literature, but to discuss the problems and the possibilities of the methods. **Consideration of the factors which influence the measurement, and thus reduce the accuracy, is important.** Attention will therefore be paid to find these factors, to estimate their quantitative

influence when possible, and to discuss how their influence can be minimized or, may be, eliminated.

We will consider the method's suitability for film thickness measurements in general as the most significant criterion for the final choice which will be made, regardless of the required spatial resolution. The reason is that a method, which can not match the required spatial resolution, can still be helpful in studying the eventual roughness deformation in the lubricated contact by using a roughness texture with longer characteristic wavelengths. Otherwise, a method will never be suitable, when it is not applicable to elastomers, even though if the spatial resolution is very high.

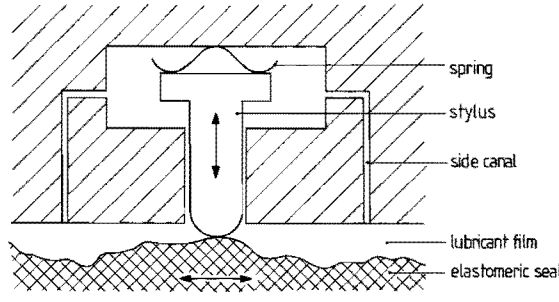
The possible spatial resolution will then be discussed in a second stage and finally, a method which is believed to be the most appropriate will be chosen.

B1 Mechanical methods

A mechanical method was used by Schrader (1978 pp. 54ff.) to measure the film profile of elastomeric piston seals. A stylus with a top radius of 0.15 mm was pressed on the seal surface by a spring with a force of about 0.015 N (fig. B1). The height position of the stylus was taken as a measure for the film thickness. Side canals are provided to balance the film pressure load on the stylus.

An important influence factor is the mechanical load on the seal's surface, which will deflect the surface. This surface deflection is measured as a virtual extra film thickness (i.e. the measured thickness is the sum of the real film thickness and the additional deflection from the stylus load) and should therefore be very small. Considering the device of Schrader, the seal deflection

Figure B1
Mechanical device for the film thickness measurement, used by Schrader (1978).



is estimated, using Hertzian theory, to be about 5 μm for an E-modulus of 100 MPa and a Poisson ratio of nearly 0.5, which are common values for elastomers. Using a smaller stylus (tip radius e.g. 5 μm) to match the required resolution, the deflection will be also about 5 μm (applying a radial load of 0.004 N as is standard for normal roughness measurements according to ISO 3274), which is obviously too large.

We must also bear in mind, that the method disturbs the flow at the measurement spot (page 20, requirement no. 2), even when the deflection due to the mechanical load of the stylus would be negligible small. This is caused by local influence of the film pressures by the presence of the transducer and the asperity deformation due to micro-EHL will therefore also be influenced.

Finally, the response time (page 20, requirement no. 6) will be too long, because of the inertia of mechanical systems.

It can be concluded that a mechanical method is not appropriate because of the mechanical load of the stylus, the disturbance of the fluid film by the stylus and because of the slow response of the transducer.

B2 Electrical methods

Electrical film thickness measurements, with an electrode on one surface and using the counter face as the other electrode, are well known from research programmes on lubricated metal to metal contacts. The widely available knowledge of and experience with these methods is perhaps a reason for their frequent application to elastomer to metal contacts, in spite of the very low conductance of elastomers. The elastomers are therefore commonly filled with conducting particles, like carbon black, to obtain the required conductance.

Alternatively, a capacitor consisting of two band electrodes side by side on the rigid surface can possibly be applied. Its capacitance is then influenced by the film thickness when the electrical permittivity of the lubricant and of the elastomer are different. The use of a conducting counter face is then not required.

In this section, we will firstly consider the suitability of electrical methods for film thickness measurements in general, not considering the requirement for the roughness detection. Both the use of the elastomeric surface as an electrode

(section B2.1) and the possibility of the use of two electrodes on the rigid surface (section B2.2) will be considered. Finally, the possibilities will be discussed to use electrical methods to detect the elastomer's surface roughness during motion (section B2.3).

B2.1 The use of the elastomeric counterface as electrode

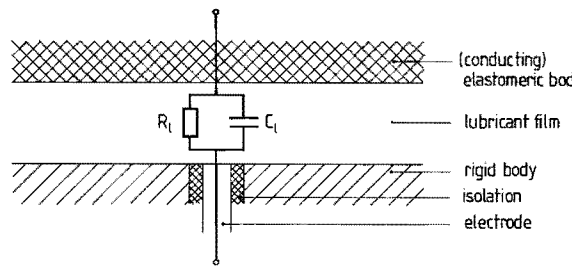
In the configuration which will now be considered, one electrode is fixed on or in the rigid surface (requirement no. 1, page 20) and the elastomeric counter face is the other electrode (see fig. B2). The lubricant film can then be modeled by a resistance (R_l) and capacitor (C_l) and can be determined by measuring the electrical impedance over the lubricant film. This impedance is given by

$$Z_l = \frac{1}{\frac{1}{R_l} + 2\pi j f C_l} \qquad \text{(B1)}$$

in which

R_l	= lubricant film resistance = $\rho h/A$	[Ω]
C_l	= lubricant film capacitance = $\varepsilon_0 \varepsilon_r A/h$	[F]
f	= frequency in the signal	[s^{-1}]
A	= electrode area	[m^2]
h	= film thickness	[m]
ρ	= specific resistance of the lubricant	[Ωm]
ε_0	= electrical permittivity of vacuum = $8.9 \cdot 10^{-11}$	[Fm^{-1}]
ε_r	= relative electrical permittivity of the lubricant	[-]

Figure B2
Configuration for electrical film thickness measurement and representation of the lubricant film resistance and capacitance.



In literature, methods are generally divided in resistive and capacitive methods. The background is, that the electrical impedance Z_l, defined in eq. (B1), is almost insensitive for the capacitance C_l, when the frequency f is low (e.g. lower than 100 Hz, depending on the total design of the electrical instrumentation). Otherwise, the influence of the resistance R_l on the impedance is negligible for high frequencies (e.g. 1 MHz).

Both resistive and capacitive methods will be reviewed and discussed below and we will find, that the capacitive methods are much more appropriate.

We must consider, that this way of film thickness measurement requires the use of a conducting elastomer or the use of a conducting coating on the contacting surface of the elastomer. The minimum required conductance (or the maximum tolerable resistance, which is the reciprocal value of the conductance) will therefore be estimated and the use of conducting elastomers and conducting coatings will be discussed.

Besides the elastomer's conductance, the local stiffness of the electrode and its isolation, fastened in the rigid body, need specific attention (requirement 3, page 20). Field (1973 p. 295) calculated that the electrode inflection can be about 0.25 μm at a pressure of only 7 MPa. This point will, however, not be discussed here, as it is considered to be a second order effect which can in principle be solved with appropriate techniques.

B2.1.1 Resistive methods

Wernecke (1983, 1987) used the resistive method to measure the film thickness of reciprocating rod seals. As discussed by Visscher (1989 pp. 11), the influence of the seal resistance is probably negligible, since the seal resistance is small compared with the lubricant film resistance R_l (e.g. 0.03 and 5 MΩ respectively).

On the other hand, Visscher (1989 pp. 10-12) found that the influence of the oil film capacitance C_l can be serious, since the frequency f in the lubricant film impedance Z_l (eq. B1) appears to be determined by the film profile and by the velocity of the seal, moving relative to the electrode mounted in the rod. What happens is, that the film thickness (and thus the resistance R_l which is proportional to the film thickness) at the electrode position varies in time. This variation is then determined by Wernecke by measuring the voltage over the lubricant film, which, in his electrical device, depends on the resistance R_l. The capacitance C_l now introduces a filtering

effect: a film thickness variation with a higher frequency has a smaller effect on the voltage over the lubricant film, since the capacitor needs time for charging and discharging.

At higher velocities (1 m/s) the 5 mm wide seal crosses the electrode in a time not longer than 5 ms. Film thickness variations with a length scale of 0.5 mm (10 percent of the contact width), result then in a frequency f of 2000 Hz. At this frequency, the lubricant film capacitance C_l has a significant influence and the measured film thickness variations are not more than about 50 percent of the real film thickness variations. At higher frequencies, the capacitance influence increases and smaller scale film thickness variations are therefore not detected. One consequence is, that the roughness on the elastomeric surface is hardly detected, even when a simple shaped regular roughness texture with a rather long wavelength is used. The resistive methods are therefore not suitable.

B2.1.2 Capacitive methods

Jagger (1957) applied the capacitive method by amplitude modulation, measuring the film thickness over the whole contact area of an axial elastomeric face seal (seal lip diameter 140 mm and contact width about 1 mm).

The measurements were performed using a supply voltage with frequency of 93 [kHz], assuring that frequencies arising from film thickness variation (being at least one order of magnitude smaller) did not influence the measurement. (The impedance Z is now determined by the frequency in the supply voltage and not by the frequencies of the film thickness variations). Otherwise, Visscher (1989 pp. 17-19) found that the influence of the seal resistance can be large. The resistance of the seal nor the specific resistance of the used material were given by Jagger, but values of about 15 Ωm and 3 Ωm are reported in literature (see e.g. Swales et al. (1972) and Field (1973) respectively). The error in the film thickness measurement could have been about 3 percent at a film thickness of 1 μm for a specific resistance of 10 Ωm, but increases more than proportional with the specific resistance (e.g. 10 percent at 20 Ωm). However, the increase of the error with decreasing film thickness is more remarkable: The measured film thickness is more than 100 percent larger than the real film thickness, when the real film thickness is about 0.1 μm and when the specific resistance is 10 Ωm!

Field (1973) and Field and Nau (1973[b]) investigated the lubrication of reciprocating seals. They measured the lubricant film capacitance by means of frequency modulation (with frequencies of the order of 1 MHz) using two types of seal materials: one with a specific resistance of 2.78 and and the other with a specific resistance of 4.76 Ωm. Field (1973 p. 295) calculated the influence of the seal resistance on the film thickness measurement and concluded, that the influence was small for a seal resistance of 200 Ω. However, no indication of the real seal resistance was given and Visscher (1989 pp. 25) reported that the resulting total seal resistance was probably about 1.75 and 3 kΩ for both seal materials respectively. In that case, the errors in the measurements of Field and Nau must have been about 20 percent at a film thickness of 1 μm and more than 100 percent at a film thickness of 0.1 μm (see Visscher, 1989 pp. 22-25).

The maximum allowable seal resistance

Visscher (1989) also performed some analysis to estimate the maximum tolerable seal resistance for both amplitude modulation and frequency modulation. The calculations were performed for the use of an electrode of 1 mm diameter and requiring a maximum uncertainty in the film thickness measurement of 10 and 1 percent for a film thickness of 0.1 and 1 μm respectively. It was estimated, that the resistance should be of the order of 1 kΩ at the maximum for amplitude modulation (the specific resistance of the seal must then probably be of the order of 1 Ωm or smaller) (Visscher, 1989 pp. 19-22). For frequency modulation, the maximum allowable seal resistance was estimated to be one order of magnitude smaller (Visscher, 1989, pp. 25-28).

Now that we have estimated the maximum allowable value of the seal resistance (or the minimum required conductance, which is the reciprocal value of the resistance), we will now discuss, how this conductance can be obtained using elastomers. There are two options for this purpose:
- The use of conducting elastomers;
- The use of a conducting coating on the contacting elastomer surface.

Both options will be discussed below.

Conducting elastomeric materials

As mentioned above the maximum allowable resistance is of the order of 1 kΩ for a capacitive method using amplitude modulation and with an electrode diameter of

1 mm. This yields e.g. a maximum allowable specific resistance of the order of 1 Ωm when the seal height is 5 mm (Visscher, 1989 p. 10), i.e. the minimum required conductance is of the order of 0.1 Sm^{-1}. Elastomers commonly do not fulfil this requirement and special measures are necessary.

Generally, the elastomer is filled with conducting particles like carbon black or metallic particles. Alternatively, real conductive polymers are available nowadays. Both kinds were subjected to a literature review by de Jong (1990) and a summary is given here.

Conductive polymers, also known as "synthetic metals", may have a specific conductance of up to 10^7 Sm^{-1}. However, their use for e.g. (test-)seals is still difficult, mainly because of the impossibility to mould them and because of the chemical instability. More information of these materials is provided by Kusy (1986 pp. 46-65).

Metal- or carbon black- filled polymers have been widely used to increase the conductance. Distinction can be made between fibres, which are long compared with their thickness, and particles, which have a roughly equal length and thickness.

Fibres can easily form a network and are therefore preferred to achieve a high conductance. Otherwise, their influence on the elasticity modulus can also be very large (about 10 to even 1000 times!, see e.g. Chow and Penwell, 1986). Such an increase in stiffness is generally not desirable, because the mechanical behaviour of e.g. a seal will then also be largely influenced.

Particles seem to have only a minor influence on the stiffness (only an increase of up to 3 times, or even a decrease), but the tensile strength seems to be decreased as well (Kusy, 1986). The mechanical behaviour of elastomers filled with such particles is not well known.

The finally resulting conductance is not only determined by the percentage and shape of the fillers, but also influenced by the production, which must be processed very carefully.

We may conclude, that the use of elastomers filled with conducting particles or fibres is not without problems. However, the required conductance can be reached, as values of the specific resistance of about 3 and 4 Ωm were reported by Field (1973). Swales et al. (1972) achieved a value of 15 Ωm, which was a compromise between the need of a high conductance and the preservation of the elastic properties. None reported the percentage of filler they needed.

More aspects on metal- and carbon black- filled polymers are provided by Bhattacharja (1986) and Sichel (1982).

Conducting coatings

Conducting coatings were also reviewed by de Jong (1990). Practical application was e.g. by Schouten and Gawlinski (1978[a], 1978[b]). According to de Jong, thin conducting coatings (e.g. a few gold atoms thick) already satisfies the requirement of a maximum resistance of 1 kΩ. Therefore, the influence of the coating on the mechanical behaviour may be neglected. However, the tangential tension of the seal surface, which is often of the order of 1 percent, may cause cracks in the coating and this could yield a serious reduction of the conductance. The probable occurrence of bad lubrication conditions, especially at the start of motion, will cause wear and thus also reduce the conductance.

The practical use of a conductive coating is not recommended unless a coating is available, the wear of which is known to be negligible.

B2.2 The use of two band electrodes on the rigid surface

As discussed in the former section (B2.1), the use of conducting elastomers is not without problems when a rather high accuracy is required, while the use of conducting coatings is not easy. Therefore, a method not requiring conducting elastomers has also been investigated (Visscher, 1989 pp. 29-33). This method uses two band electrodes side by side on the rigid body. Fig. B3 shows a cross section of the configuration.

An isolation layer between the electrodes and the rigid body is of course necessary, when the latter is metallic. The principle is now, that the electrical field, and thus the capacitance, between the electrodes is influenced by the film thickness as long as the electrical permittivities of the elastomer and

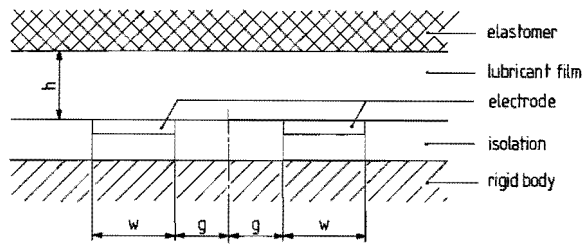


Figure B3
Cross section of a transducer, formed by two band electrodes on the rod.

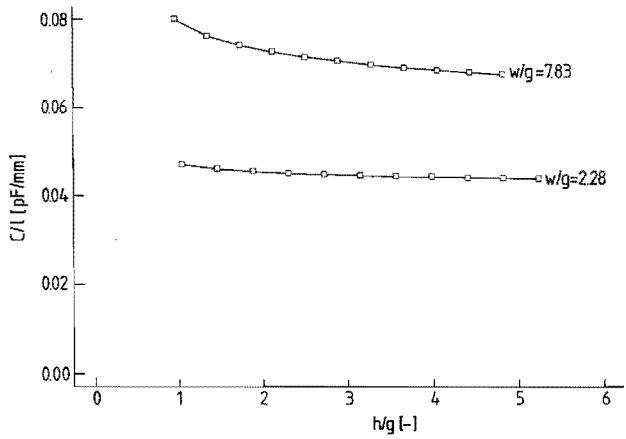
the fluid differs. The relative electrical permittivity ε_r is commonly about 2.3 for oils, 2.1 for unfilled PTFE (but may increase significantly for brass-, glass-, or carbon black filled compounds) and about 4 to 6 for unfilled polyurethane (see e.g. Grzegorczyk and Feineman, 1974; Saure, 1979).

The capacitance has been measured on a system, in which the lubricant was replaced by air, while both the elastomer and isolation layer (with a large thickness) were replaced by glass. The results are presented in fig. B4.

This figure shows that the sensitivity of the method decreases rapidly with increase of h/g. The gap width $2g$ should therefore be large, e.g. 20 μm, to obtain an h/g value smaller than 1 for film thicknesses up to 10 μm.

A reasonable capacitance can be obtained by a relatively long and/or a relatively wide transducer, compared with the dimensions of the contact area. The capacitance would vary from about 8 pF at a thick film to 12 pF at a thin film, when the electrode is 100 mm long, the gap width $2g$ is 20 μm wide and the electrode width w is 80 μm. The total transducer width in this example is 180 μm, which is e.g. more than the contact width of radial lip seals, and the electrode width w should therefore not be larger. A longer electrode is also not realistic, since the electrode length is limited by e.g. the contact length (e.g. the shaft diameter in the case of seals). Both the capacitance and the sensitivity appears thus to be very low in a practical configuration and this kind of transducer is thus not applicable.

Figure B4
The capacitance per unit of electrode length C/l versus the relative film thickness h/g in an air gap between two glass blocks. (the relative electrical permittivity ε_r of the glass is 8.4)
(Source: Visscher, 1989)



B2.3 The applicability of electrical methods for roughness detection

Thus far, we considered the electrical methods for general film thickness measurements, not accounting for the requirement of simultaneous determination of the roughness behaviour during the motion. This will now be discussed.

Dowson and Swales (1969), measuring the lubricant film thickness of carbon black filled elastomeric specimen with a capacitive method, found variations, occurring on a small length scale in a part of the measurements, which were attributed to the surface roughness of the elastomer. The contact width was 25 mm and the electrode in the rigid surface had a diameter of 0.9 mm. The wavelength of the variation was of the order of 1 mm. The amplitude of the variation was roughly about 0.5 μm at nominal film thicknesses of about 2 to 5 μm. Roughness values, like the root mean square of the height distribution and characteristic wavelengths, were not given.

Normal technical surfaces have often a (nearly) Gaussian height distribution and the characteristic wavelengths are rather short, i.e. some micrometers (appendix A2) The electrode dimensions should be of the same order, but this would yield very small capacitance values (e.g. 0.016 pF at an electrode diameter of 10 μm and a film thickness of 0.1 μm), making accurate measurement doubtful.

When the measurements are not performed on an elastomer with a common, more or less Gaussian, roughness texture, but with a more or less regular roughness texture (e.g. a sine wave) instead, the electrode may be larger and the capacitance will thus be generally larger. If the roughness texture is transversal (i.e. no roughness height variation perpendicular to the direction of motion), then the electrode may be also extended in the direction perpendicular to the direction of motion, also increasing the capacitance.

B2.4 Conclusions

The final conclusions are:

- Resistive methods are not suitable for film thickness measurements because of a filtering effect caused by the lubricant film capacitance: small scale film thickness variations are not measured accurately;
- Capacitive methods can be suitable for film thickness measurements in general, but only when a conductive elastomer is used as electrode. The required conductance can be obtained using metallic or carbon black fillers mixed in

the elastomer, but one must consider that the mechanical properties of the elastomer can be dramatically changed.

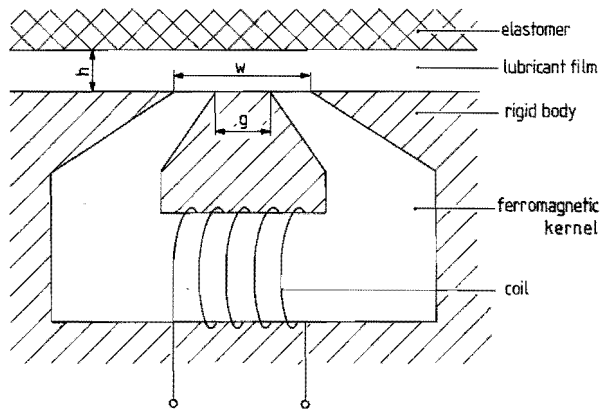
- Electrical methods are not suitable to measure the lubricant film thickness on a micrometer scale. Therefore, the roughness behaviour in the lubricated contact can only be studied using a simple shaped regular roughness texture with a rather long wavelength.

B3 Magnetic induction methods

Using a magnetic induction method, a transducer like the recording head of a tape recorder is mounted in a non-magnetic rigid body (fig. B5). It consists of an electrical coil and a double bent kernel forming the two poles. During the measurement, the magnetic flux lines, which are induced by an electrical current through the coil, cross the gap between the poles through the lubricant film and (partly) through the elastomer. The inductance of the coil will depend on the film thickness, when the magnetic permeability of the elastomer and of the lubricant are different. The film thickness can thus be determined by measuring the electrical impedance of the coil.

A first order approximation of the inductance is given by Poll and Gabelli (1992[a]), who investigated the lubrication of radial lip seals, for the case of a magnetic lubricant

Figure B5
Cross section of an inductive film thickness transducer.



$$L = n^2 \, l \, \mu_0 \, \mu_r \, \frac{h}{g} \qquad \text{(B2)}$$

in which

n	= number of coil windings	[-]
l	= transducer length	[m]
μ_0	= magnetic permeability of vacuum	[$\Omega s m^{-1}$]
μ_r	= relative magnetic permeability of the lubricant	[-]
h	= film thickness	[m]
g	= gap width between the poles	[m]

This approximation suggests a linear relation between the film thickness and the inductance, but in reality, the relation is less than proportional (Poll and Gabelli, 1992[a] fig. 7) and the sensitivity will decrease with increasing h/g ratio. (This was also found for the analogous electrical capacitance method with two electrodes side by side on the rigid surface, see section B2.2 pages 114). This means, that the ratio h/g must not be too high, and the gap width should thus not be too small. The minimum value for the gap width is, however, not known and needs more investigation, if application of the method is considered.

We must also consider, that the effective dimension of the transducer is larger than the gap width g. How much larger is not easily determined. Poll and Gabelli (1992[a]) reported, that the measurement was clearly influenced by the total amount of oil in the vicinity of the contact. The gap width g of the transducer was about 5 to 10 μm, while the contact width was about 1 mm and the total transducer width w was about 3 mm.

A practical complication using a magnetic induction method is, that only a few elements are ferromagnetic, i.e. they have a magnetic permeability significant larger than one at room temperature. The permeability of other elements and of materials not containing ferromagnetic elements, like both oil and elastomers, is very close to one (deviating less than 0.1 percent). Application of an inductive method needs therefore the use of a lubricant or an elastomer filled with ferromagnetic particles, which will generally lead to a change in the material properties.

Use of an elastomer filled with magnetic particles

Schrader (1978 p. 54) first tried an inductive method with an iron filled seal. The method appeared to be very sensitive, but also to be disturbed by the

pressure dependence of the magnetic permeability of the seal at pressures up to 20 MPa (Schrader, 1978 p.48). Another problem was the high wear rate of the seal. Therefore Schrader finally applied another method (the mechanical method described in section B1, page 107) instead of a magnetic method.

Use of a magnetic lubricant

Poll and Gabelli (1992[a]) applied the method, using a normal radial lip seal and a magnetic fluid (an oil with suspended magnetic particles) with a relative magnetic permeability of about 2. The sensitivity appeared to be good, but there is no indication of the accuracy.

One influence factor mentioned is the temperature rise in the contact, which may be high for radial lip seals. An estimation of the resulting error is, however, not given.

Another influence factor is the probable pressure dependence of the permeability of the oil, just like the permeability of the seal of Schrader was pressure dependent. Poll and Gabelli did not mention it and this influence might be negligible small indeed, because the contact pressures of radial lip seals are rather low (of the order of 1 MPa, see e.g. Stakenborg, 1988 section 3.6).

A matter of discussion is, whether the magnetic particles, with 10 nm average size and 80 nm maximum size, disturb the fluid film, which is not unlikely when thin films locally occur, e.g. due to surface roughness effects. According to Poll and Gabelli, the particles will not disturb the formation of a lubricant film nor cause wear. However, the measured minimum film thickness appeared to be of the same order as the combined roughness of the seal and the shaft at low velocities and the particles may therefore influence simultaneous friction measurements.

Conclusion

The final conclusion is, that the method needs more investigation on among others the temperature and pressure influence for general application. The effective transducer dimensions are also a problem to be coped with.

Considering the requirement to determine the roughness of the elastomeric surface, it may be concluded that the transducer should be much smaller. The

sensitivity will then probably decrease and whether the method is still accurate enough is not known at the moment.

Finally, the use of an iron filled elastomer seems not to be preferable. Otherwise, the use of a magnetic fluid can be undesirable when thin fluid films are locally expected.

B4 Optical methods

Optical methods are frequently applied to measure the lubricant film thickness (see e.g. section 1.2.2, page 13, and this section below). A distinction can be made between different principles of which interferometry is most widely used. The word "optical" commonly refers to visible light, but all kinds of electromagnetic radiation can be applied to the same principle.

Optical methods generally requires one of the mating bodies to be (partly) transparent for the applied range of frequencies and it will be obvious to make the rigid body transparent, mainly because the optical properties of transparent elastomers are generally poor. Otherwise, a lot of transparent materials with a high stiffness and high optical quality are available and can be used for the rigid body.

Most optical methods need reflection. Interferometry (section B4.1) e.g. is based on reflection on both the elastomer to lubricant interface and the interface between the lubricant and the (transparent) rigid body, while the focusing method (section B4.4) uses the reflection on the elastomer to lubricant interface. Therefore, application of these methods introduces requirements on the reflectance of one or both interfaces. Considering nonmetallic materials, the reflectance is determined by the indices of refraction and the angle of incidence. When e.g. the indices of refraction of two materials are equal, no reflection will occur on their interface. In practice, the indices of refraction of glass, oil and elastomers are not very different. They all range from about 1.3 to 1.7. Consequently, the combination of the solids and the liquid should be chosen carefully.

The optical methods can further be distinguished in methods, producing a (two-dimensional) fringe pattern (i.e. a pattern of light and dark lines), containing information about the (three-dimensional) height structure of (a part of) the surface (e.g. interferometry and moiré topography), and methods scanning a (two-

dimensional) height profile along a line.

We must emphasize here, that the methods producing a fringe pattern do not require the instrumentation to be mounted in (or on) the rigid body as was specified in requirement no. 1 (page 20). Instead, the height contour of the lubricant film can be observed with a suitable optical instrument (e.g. a microscope), when the elastomer is not moving, but the rigid body moves. These methods have then two advantages: The whole contact can be observed (instead of only a line) and the film thickness is not determined by a signal measured in time but by a picture of the contact area. Therefore, the requirement on the fast response (requirement no. 6, page 20) can be dropped.

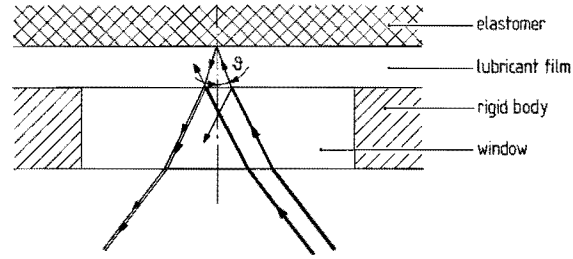
However, the resolution is generally low and the picture of the three-dimensional contour often provides only the height differences between neighbouring fringes, as will be shown below in the discussion of the methods producing a fringe pattern (interferometry and moiré). Neither the absolute value of the surface height at the position of a fringe, nor whether a fringe represents a higher or a lower position than a neighbouring fringe, can be derived from the fringe pattern itself. Special measures are then necessary to derive the absolute film thickness. Such measures can consist of starting with an (almost) unlubricated situation (in which no (or only a very thin) film exist and no fringe is visible) and counting the number of newly generated fringes, when the film thickness increases under lubrication conditions. Other measures are possible using more optics and/or a lot of calculations.

Having considered some general items of most optical methods now, these methods will be discussed individually in the following. We will first consider the practical use of the method for film thickness measurements in the contact of an elastomeric and a rigid surface in general, followed by a further discussion on the suitability for the measurement of the surface roughness of the elastomeric surface during the lubrication.

B4.1 Interferometry

Interferometry uses the reflection of a coherent collimated beam on both contacting surfaces with the lubricant film in between. In fig. B6 two rays of the incident beam are drawn and these two rays are in phase, since the incident beam is coherent. In the reflected beam one ray, reflected on the elastomer to lubricant interface, interferes with another ray, reflected on the lubricant to

Figure B6
Film thickness measurement by means of interferometry.



window interface. These two interfering rays have propagated over a different distance and are therefore generally out of phase, the phase shift being dependent on the difference in distance (and thus on the lubricant film thickness) and on the index of refraction of the lubricant.

The amplitudes of the interfering rays are superposed, resulting in a bright spot where the amplitude, and thus the irradiance[1], is maximum (constructive interference), This occurs where the phase shift is zero (or i times the wavelength, i being 1, 2, 3, ...). A dark spot is present where the amplitude, and the irradiance, is minimum (destructive interference), which occurs where the phase shift equals half the wavelength (or $(i + {}^1/_2)$ times the wavelength). More incident rays will altogether form an interferogram, consisting of dark lines (fringes) representing contours of equal film thickness which can be seen directly with the eye or, magnified, using a microscope.

A practical problem, already discussed on page 121, is that the fringes only represent height differences when using monochromatic light, due to the periodicity in the phase shift, and not absolute values of the film thickness. There are, however, some possibilities to overcome this problem as will be shown below.

Further, the vertical resolution of the method, the applicability to elastomers and the applicability to rough surfaces will be discussed.

[1] The term *intensity* was generally used in the past, but it is nowadays often replaced by *irradiance* in optics.

B4.1.1 Derivation of the absolute film thickness

Monochromatic light is used in most applications of interferometry. However, a problem is that only film thickness differences can be determined and not the absolute value of the film thickness. To overcome this problem, the contact is often observed continuously and the number of newly generated and the number of disappearing fringes is counted from the start (when the film is not present or very thin) until the lubricant film becomes stationary. Then, the order of the fringes is known and the absolute value of the film thickness can be derived.

Another solution is the use of white light, consisting of a range of wavelength. Every wavelength has its own fringe pattern, with a unique height distance between the fringes, and the overlapping of these fringes, with different wavelength, yields a coloured interference pattern from which the absolute film thickness can be determined directly from the colour of a fringe after calibration. However, the application of white light interferometry appears to be limited to films thinner than about 1 μm, according to McClune (1974 p. 45) and Kalsi (1975 p. 28), since so many different colours are present in one point at thicker films, that the picture is nearly white and the contrast of the pattern is thus very low.

A probable better solution is the use of bichromatic light, as was performed by Krauter (1982). Now two monochromatic interference patterns overlap and the combined periodicity is larger than of the two single patterns (of the order of 1 μm or larger, instead of 0.1 μm) and the absolute film thickness is derived directly, since the film thickness is commonly of the order of 1 μm.

B4.1.2 The vertical resolution

The height difference Δh between two neighbouring (dark) fringes is (see e.g. Hecht, 1987)

$$\Delta h = \frac{\lambda}{2n \cos\vartheta} \tag{B3}$$

in which

λ	= wavelength in air	[m]
n	= index of refraction of the lubricant	[-]
ϑ	= angle (see fig. B6)	[rad]

At normal incidence (ϑ = 0) and with an index of refraction of 1.5, the height difference Δh will be e.g. 0.14 μm for λ = 0.42 μm (blue) and 0.26 μm for λ =

0.78 μm (red).

The vertical resolution can be up to 10 times the height difference Δh using interpolation techniques. Then a vertical resolution of roughly 0.02 μm is possible, which is not fully sufficient to realize the required accuracy (the maximum uncertainty in the measured film thickness should be about 0.01 μm for films in the range of 0.1 to 1 μm and about 1 percent for film thicknesses in the range of 1 to 10 μm, as stated in chapter 2, page 19).

B4.1.3 The applicability to elastomers

Interferometry has been widely applied in studying lubricated metal to metal (or in fact: metal to glass) contacts. The technique is also favoured in the investigation of elastomer to metal (glass) contacts. The main problem to be solved was the creation of a sufficient reflecting elastomeric surface.

Blok and Koens (1966) considered, that the reflectance of the oil-elastomer interface would be too low to obtain a sufficiently clear interferogram. They solved this problem with a thin and flexible aluminized plastic foil, bonded on the rubber surface. Krauter (1982) applied a thin elastic lacquer coating on the elastomer.

Roberts (1968) and Roberts and Tabor (1968, 1971) successfully produced optically smooth elastomeric surfaces and obtained clear interference patterns without the need to attach a highly reflecting material to the surface. Subsequent work was presented by McClune (1974), McClune and Briscoe (1977) and McClune and Tabor (1978).

Later investigators adopted the technique of Roberts to produce optically smooth elastomers and also used interferometry, e.g. Field (1973), Field and Nau (1973[a], 1976), Austin, Flitney and Nau (1977), Flitney (1982) and Kalsi (1975, 1981).

B4.1.4 The applicability to rough surfaces

The work of Roberts (1968) and Roberts and Tabor (1968, 1971) (see also section B4.1.3 above) proves, that the low reflectance on the oil-elastomer interface is not the real problem for interferometry, but the roughness of the surface. The surface roughness has two effects.

One effect is that each asperity is surrounded by a number of interference fringes, which are generally very close to each other because of the rather steep slopes in the surface. The fringes can then be too close to enable proper distinction (see Jackson and Cameron, 1976) and the maximum tolerable surface slope is thus limited by the lateral resolution (i.e. the smallest distance between two fringes which is resolved) of the interference microscope.

The other effect is, that the roughness cause light scatter, which may disturb the interference pattern seriously (see Tønder and Jakobsen, 1992). This effect is also determined by the surface slopes (varying with position), which cause the rays to be reflected in different directions.

Nevertheless, interferometry is used for surface roughness measurement (see e.g. Wyant et al., 1986), but the roughness height and the surface slopes appear to be small in the presented measurements: The peak to valley roughness height is then e.g. smaller than 0.2 μm and the slopes are smaller than e.g. 0.03 rad. Therefore these measurements do not indicate whether interferometry can be useful for our application, in which the roughness height is of the order of 1 μm and the slope of the order of 0.1 rad.

Now both the lateral resolution and the influence of scatter will be discussed in more detail to estimate the maximum allowable surface slope for application of interferometry.

The maximum surface slope

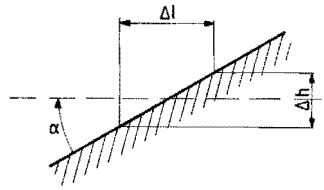
As mentioned above, the slopes in a surface with a normal roughness texture can be so large that two neighbouring fringes are too close to be distinguished by the microscope. They are then seen as one fringe instead of two (or more) in the interferogram, leading to misinterpretation. Therefore, the maximum slope which can be allowed is determined by the lateral resolution (apart from the scattering effect, which will be discussed below).

Given a slope with an angle of inclination α (fig. B7), we can define the lateral distance Δl between two neighbouring fringes (with a height difference Δh, as expressed in eq. (B3) on page 123)[2]

$$\Delta l = \frac{\Delta h}{\tan \alpha} \tag{B4}$$

[2] We are only interested in a rough estimation here and will therefore not account for the fact, that the direction of propagation of the reflected ray is determined by the surface slope.

Figure B7
Definition of the angle of inclination of the surface.



The lateral resolution of the microscope (and of all optical image forming instruments) is limited by diffraction. Using Rayleigh's criterion we find for the resolution

$$(\Delta l)_{\min} = 1.22 \frac{f\lambda}{D} \tag{B5}$$

in which f is the focal length and D the aperture of the microscope objective lens (see e.g. Hecht, 1987 p. 422)[3]. For normal incidence, we can now estimate the maximum allowable angle

$$\tan \alpha_{max} = \frac{\Delta h}{(\Delta l)_{\min}} = \frac{D}{2.44\ nf} \tag{B6}$$

When a lens with a numerical aperture[4] of 0.5 (i.e. $D/f = 1.15$) is used and the index of refraction n is 1.5, we find for the maximum slope

$$\tan \alpha_{max} \approx 0.3$$

and α_{max} is thus about 0.3 rad (17.5°), which compares to the slopes found in roughness measurements on elastomeric seals (see appendix A2).

The light scatter

The slopes in the surface have a different angle of inclination and the rays reflected at the elastomer to oil interface will therefore propagate in different directions, although the incident beam is collimated. The consequence is,

[3] The "Rayleigh's criterion" is derived for incoherent radiation, but the difference with the formula for coherent radiation is not large (see e.g. Born and Wolf, 1970, pp. 418-424).

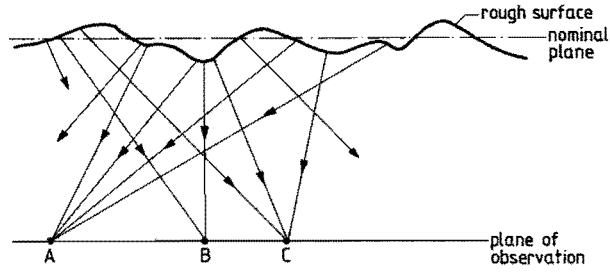
[4] See for definition appendix F1

that the interference pattern can be seriously disturbed, since interference, even between rays reflected on different parts of the elastomer (fig. B8) now occurs more or less randomly and a speckled interferogram is formed instead of a nice fringe pattern, even when the slopes are small.

This "speckle effect" always occurs when a free rough surface (i.e. without oil film and window) is illuminated by coherent light (see e.g. Hecht, 1987 p. 592) and is sometimes used too determine the standard deviation of the roughness height distribution (see e.g. Sherrington and Smith, 1988), but is only useful for roughness heights smaller than the wavelength λ ($R_a < \lambda/5$ according to Sherrington and Smith). A roughness profile is also not obtained.

It is because of this speckle effect, that film thickness measurement by means of interferometry requires a smooth surface or a roughness texture with very small slopes.

Figure B8



The speckle effect: Rays reflected on different parts of the rough surface converge to one point (e.g. A) in the plane of observation, where they interfere.
(The incident beam, which is not drawn, is perpendicular to the nominal plane).

Conclusions

Interferometry is suitable to measure the lubricant film thickness on elastomers. However, application to surfaces with a normal (more or less random) roughness is not possible because of the "speckle effect": The interference pattern is totally disturbed by interference of light, scattered in several directions due to the variation in surface slopes. Therefore, application of interferometry requires the use of a smooth surface, or a regular "roughness" texture with small slopes.

B4.2 Moiré methods

The basic idea of moiré methods is, that two overlapping transmission gratings, with equal line spacing and a slightly different grating orientation, show a pattern of light and dark lines, as shown in fig. B9. The distance d of these pattern lines depends on the grating pitch p and the angle ϑ between the grating orientations. The principles of moiré are presented by e.g. Kafri and Glatt (1990).

Film thickness measurements by means of moiré topography (to be more specific: shadow moiré) was performed on a rubber specimen by Hori et al. (1981). The rubber was in contact with a glass plate, having a grating on its contacting surface. The essence of the method is, that the contact is illuminated under an angle and a shadow of the grating is formed on the rubber contact face. The lateral position of the shadow lines depends on the angle of illumination and on the local film thickness (fig. B10). The real grating lines and their shadow lines form together a moiré pattern when observed from above, showing contour lines of equal height, i.e. equal film thickness (see e.g. Meadows et al., 1970; Takasaki, 1973; Kafri and Glatt, 1990 p. 61 ff.).

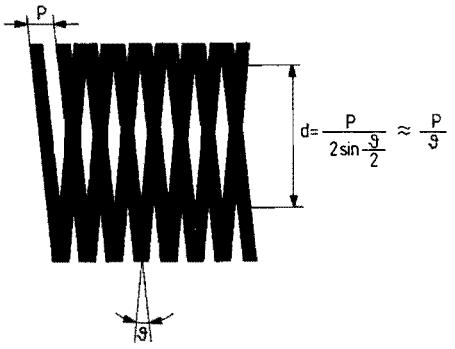


Figure B9
Moiré pattern of two overlapping transmission gratings.

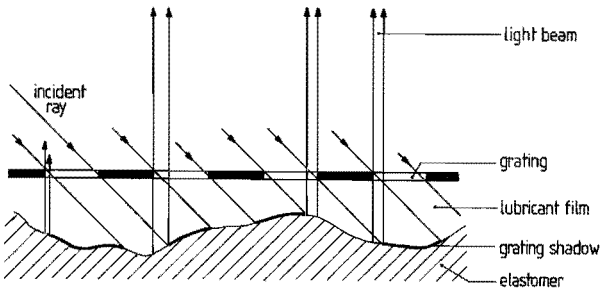


Figure B10
Film thickness measurement by means of shadow moiré.

Moiré methods are in principle incremental, just like monochromatic interferometry. The contour lines provide only information on height differences and which line is the higher and which one the lower can at the best only be guessed from previous knowledge of the shape. The absolute film thickness of a contour line is then of course also unknown. To overcome this problem, Takasaki (1973) suggests among others to make stereo photographs. The additional information can then be obtained from the relative differences of the position of contour lines on the two pictures.

The lateral and vertical resolution

In the configuration of Hori et al. (1981), the grating pitch, which determines the lateral resolution, was 63.5 μm. They also reported, that the height difference between two adjacent contour lines (i.e. the vertical resolution) was almost equal to this grating pitch and 10 contour lines could be observed. The films investigated were thus relatively thick and the spatial resolution, which is of the order of the grating pitch, was rather low.

Application of the method for thinner films and a better spatial resolution (both of the order of 1 μm) requires a grating pitch of the order of 1 μm. However, the grating pitch will then be of the same order as the wavelength of the light, causing serious problems because of diffraction. Moiré topography will therefore only be applicable, when the range of film thicknesses is increased by at least 10 times.

The applicability to rough surfaces

As we have just concluded, the moiré method has a low lateral and vertical resolution, which means that the surface roughness deformation can only be studied, when both the height and the wavelengths of the roughness texture are at least 10 times larger than for normal roughness textures, making the roughness height and wavelength of the order of 10 μm instead of 1 μm. In practice, an artificially manufactured "roughness" texture must be used.

B4.3 Ellipsometry

Film thickness measurement by means of ellipsometry is based on the reflection of a parallel, linearly polarized beam on the film. The state of polarization of the reflected beam, which depends among others on the film thickness, is measured. The relation between the state of polarization and the film thickness is, however, complicated and the evaluation of the measured data therefore not simple. The principles are described by e.g. Neal and Fane (1973), Azzam and Bashara (1977), Hanekamp (1983) and Riedling (1988).

Ellipsometric measurements of the lubricant film thickness in a (full film) lubricated contact have not been published, to the author's knowledge. An example of oil film thickness measurement on a free metallic surface is given by Meyer and Loyen (1975), while Çavdar and Ludema (1991) measured the thickness of boundary films in steel contacts during motion.

Meyer and Loyen (1975), using a He-Ne laser with a wavelength of 632.8 nm, reported the uncertainty of their measurements to be about 10 percent and considered the maximum film thickness, which can be detected, to be 500 nm. The reason for this limitation is the simultaneous occurrence of interference, which leads to alternating light and dark fringes. At the place of the dark fringes is no light and the state of polarization can obviously not be measured there. This would mean that the method is not suitable for the proposed application of this study, since film thicknesses of the order of 0.1 to 1 μm are expected (see point 6 on page 19).

However, a solution can be the use of infrared radiation. Then, the first dark fringe is present at a thicker film, because of the longer wavelength, and the method can be applied to thicker films (e.g. of the order of 1 μm).

Measurement on rough surfaces

It is not apparent from literature, whether and to what extent ellipsometry is applicable to rough surfaces. The measurements of Meyer and Loyen (1975) were performed on polished surfaces with R_a values of 0.05 and 0.1 μm, but information on other important matters, like the slopes, is not provided. The influence of the roughness height and slopes has been studied (see Azzam and Bashara, 1977 pp. 361-363, for a review) and it appeared that it can be very large. Some

simplified models to account for the surface roughness have been proposed.

Most of the literature considers the standard configuration of ellipsometry using a parallel laser beam, having a diameter of about 500 μm or more. It appears to be also possible to apply a very small spot (the diameter being of the order of the wavelength) by focusing the beam on the surface of investigation and in doing this, the required spatial resolution can be obtained, according to Svitashev et al. (1971, 1973).

Conclusion

We can conclude now, that ellipsometry needs more investigation to decide, whether it is applicable to measure the film thickness and the roughness deformation in the lubricated contact.

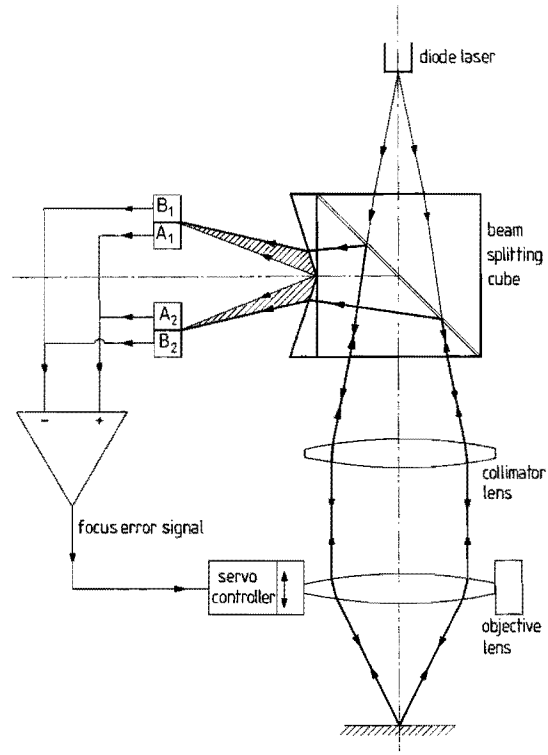
B4.4 Focus error detection

In the last decade, optical methods have been developed for non-contacting surface roughness measurements. Most of them are based on focus error detection, meaning that a lens is focused onto the surface by an active control system (fig. B11). The heart of the system is an opto-electronic device, detecting whether the focus spot of the lens is above, on, or (virtually) below the surface: All four photodiodes receive the same amount of light when the surface is in focus (as drawn); The outer diodes (B_1 and B_2) receive more or less light than the inner diodes (A_1 and A_2), when the surface is above or under the focus respectively. The output of the device is a so-called focus error signal, being positive, zero or negative respectively. Measurement of this focus error signal thus yields the surface height relative to the focal point. The measurement range is then e.g. some micrometer (see e.g. Struik and Chang, 1987).

Increase of this measurement range is possible using a servo controller (see fig. B11) which keeps the focus error signal on zero by repositioning the lens. In this way, the lens remains being focused on the surface, the roughness of which is determined from continuous measurement of the lens position, while the surface moves in a vertical direction (perpendicular to the optical axis).

Many opto-electronic systems for focus error detection are available and a lot of them have been applied (see e.g. Bouwhuis and Braat, 1978; Mignot and Gorecki, 1983; Kohno et al., 1985; Mitsui et al., 1985; Mitsui, 1986; Bouwhuis

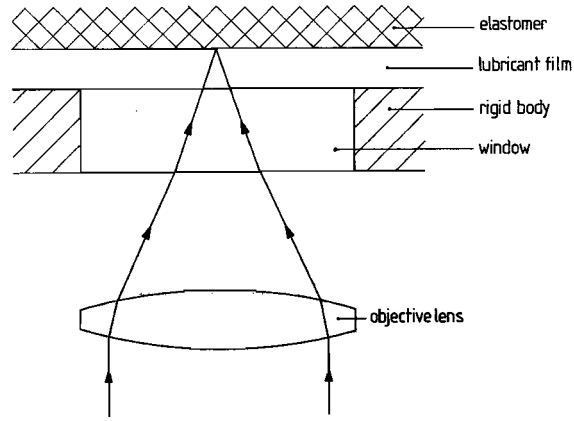
Figure B11
A focus error surface roughness sensor. (Struik and Chang, 1987)



et al., 1987 pp. 75-80; Struik and Chang, 1987; Sherrington and Smith, 1988; Kagami et al., 1989). Though the different methods all have to some extent their own characteristics, regarding e.g. linearity, sensitivity, accuracy, and the possibilities to influence these, we will not discuss them individually. Our concern is now to evaluate the suitability of the method for our purpose in general.

To the authors' knowledge, focus error detection has not been applied to film thickness measurements yet. The method could be helpful for us, scanning the height profile of the elastomeric surface through a glass window in the rigid surface (fig. B12), as we are not only interested in the (nominal) film thickness, but in the real roughness structure in the lubricated contact as well. The greatest advantage is, that the dimension of the measuring spot (i.e. the focus spot) and the vertical resolution are of the desired order (about 1 μm and 1 nm respectively).

Figure B12
Scanning of the height of the (rough) elastomeric surface through a glass window.



Possible problems are:
- The poor reflection on the elastomer to lubricant interface;
- the influence of the reflection on the lubricant to glass interface;
- the dynamic behaviour of the system;
- the influence of surface slopes on the measurement.

These will now be discussed briefly.

The reflection on the elastomer to lubricant interface

The reflection on the elastomer to lubricant interface will be very low, e.g. 1 percent when the indices of reflection of the elastomer and the lubricant are 1.7 and 1.4 respectively (appendix D), and only 0.1 percent when the indices of refraction are 1.5 and 1.4. Whether this low reflectance will actually be a problem, depends on the laser power, the sensitivity of the photodiodes and on eventual background scatter. The use of a thin reflecting coating on the elastomeric surface can be considered, if the reflectance is not large enough.

The reflection on the lubricant to glass interface

The second possible problem is the influence of reflection on the window to lubricant interface. The effect of this reflection is probably, that the objective lens will not focus on the elastomeric surface, but somewhere between the elastomeric and the glass surface (see appendix G). The measurement will thus yield a film height, which is lower than the real film thickness. How much

lower will depend on the ratio of the reflectances and must be derived from e.g. calculation and/or calibration.

The dynamic behaviour of the system

Another problem may arise from the dynamic behaviour of the system. Continuous refocusing of the lens will not be possible, because of the relatively slow response of the mechanical system: The response time is much higher than the required 1 μs (requirement number 6, page 20). Therefore, the lens must be fixed and the focus error signal itself must be recorded as a measure for the local film thickness. This will restrict the measurement range for two reasons. One is, that the focus error signal is only well defined in a small range around the in-focus position (e.g. some micrometer in the device of Struik and Chang, 1987). The other reason is, that the real dimension of the measurement spot increases with increasing distance between the focus spot and the scanned surface, because of the conical shape of the beam. Both facts limits the maximum film thickness which can be detected properly. However, a range of some micrometers is possible (Mitsui et al., 1985) and this figure can be influenced by variation of design parameters (e.g. the focal length of a lens) (Mitsui, 1986).

The influence of surface slopes

When the surface is not perpendicular to the optical axis, the cone of the reflected rays will not be symmetrical around the optical axis, but skew. Consequently, not all light will return into the device and the amount of light, received by the photodiodes, will be reduced. When the surface slope is larger than the half top angle of the incident light cone, no light will reenter through the objective lens. At a smaller angle (e.g. about 10° or lower for the device of Struik and Chang, 1987), the effect has no significant influence on the lens focusing.

If, however, the lens is fixed, as required in our application by the dynamics, the focus error signal is significantly influenced. The maximum surface slope, which has a negligible influence, is 1° for the device used by Struik and Chang (1987) and 5° for the device of Kohno et al. (1988), who adopted a different principle for focus error detection. The slope influence appears thus to depend on the kind of transducer. The influence of the surface slope may thus be neglected, when a suitable focus error device is chosen.

If the slope influence is not negligible, simultaneous measurement of the so-called "radial error signal" can be considered to eliminate the slope influence. In the compact disc sensor, the radial error signal is used to position the sensor in the radial direction in order to keep in track. If the focus is not in the middle of the track, and thus not in the centre of a pit when it passes, the reflecting light beam is not symmetrical around the optical axis, but skew (just as when the surface is not perpendicular to the optical axis). As a result, one pair of photodiodes (A_1B_1) will receive a different amount of light than the other pair (A_2B_2) (Bouwhuis et al., 1987). The radial error signal measures this difference and a servo controller is used to position the sensor in the radial position, keeping the radial error signal zero and thus assuring that the focus remains in track.

Also this radial error signal can in principle be used in our application to measure the local surface slope, enabling the correction for the slope influence when the film thickness is derived from the focus error signal (see appendix E for more details).

Conclusion

The focus error detection method has the advantage, that the measurement spot dimension is principally of the desired order to detect the roughness on the elastomeric surface during motion. There are still some questions on the proper working. These are the low reflectance on the elastomer to lubricant interface, the influence of the reflectance on the lubricant to glass interface and the influence of the surface slopes. However, these possible problems are probably solvable.

B4.5 Absorption methods

Absorption methods are based on the fact that one or more wavelengths from an incident beam is absorbed by the lubricant. When the beam propagates through the film, the irradiance of the transmitted beam will be lower when the film is thicker and the total absorptance higher. The film thickness can therefore be determined from the ratio of the transmitted beam irradiance and the incidence beam irradiance.

Measurement of the lubricant film thickness using absorptance techniques are not known from literature. Cann and Spikes (1984) applied a similar technique (Reflection-Absorption Infrared Spectroscopy) to measure the oil film thickness on a free, smooth metallic surface. They reported, that their technique was suitable for films of 0.05 to 1 μm thick. The reason was, that the sensitivity of the method decreases with increasing film thickness. Therefore, the measurement of thicker films (e.g. up to 10 μm) seems to be possible when a lubricant and wavelength combination with a lower absorptance is chosen, i.e. the same absorptance is achieved at a thicker film.

The beam which is transmitted through the film must of course be captured by a photodetector. It can be transmitted through a transparent elastomer to reach the detector, or it can be reflected on the elastomeric surface. The former option is not preferable because of the commonly relative bad optical properties of elastomers for light transmission. Therefore, reflection appears to be necessary to capture the beam on a photodetector. However, possible difficulties may arise from the very low reflectance on the lubricant to elastomer interface.

Surface roughness detection is perhaps possible with a focused beam, yielding a small spot, but it is not known, whether the required accuracy can then be obtained or not. More research is therefore necessary before eventual application.

B4.6 Fluorescence

A fluorescing material has the property, that the energy of absorbed radiation with a certain frequency is subsequently emitted in the form of radiation with a different frequency. The absorption and fluorescence spectra are typical for the material: Some frequencies of the incident beam are strongly absorbed with an also strong emission of fluorescence radiation, while other frequencies of the incident beam have hardly or no fluorescence effect.

When the used lubricant fluoresces, either as a natural property or originating from a fluorescent dye solved in it, the film thickness can be determined by measuring the irradiation ratio of the incident and the fluorescent light beam.

The great advantage of fluorescence for film thickness measurements in elastomer to metal (or glass) contacts is, that no physical property of the elastomer (or elastomeric surface), like e.g. reflectance, is used. Therefore,

all kinds of filled or unfilled elastomeric materials can be used, as long as eventual fluorescence of the elastomer does not yield too much background signal.

However, fluorescence is hardly used for such measurements up till now. Kassfeldt (1987) applied the method to reciprocating piston seals, but little information is provided on the quality of the results. She only reported the measurement spot to be smaller than 1 mm in diameter, the seal contact width being about 5 to 7 mm. An indication of the accuracy was not given and it seems that the calibration was performed for a film thickness range one order of magnitude larger than the real film thicknesses measured.

Recently, the method was also adopted by Poll et al. (1992[b]) to measure the lubricant film thickness of radial lip seals. They discussed the performance of the system and concluded that especially the concentration of the fluorescent dye, solved in the lubricant, is an important factor considering the accuracy and the sensitivity. Remarkable is that the lubricant's viscosity is significantly decreased by the solved dye according to Poll et al. At the "optimum dye concentration"[5] the kinematic viscosity would be decreased from 68 mm^2/s (for the lubricant without dye) to 57 mm^2/s (at a temperature of 40 °C). Possible influence on other lubricant properties (e.g. concerning the boundary lubrication behaviour or non-Newtonian behaviour) was not discussed.

A number of questions still remains, mainly because the final achieved accuracy and spatial resolution were not given. Also it is not clear whether the fluorescence method is accurate enough when the spot size is of the order of 1 µm. This information is also not provided by other literature on similar fluorescent film thickness measurements, e.g. applied to oil films on a free metallic surface (e.g. Smart and Ford, 1974; Ford and Foord 1978; Köhnlechner 1980; and Schmutz, 1984), or applied to the lubrication of two metallic counterfaces (e.g. Ting, 1980[a] and 1980[b]; and Hoult and Takiguchi, 1991).

The conclusion is now, that the fluorescence method seems to be a very attractive method, since no requirements on the properties of the elastomer have to be made. Otherwise, the accuracy is still questionable, as well as the spatial resolution which can be reached. It is especially unclear whether the accuracy will be sufficient at a small spot of about 1 µm in diameter.

[5] It is not reported what this concentration actually was, but it was probably of the order of 0.1 percent.

More investigation is needed to find out whether fluorescence can really be used with such a small spot.

B5 Ultrasonic methods

Ultrasonic methods are qualitatively comparable to optical, as in both cases the film thickness is measured by use of propagating waves. A difference is the media in which the waves propagates: Light propagates in electromagnetic fields, while (ultra)sound propagates in pressure fields.

Application of ultrasound for film thickness measurement is not known by the author. Ultrasonic surface roughness measurements were reported by Blessing and Eitzen (1988, 1989), who used a pulse-echo method (see e.g. Szilard, 1982 pp. 41ff.) to measure the travel time of a wave from the transducer to the surface and back again. The transducer to surface distance can then be derived when the sound speed in the used medium (air or a liquid) is known and thus the (nominal) surface profile can be derived. The average of a typical roughness parameter (like e.g. the Root Mean Square) was derived from the amplitude of the reflected ultrasound beam.

In most experiments, the measurement spot was several millimeters in diameter. Some measurements were performed using a focused beam, which yielded a spot of some 100 μm in diameter. A spot of about 1 μm in diameter is also possible, as reported by e.g. Weglein and Wilson (1977), which enables the scanning of e.g. a roughness profile, but the vertical resolution and the accuracy of such measurements are not given. A literature review by Breeuwer (1991) on position measurements indicates, that a vertical resolution of 0.1 μm has been reached using a non-focused beam of 6 mm in diameter (Fox et al., 1984), while a resolution of 0.6 μm has been reached for a focused beam with a spot of 400 μm in diameter (Fox et al., 1985).

In general, ultrasonic methods seems not to be favoured in measurement of roughness profiles nor in film thickness measurement. It is not clear whether they can be applied. Especially the accuracy which can be achieved is uncertain.

B6 Conclusions and choice of the method

A number of methods, which can in principle be used for film thickness measurements, have been discussed. Not all methods, which could perhaps be applied, are mentioned, but the discussed methods can be considered to be the most important, since they have been applied for either film thickness measurement or for similar matters. Now we must make a choice of a method which will be further developed and applied.

First, the suitability of the different methods will be summarized and compared using criteria derived from the requirements mentioned in chapter 2 (page 19-21). These criteria will now be mentioned in sequence of importance (i.e. a crucial criterion, which must be fulfilled, is mentioned first and a recommendation, which is a criterion that can be dropped if necessary, is mentioned later).

The criteria

The most important criterion is derived from the fact that the lubricant film thickness must be measured. This measurement only make sense if the transducer does not influence the measured quantity. Therefore the first and most crucial criterion is:

1. *The lubricant film thickness must be measured without disturbance of this film at the spot of measurement.*

Secondly, the method must be able to detect the film thickness accurately. This means for the expected range in film thicknesses:

2. *Film thicknesses in the range of 0.1 to 10* μm *must be measured accurately.*

Thirdly, the eventual deformed roughness texture in the lubricated contact must be detected. It is very recommended to use an elastomeric surface with a roughness texture comparable to the roughness of seals, having a characteristic wavelength and a roughness height of the order of 1 μm. Use of an elastomer with a "model roughness" having a significantly longer wavelength can be considered, if necessary, but is not recommended, because there is evidence that the roughness deformation in a lubricated contact is less pronounced when the wavelength is smaller (Kweh et al., 1992; see also section 1.2.1.1 page 10).

The requirement to measure the local roughness height in the lubricated contact has a consequence for the spatial resolution, and thus for the dimension of the measurement spot, as well as for the maximum tolerable response time of the system in relation with the sliding velocity in the contact area (this latter point does not apply to some optical methods, as discussed in section B4 page 121). This yields the following two criteria[6]:

3. *The spot size must be of the order of 1* μm;
4. *The maximum response time must be limited to the order of 1* μs.

The final criterion concerns the physical properties of the elastomer and the lubricant:

5. *Change in the mechanical behaviour of the elastomer and/or the lubricant is not preferred.*

If, however, such a change can not be avoided, because of an unavoidable change in a physical property to enable the measurement, the consequence for the lubrication problem must be quantified.

The pressure and temperature influence is not mentioned as a criterion. The reason is, that there is very little information on this subject in literature and can therefore not be used to compare the different methods. However, it must be investigated for the method which will be chosen.

---

[6] Of these two criteria is n. 3 more crucial because it is determined by the characteristic wavelength in the roughness texture only, while criterion 4 is also determined by the velocity. In a general full film lubrication problem (see section 1.1.1 page 3) the velocity is not an explicit parameter, but the problem is characterized by some dimensionless parameters (see e.g. Hamrock and Dowson, 1978; Venner, 1991; see also appendix 10), involving e.g. the product of the velocity and the lubricant's viscosity. Consequently the velocity can be reduced without affecting the lubrication problem (e.g. the ratio of the film thickness to roughness height) by using a lubricant with a higher viscosity. Then a slower response of the transducer can be tolerated because of the lower velocity.

Comparison of the methods

The methods will now be compared, using a scheme in which is indicated whether and to what extend a criterion is or can be fulfilled. This scheme is given below in table B1.

Table B1 Comparison of the methods discussed in the former sections.
The numbers 1 to 5 refer to the criteria mentioned above:
1: No disturbance of the lubricant film;
2: Accuracy for a film thickness in the range of 0.1 to 10 μm;
3. Spot size of the order of 1 μm;
4. Response time of 1 μs;
5. No change in mechanical properties of elastomer and lubricant.
("Capacitive 1" is the capacitive method in which the elastomeric counterface is used as electrode; "Capacitive 2" is the method in which two band electrodes are applied onto the rigid body).

Method	(page n.)	**1**	**2**	**3**	**4**	**5**
Mechanical	(107)	--	-	+	--	+
Electrical:	(108)					
Resistive	(110)	+	+	?	--	□
Capacitive 1	(111)	+	+	--	+	□
Capacitive 2	(114)	+	-	--	+	+
Magnetic Induction	(117)	+	?	?	+	□
Optical:	(120)					
Interferometry	(121)	+	□	-	/	+
Moiré	(128)	+	--	--	/	+
Ellipsometry	(130)	+	□	+	?	+
Focus error detection	(131)	+	+	+	+	?
Absorption	(135)	+	?	?	+	?
Fluorescence	(136)	+	?	?	+	+
Ultrasonic	(138)	+	?	+	+	?

\+ = This criterion can be fulfilled;
□ = This criterion is not easily fulfilled;
\- = This criterion is hard to fulfil;
-- = This criterion can not be fulfilled;
? = The method needs more investigation on this point;
/ = This criterion does not apply to this method.

This scheme indicates among others the eventual disturbance of the lubricant film (point 1) and the possibility to realize the required measurement speed (point 4). Both points were not discussed for all methods individually, but it was considered that criterion 1 is fulfilled (or can be easily fulfilled) for all methods, apart from the mechanical, since the transducer is mounted in the rigid body. In the same way, criterion 4 is considered to be fulfilled when the response time depends in essence on the electronics only.

Choice of the method

As shown in table B1 the focus error detection method is the only one which in principle fulfils the first four criteria. The other methods either do not fulfil one or more of these criteria or need more investigation to find out whether these criteria are fulfilled. Focus error detection is therefore chosen for our film thickness measurements and will be further investigated.

Application of the focus error detection method needs more investigation on criterion 5, which in fact concerns the low reflectance of the elastomer to lubricant interface. A thin reflecting coating on the elastomer's surface should be applied if the reflectance on the elastomer to lubricant interface is not sufficient to obtain the required accuracy, but the eventual influence of such a coating is not quite clear yet.

APPENDIX C SET UP FOR THE TESTS

C1 Set up for the measurement of the signals

In chapter 3 and in appendix F and G, several measurements of the photodiode signals and the focus error signal, versus the distance between the focal point and a test surface, are presented. A simple set up was constructed for these measurements and will be described below.

It must be considered that the focus error device is very sensitive. The gradient in the focus error signal around the in-focus position (height $z = 0$) is e.g.

$$\frac{dffs}{dz} = 2.3 \quad [V \cdot \mu m^{-1}] \tag{C1}$$

for measurements on metals with the DWS (see section 3.1.2 page 26). The surface height must therefore be adjusted and measured very accurately (e.g. with an accuracy of 0.1 μm).

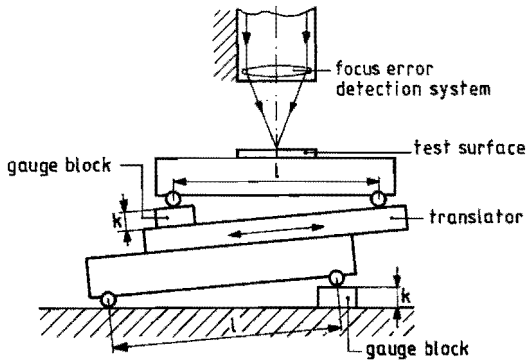
Also, the measurements must be performed on both a horizontal test surface as well as on a skew surface, with variable slope, to study the influence of the surface slopes on the signals.

Finally, the influence of the reflection on the window surfaces must be studied for different gap heights between the test surface and the window.

C1.1 Set up for the signal measurement on a horizontal test surface

The set up for the signal measurement on a horizontal test surface is shown in fig. C1. The height position of the test surface is adjusted with a translator, which is mounted under a small slope. The slope is determined by a gauge block under the translator, while a second gauge block with equal height is used to keep the test surface horizontal. A displacement of the translator will then cause a vertical displacement of the test surface, which is k/l-times smaller than the translator displacement. The signals can thus be measured as a function of the test surface height with a high resolution, provided that the test surface is flat and smooth.

Figure C1
Set up for the measurement of the signals of the focus error device as a function of the height of a horizontal surface.

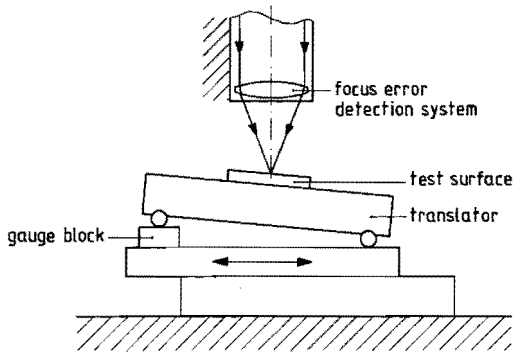


The distance l between the supports is 100 mm (with a maximum deviation of 0.05 mm) and the resolution of the adjustment of the translator displacement is 10 µm. The resolution of the surface height adjustment is then e.g. 0.1 µm, when gauge blocks of 1 mm high are used. Measuring the translator displacement with an accuracy of 1 µm decreases the uncertainty in the surface height to 0.01 µm. However, the height uncertainty is determined by the shape and roughness of the test surfaces, which were flat and smooth within 0.1 µm. The accuracy in the signal measurements is therefore 0.1 µm or better.

C1.2 Set up for the signal measurement with varying slopes of the test surface

The set up used to measure the influence of the surface slopes on the signals (fig. C2) is similar to the set up presented in the former section (see fig. C1), with the exception of the lower gauge block which is now missing. Now, the

Figure C2
Set up for the measurement of the signals of the focus error device as a function of the height at varying surface slopes.



test surface is not horizontal and its slope is determined by the gauge block height k and the support distance l (which is again 100 (± 0.05) mm). The translator is again used to adjust the surface height at the measurement spot accurately.

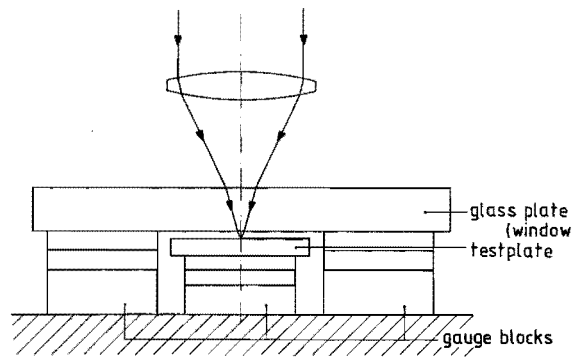
C1.3 Signal measurement with a glass plate on the test surface

The reflection on the window surfaces will probably influence the focus error signal. This influence is also studied by measuring the signals as a function of the surface height using the configuration shown in fig. C1, but with an additional 1.2 mm thick glass plate above the surface of the test plate (fig. C3). This glass plate can of coarse be laid on the test surface directly, as was e.g. done for the initial tests presented in fig. 3.15 (section 3.2.5 page 42), but the gap height between the test surface and the window must also be variable. The adjustment of this gap height is also performed using gauge blocks, as shown in fig. C3.

Two stacks of gauge blocks, with equal height, are used to support the glass plate, while the test plate is supported by a third stack, the height of which is varied to realize different gap heights. Both the glass and the test plate must be flat and smooth to obtain proper results.

When measurements are performed with a glass plate on or just above the test surface we must be aware of the light refraction on the window surfaces, causing spherical aberration (see section 3.2.4, pages 40, and appendix F3). The effect of aberrations is that the spot dimension is increased and this must be avoided,

Figure C3
Set up for the gap height adjustment between the test surface and the glass plate.



since it reduces the lateral resolution.

As already mentioned in section 3.1, the device for displacement and roughness measurements is derived from the compact disc transducer, the objective lens of which is specially designed to scan the CD-surface through the 1.2 mm thick protective layer, without suffering from significant aberrations. When the transducer is used for displacement or roughness measurements, the protective layer is not present and a 1.2 mm thick window is mounted just near the lens instead. Now, we can remove this window and use a 1.2 mm thick glass plate on or above our test surface without introducing spherical aberration.

C2 Set up for the measurement of the (roughness) profile

Some measurements were performed to test the behaviour of the optical sensor, when the profile and roughness of a test surface is scanned through a "window" (i.e. glass plate) laid on it (appendix G2.3). These tests were performed with the "double wedge" focusing device (see section 3.1.2, page 26) on the rig described by Struik and Chang (1987) and schematically shown in fig. C4.

The optical head, containing the focus error device, is moved downwards to the surface before the start of a measurement. As soon as the surface is in the focal point of the objective lens, which is determined by a zero focus error signal (see e.g. section 3.1.4.2), the motion of the optical head stops automatically. Now, the measurement can be started by moving the surface in the X-direction. The objective lens is then continuously focusing and will thus follow the height contour over a line track, while the optical head remains stationary.

The Y-axis translator can be used to scan several adjoining line tracks to obtain a three dimensional picture. Then two adjoining tracks are scanned in different direction, one in the positive and the other in the negative X-direction, as shown in fig. C5.

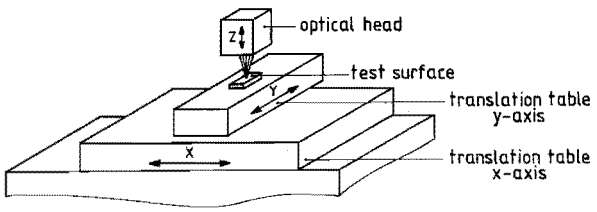
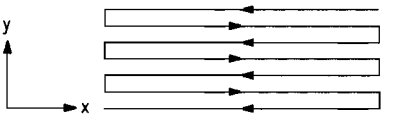


Figure C4
Test rig for optical profilometry.

Figure C5
Surface motion during 3-dimensional scanning.

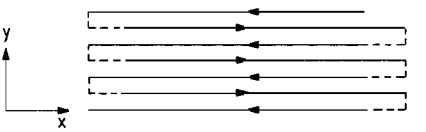


Scanning the test surface through a glass plate can be performed by laying the plate directly on the test surface. The set up shown in fig. C3 can be used, when a gap, with adjustable height, between the glass plate and the test surface is wanted. This glass plate must be about 1.2 mm thick and the window before the objective lens must be removed to keep the spherical aberration small (see also section C1.3 of this appendix).

Finally a problem, experienced with three-dimensional scanning and will now be briefly discussed. It originates from the fact that the test rig was initially developed for two-dimensional measurements only and it concerns the X-positioning of the tracks relative to each other.

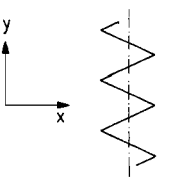
During the experiments of chapter 5, it appeared in some of the measurements that the X-position of the starting point of a new track was not equal to the final X-position of the former track. As a result the tracks, scanned in the negative X-direction, are shifted over some distance relative to the tracks in the positive X-direction (see fig. C6a). This shift appear as a ripple in the height contour lines, as e.g. shown in fig. C6b for a step-profile.

Figure C6
Shift between the tracks in the X-direction.

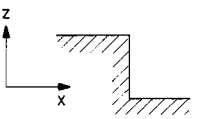


a. Pattern of the tracks.

b. Resulting height contour line of a step profile (see below).



c. Step in the profile



APPENDIX D THE REFLECTANCE ON THE GLASS AND ELASTOMER SURFACES

This appendix deals with the calculation of the reflectances on the glass and on the elastomeric surfaces. As discussed in section 3.2.5 and 3.2.6, reflection on the window to lubricant interface (fig. 3.13 page 38) should be avoided while the reflection on the elastomer to lubricant interface must be reasonable high to obtain the required accuracy.

Generally the reflectance $\mathcal{R}$ (i.e. the irradiance ratio of the reflecting and the incident beam) on the interface of two adjoining non-metallic media is given by

$$\mathcal{R} = \left(\frac{n_1 - n_2}{n_1 + n_2} \right)^2 \qquad \text{(D1)}$$

for a beam at normal incidence (see e.g. Hecht 1987 pp. 94-104) (n_1 and n_2 are the respective indices of refraction of the adjoining media).

In our configuration, the incident beam is not collimated but convergent, i.e. the angle of incidence is different for the different rays. The maximum angle of incidence is 0.47 rad (27°) for a numerical aperture of the objective lens of 0.45 (see section 3.1.2 and 3.1.3, pages 26ff.). At this angle the reflectance is not much higher than for normal incidence (see e.g. Hecht, 1987 fig. 4.28 p. 103). The reflectance of the convergent beam will therefore be estimated by the normal incidence equation (D1).

D1 The reflectance on the glass to lubricant interface

Most optical glasses have an index of refraction in the range of roughly 1.4 to 1.7. The reflectance on an air to glass interface is then in the range of 2.8 to 6.7 percent, since the index of refraction of air is 1.

The presence of the lubricant film between the window and the elastomer (see fig. 3.13 page 38) reduces the reflection on the window surface at the side of the lubricant film and the elastomer, since most lubricants have an index of refraction also in the range of 1.4 and 1.7 (the reflectance $\mathcal{R}$ is e.g. 0.94 percent for the extreme case that $n_1 = 1.4$ and $n_2 = 1.7$).

As discussed in section 3.2.5 and appendix G2.3 the reflection on the lubricant to window interface should be eliminated. This is realized when a liquid is used with an equal index of refraction as the window, as expressed by eq. (D1). To find such a lubricant, the index of refraction has been measured and is given in table D1. These measurements could, however, not be performed for the wavelengths of the focus error devices (820 nm for the double wedge system (section 3.1.2, page 26) and 780 nm for the diffractive system described in section 3.1.3, page 28). Therefore, the indices of refraction were measured at two different wavelengths to get an idea of the wavelength dependence.

Table D1 The index of refraction n of different lubricants, measured for two wavelengths λ at a temperature of 20 °C and at atmospheric pressure. (The values for the Shell Flex oils were obtained from Shell)

Lubricant:	λ = 598.3 nm	λ = 670 nm
Shell Omala 220	1.4909	1.4874
Shell Ondina 32	1.4709	1.4684
Shell Ondina 68	1.4800	1.4771
Shell Tellus 46	1.4818	1.4768
Shell Tellus 100	1.4855	1.4821
Shell Tellus T46	1.4791	1.4764
Shell Tellus C320	1.4920	1.4890
Shell Tonna 220	1.4890	1.4855
Shell Flex 410 HP	1.4849	
Shell Flex 790 HP	1.4939	
Shell Flex 532 GH	1.499	
Shell Flex 782 GH	1.511	
Shell Flex 792 GH	1.520	
Mobil Vactra 4	1.4882	1.4851

The window used in the experiments of chapter 3 and 5 was a 1.2 mm thick plate of Duran glass. The index of refraction of this glass is

1.4722 (λ = 589.3 nm)
1.4701 (λ = 670 nm)

Using Duran glass in combination with a mixture of 75 percent Shell Ondina 15 and 25 percent of Shell Ondina 68 appears to eliminate the reflection on the window to lubricant interface adequately (see appendix G2.3). The index of refraction of this mixture is

1.4733 (λ = 589.3 nm)
1.4708 (λ = 670 nm)

D2 The reflectance on the elastomer to lubricant interface

The film thickness and the (eventual deformed) roughness texture will be determined by measurement of the height position of the elastomeric surface relative to the focal point of the objective lens (see section 3.2, page 37ff.). Therefore a reasonable reflectance on the lubricant to elastomer interface is needed for proper measurement (see section 3.2.6, page 45). This reflectance can also be estimated using eq. (D1), provided that the respective indices of refraction of the lubricant and the elastomer are known.

The index of refraction of different lubricants is given in table D1 above. The index of refraction of soft elastomers (which are used for seals e.g.), however, is not easily determined, mainly because of the roughness of the surfaces which scatter the light too much for proper measurement of the index of refraction. This roughness is formed during the injection moulding process and is not easily avoided (see e.g. Kanters, 1990 section 4.4.1.1 p. 66, where it is shown that the use of a very smooth moulding form did not yield very smooth elastomeric specimen). Nevertheless some values of the index of refraction of elastomers can be found in literature (see table D2), but this information is not very detailed.

We can conclude from table D1 and D2 that the indices of refraction of lubricants and elastomers are in the same range. The reflectance on the lubricant to elastomer interface is therefore expected to be low (smaller than 1 percent when both indices of refraction are in the range of 1.4 to 1.7).

Table D2 The index of refraction n of some elastomers at wavelength λ

Material:	λ [nm]	n [-]	References
Polyurethane	not given	1.488	Field (1973 p. 72a)
Silicon rubber	not given	1.40	Field (1973 p. 72a)
Silicon rubber	546.1	1.43	Kalsi (1975 p. 27)
Synthetic Polyisoprene	546.1	1.527	McClune (1974 p. 35)

APPENDIX E THE INFLUENCE OF SURFACE SLOPES ON THE FOCUS ERROR SIGNAL

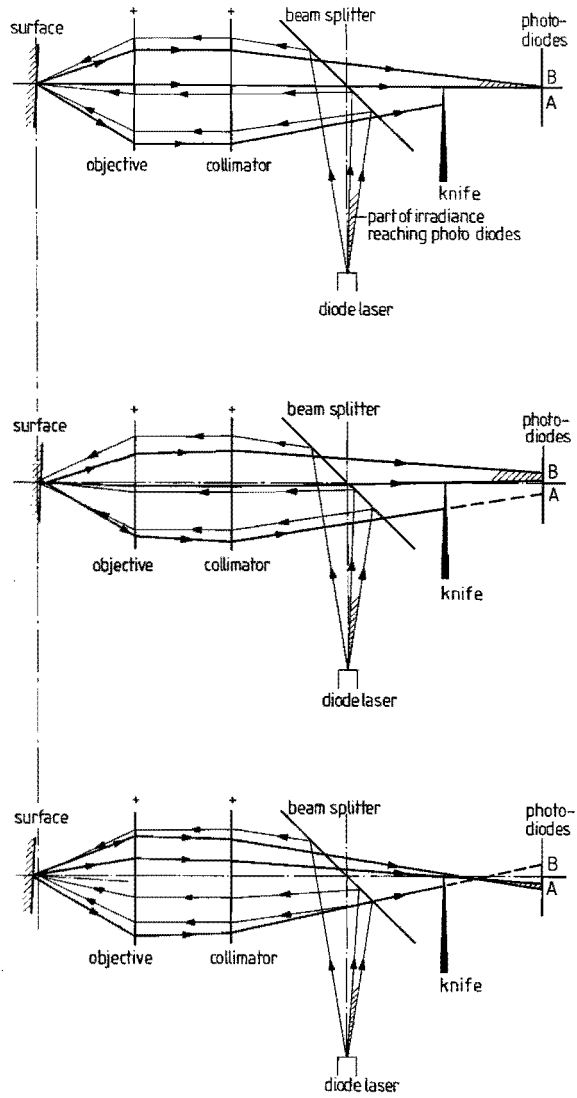
In section 3.2.2 (page 40) was discussed that the objective lens must be fixed for the film thickness measurements because of the required frequency range. The film thickness will then be derived from the measured focus error signal directly and consequently we must account for the possible influence of the surface slopes on the focus error signal. This influence of the surface slopes is illustrated in fig. E1 and, for slopes in the opposite direction, in fig. E2. These figures can be compared with fig. 3.2 (page 25), where the slopes are zero, i.e. the surface is perpendicular to the optical axis.

Both figures show that the beam reflected on the surface is still focused on the boundary of the photodiodes when the surface is in focus. Therefore the in-focus position still yield a zero focus error signal (this is why the surface slopes have no influence when the measurements are performed in the open loop mode, see section 3.1.4.2 page 34), but the irradiance received by the photodiodes is now different (the irradiance is e.g. larger in fig. E1a and smaller in fig. E1b compared with the irradiance at zero slope, fig. 3.2 page 25).

Fig. E1 further show that part of the light reaches the "obscured" photodiode (diode A in fig. E1b and diode B in fig. E1c) and the "obscured" photodiode will therefore yield a larger signal than at zero slope. When the slope is in the opposite direction, no light reaches the obscured photodiode (fig. E2b/c) and they will therefore a smaller signal than at zero slope.

In the systems which are used in practice for the displacement and roughness measurements (see section 3.1.2 and 3.1.3, pages 26ff.), the situation of fig. E1 and fig. E2 occur simultaneously because of the symmetry of the devices. One diode pair (e.g. A_1B_1) will receive more light, as in fig. E1, when the other pair (A_2B_2) receives less light, as in fig. E2. Then it is perhaps possible that the effect on one diode pair cancels the effect on the other pair, with the result that the signals A and B (being $(A_1 + A_2)$ and $(B_1 + B_2)$ respectively), and thus the focus error signal) are not influenced by the surface slopes. However, it is not sure that the signals are really independent of the slope and some measurements have been performed to test this.

Figure E1
Influence of the surface slopes on the measurement.



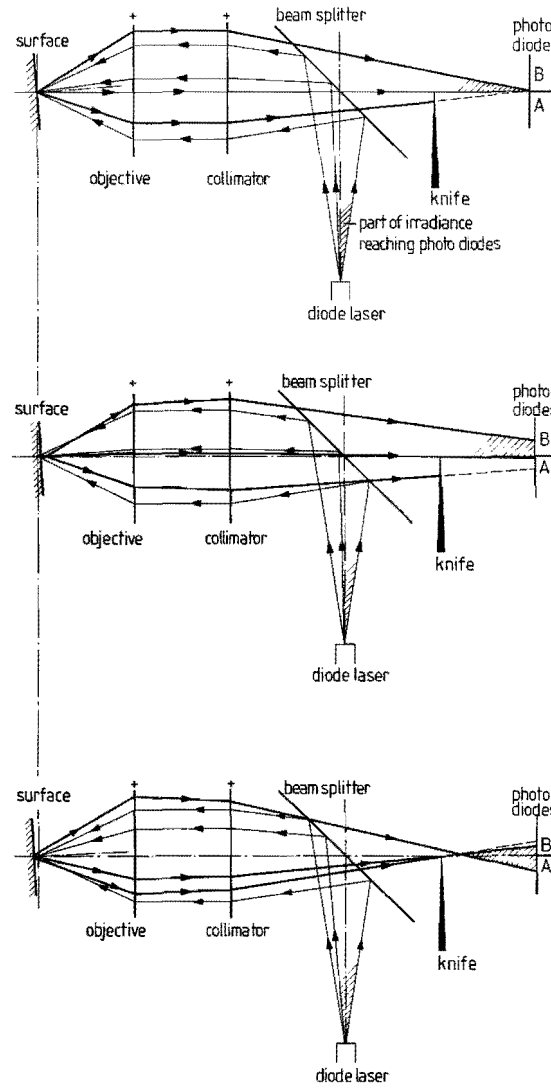
a. Surface in focus.

b. Surface in front of the focus.

c. Surface beyond the focus.

When the slope is in a direction perpendicular to the direction shown in fig. E1 and E2, the signals will probably be independent of the surface slopes. This is illustrated by fig. E3, where a cross section of the focus error device, perpendicular to the cross section of fig. E1 and fig. E2, is shown. Now the out-of-focus position of the surfaces is the same in all three situations drawn, but the slopes are different. At these different slopes, the spot on the

Figure E2
Influence of the surface slopes on the measurement (surface slope in the opposite direction compared with fig. E1).



a. Surface in focus.

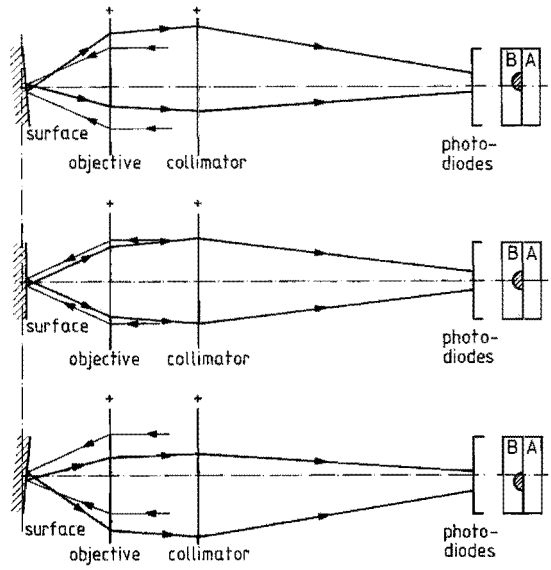
b. Surface in front of the focus.

c. Surface beyond the focus.

photodiodes is at a different position, but it is not (whole or partly) shifted from one diode to the other. Therefore the focus-error signal will be equal, regardless of the slope.

The measurements (presented in section E1 below) show that the influence of the slopes on the focus error signal is indeed negligible for slopes the direction shown in fig. E3 (in the following referred to as the "X-direction"), while the influence is significant for surface slopes in the other direction shown in fig.

Figure E3
Influence of the slope on the measurement (surface slope perpendicular to the direction shown in fig. E1 and E2).



E1 and E2 (in the following referred to as the "Y-direction"). This influence will be eliminated using the so-called radial error signal, as discussed in section 3.1.4.2 (page 35).

Now the measurements of the slope influence on the focus error signal will be presented, as well as some measurements of the radial error signal.

E1 Measurement of the signals for different surface slopes

The measurements to determine the slope influence on the photodiode signals and on the focus error signal were performed using the set up described in appendix C2.1 and a displacement and roughness sensor based on the diffractive focus error detection system described in section 3.1.3 (page 28). This system was chosen for these experiments (instead of the double wedge system) because the film thickness transducer will be based on this diffractive element.

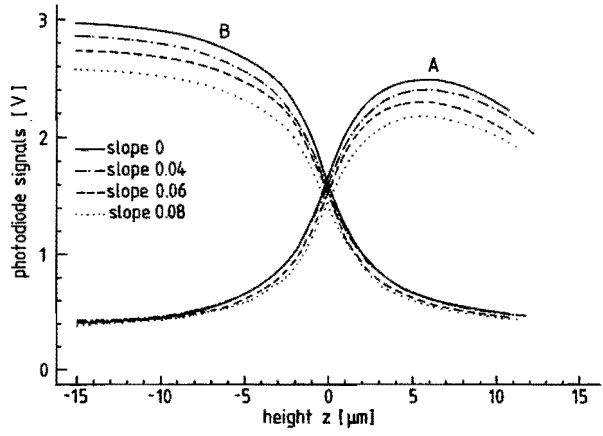
The measurements were performed for surface slopes in both the X-direction (fig. E3) and the Y-direction (fig. E1 and E2).

Measurement in the X-direction

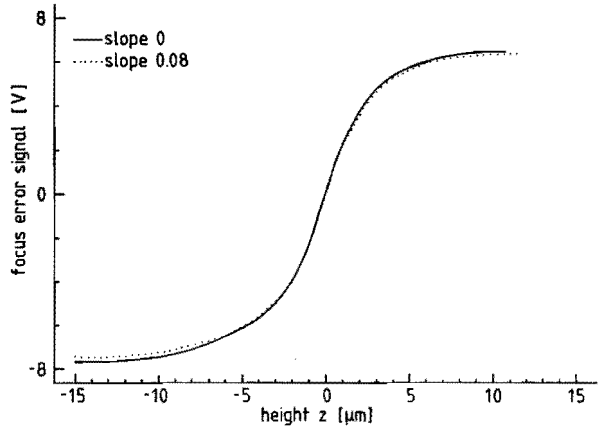
The results of the measurements on a slope in the X-direction (fig. E3) are given in fig. E4. As shown in fig. E4a the photodiode signals A and B are influenced by the surface slope. This is caused by the fact that the reflected light cone, returning into the system through the objective lens, is partly blocked by the diaphragm, as illustrated in fig. E5. Consequently the irradiance measured by the photodiodes, and thus the signal from the photodiodes, is lower.

The relative decrease in the irradiance (measured by the photodiodes) is equal for all photodiodes and the focus error signal (which is $(A - B)/(A + B)$, see section 3.1.4.2 page 32ff.) is therefore independent of the slopes, as shown in fig. E4b.

Figure E4
Measures signals at different slopes in X-direction (fig. E3)

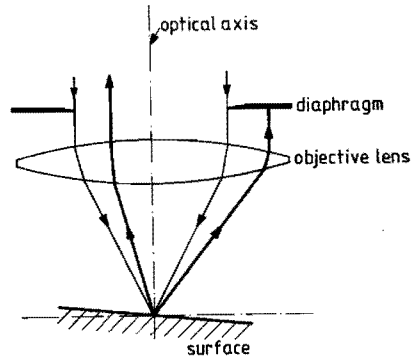


a. Photodiode signals



b. focus error signal

Figure E5
Blockage of part of the returning light by the diaphragm.



Measurement in the Y-direction

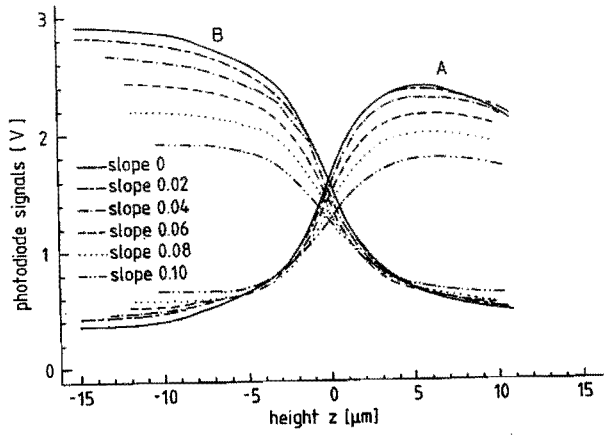
As shown in fig. E6a the influence of the surface slopes in the Y-direction on the photodiode signals *A* and *B* is different from the influence of slopes in the X-direction. In the region where the signal is large (e.g. at height $z > 0$ μm for signal *A*), the signal decreases at increasing slope, and this decrease is larger than found for slopes in the X-direction (fig. E4a). In the region where the signal is low (e.g. at height $z < -5$ μm for signal *A*) it increases at increasing slope and this was not found at slopes in the X-direction. Consequently the focus error signal is influenced by surface slopes in the Y-direction (see fig. E6b).

As discussed in section 3.1.4.2 the radial error signal can be used to eliminate this slope influence on the focus error signal (and thus on the film thickness measurement). Therefore some measurements of the radial error signal will now be presented.

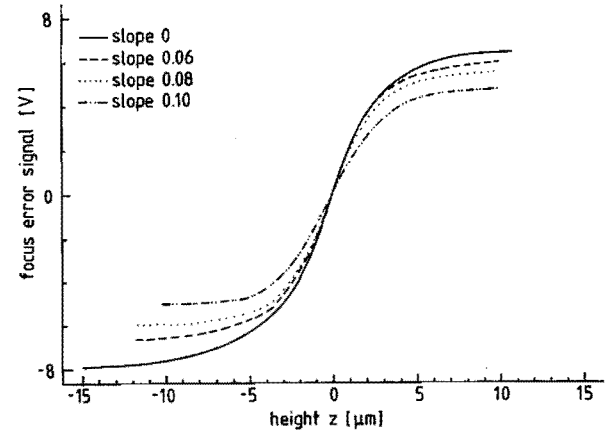
E2 Measurement of the radial error signal

The radial error signal was measured for surface slopes in the Y-direction, where the slope influence on the focus error signal appeared to be large. As shown in fig. E7 the radial error signal is almost independent of the surface height *z*, when *z* is roughly between -5 and + 5 μm, i.e. in the range where the slope in the focus error signal is steep (fig. E6b). Therefore the surface slope can be derived from the focus error signal directly and used to determine the surface height from the focus error signal.

Figure E6
Measured signals at different slopes in Y-direction (fig. E1 and E2)

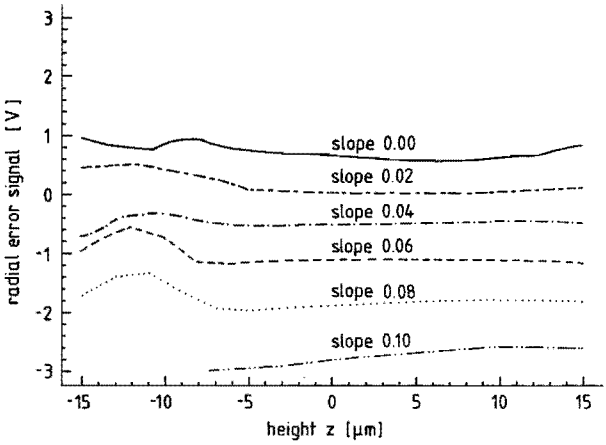


a. Photodiode signals



b. focus error signal

Figure E7
Measured radial error signals at different slopes in Y-direction (fig. E1 and E2)



APPENDIX F THE DIMENSION AND THE IRRADIANCE DISTRIBUTION OF THE FOCUS SPOT

An aberration free, diffraction limited focus spot is preferred to obtain the smallest spot dimension and thus the highest lateral resolution (see section 3.1.4.1 page 29). However, the non-uniform irradiance distribution of the incident beam increases the spot width. The window (necessary in our experiments, see section 3.2 page 37), introduces spherical aberration which also spread the radiation over a larger area. Both effects thus reduce the lateral resolution.

This appendix deals with the influence of the irradiance distribution of the incident beam and with the influence of the spherical aberration on the irradiance distribution and the dimension of the focus spot.

First the dimension of the diffraction limited spot will be calculated (assuming a uniform irradiance distribution of the incident beam), followed by consideration of the influence of the non-uniform irradiance distribution of the incident beam. This will yield a criterion for the maximum allowable numerical aperture of the collimator lens, assuring that the irradiance distribution of the incident beam is sufficiently uniform to avoid significant increase of the spot size.

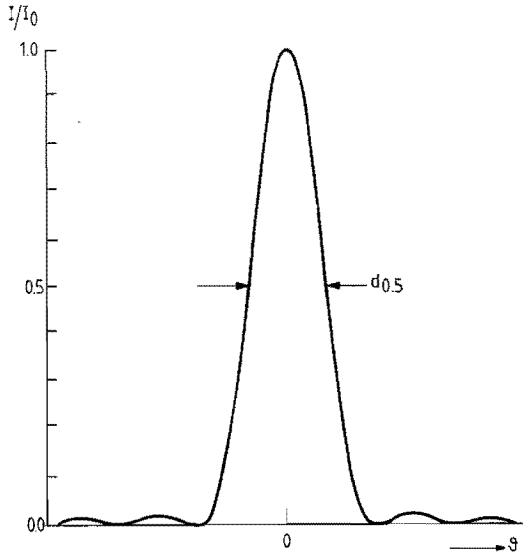
Finally the influence of the spherical aberration on the irradiance distribution of the spot will be quantified and a criterion for the maximum tolerable spherical aberration (in terms of the window thickness, its index of refraction and the numerical aperture of the objective lens), for which we can still regard the spot as diffraction limited, will be formulated.

F1 The dimension of the diffraction limited spot

The theory of this section can be derived from basic works on optics, e.g. Born and Wolf (1970, section 8.5.2, pp. 395ff.) or Hecht (1987, section 10.2.5, pp. 416ff.). The main assumptions are, that the irradiance distribution of the incident beam is uniform and the system is free of aberrations.

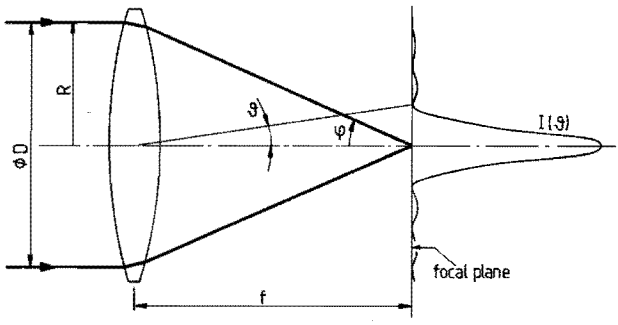
The irradiance distribution of a diffraction limited spot, the so-called Airy pattern, is given in fig. F1. In this figure, v characterizes the radial position in the focal plane (fig. F2). The dimension of the spot will now be

Figure F1
Irradiance distribution of the diffraction limited spot (Airy pattern).



($d_{0.5}$ is the fifty percent-irradiance width, i.e. the width of the irradiance distribution at half of the irradiance maximum I_0)

Figure F2
Focusing arrangement of an aberration free objective lens.



characterized by the fifty-percent-irradiance width $d_{0.5}$, i.e. the position, where the irradiance is half of its maximum.

The irradiance at position ϑ is given by

$$I(\vartheta) = I_0 \left[\frac{2\ J_1(kR\ \sin\vartheta)}{kR\ \sin\vartheta} \right]^2 \qquad \text{(F1)}$$

in which

I_0 = irradiance maximum (occurring at $\vartheta = 0$)
J_1 = Bessel function of the first kind and of order zero
k = propagation number = $2\pi/\lambda$
R = half diaphragm diameter = $D/2$

To find the fifty-percent-irradiance width, we must solve eq. (F1) for

$$I(\vartheta_{0.5}) = I_0/2$$

in which $\vartheta_{0.5}$ is the angle of the position, where the irradiance is half of its maximum.

Therefore

$$I_0 \left[\frac{2\ J_1(kR\ \sin\vartheta_{0.5})}{kR\ \sin\vartheta_{0.5}} \right]^2 = \frac{I_0}{2}$$

and thus

$$J_1(kR\ \sin\vartheta_{0.5}) = \frac{1}{4}\sqrt{2}\ kR\ \sin\vartheta_{0.5}$$

$$\approx 0.354\ kR\ \sin\vartheta_{0.5}$$

which is true for

$$kR\ \sin\vartheta_{0.5} \approx 1.6$$

(then, $J_1(kR\ \sin\vartheta_{0.5}) = 0.5699 = 0.356\ kR\ \sin\vartheta_{0.5}$)

Considering that $\vartheta_{0.5} \ll 1$ (which is generally true since the focus spot is much smaller than the focal length f), the fifty-percent-irradiance width can be written as

$$d_{0.5} = 2f\tan\vartheta_{0.5} \approx 2f\sin\vartheta_{0.5} = 2f\frac{1.6}{kR} = \frac{3.2}{\pi}\lambda\frac{f}{D}$$

Introducing the numerical aperture

$$NA = n\sin\varphi \qquad \text{(F2)}$$

in which n is the index of refraction of the medium in which the light propagates.

Then we can write for small φ and for $n = 1$ (air)

$$NA \approx \frac{D}{2f}$$

(the error is then e.g. only 2 percent for $NA = 0.2$)

and we find finally

$$d_{0.5} \approx \frac{\lambda}{2\ NA} \qquad \text{(F3)}$$

F2 The influence on the spot size of the nonuniform irradiance distribution of the incident beam

The dimension of the diffraction limited spot, as calculated in section F1 above, is only valid for an incident beam with a uniform irradiance distribution over the diaphragm cross section. In practice, however, the irradiance distribution is rarely uniform and consequently the spot diameter will be larger.

In this section we will first consider the irradiance distribution of diode laser beams, followed by calculation of the maximum diameter of the beam (in terms of the numerical aperture of the collimator lens, see section F2.1 below) to assure that the irradiance distribution over the diaphragm cross section is sufficiently uniform to prevent too much increase in the spot size.

F2.1 The irradiance distribution of diode laser beams

Diode lasers have a Gaussian beam, i.e. the electrical field amplitude distribution is Gaussian. Referring to fig. F3 and assuming the beam to be

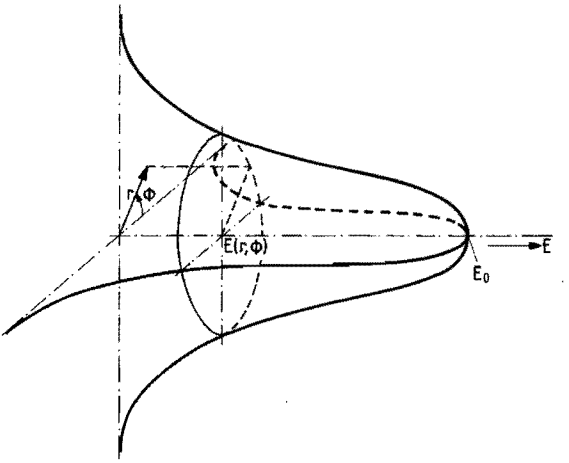


Figure F3
The Gaussian amplitude distribution.

cylindrical, the electrical field amplitude distribution $E(r_n,\phi)$ is given by (see e.g. Bouwhuis et al., 1987 section 2.3.2 p. 30)

$$E(r_n, \phi) = E_0 \, e^{\left(-\frac{\sigma}{2} r_n^2 \right)} \qquad \text{(F4)}$$

in which

E_0	=	the electrical field amplitude maximum	[Vm^{-1}]
r_n	=	the *normalized* radial position (i.e. r_n is 1 at the rim of the diaphragm)	[-]
ϕ	=	circumferential position	[rad]

The electrical field amplitude at the rim of the diaphragm is then

$$E_r = E(1, \phi) = E_0 \, e^{-\sigma/2}$$

and the irradiance

$$I_r = I(1, \phi) = I_0 \, e^{-\sigma}$$

Further the beam from a diode laser is not collimated but divergent and a cross section of the beam is not cylindrical but elliptical (see e.g. Hall and Jackson (1989) for the theoretical background).

According to Bouwhuis et al. (1987 section 2.6.1 pp. 81-82) the electrical field amplitude of the AlGaAs laser, used in the focus error detection systems, is about 50 percent of the amplitude maximum E_0 (i.e. $\sigma \approx 1.5$) for rays diverging at an angle of 5° in one direction and 20° in the perpendicular direction.

Now we will consider the smallest allowable value of the electrical field amplitude at the rim of the diaphragm, for which the irradiance distribution over the diaphragm cross section is sufficiently uniform to keep the spot diffraction limited. This means that the maximum tolerable diaphragm diameter will be calculated for a given amplitude distribution of the beam. Or, in the case of the diode laser, the maximum tolerable numerical aperture, since the maximum tolerable diaphragm diameter depends on the axial position of the diaphragm relative to the laser because of the divergence of the beam.

F2.2 The maximum tolerable numerical apperture of the colloimator lens

As shown by Bouwhuis et al. (1987, fig. 2.16 p. 31) the spot size is not much influenced by the Gaussian, non-uniform, amplitude distribution, when

$$\sigma \leq 1.5$$

When $\sigma = 1.5$ (i.e. the electrical field amplitude at the rim E_r is 0.47 times the maximum amplitude E_0, the fifty-percent-irradiance width of the spot is increased by not more than 10 percent. Now we will allow a maximum amplitude difference between the centre of the diaphragm and the rim of about 50 percent. For the AlGaAs laser, this amplitude is reached for rays diverging at 5° ($\approx$ 0.087 rad) in the direction of the shortest axis of the beam ellipse. The criterion for the numerical aperture of the collimator lens is therefore

$$(NA)_{coll} \leq 0.1$$

F3 The decrease in the irradiance maximum due to spherical aberration

Consider now fig. F4. The rays at the left-hand side of the window are regarded as aberration free, i.e. the wave front is perfectly spherical and when the window would not be present, all rays would intersect the optical axis at the same position F where a diffraction limited spot would be formed. The window, however, refracts the rays and consequently the wave fronts are aspherical while

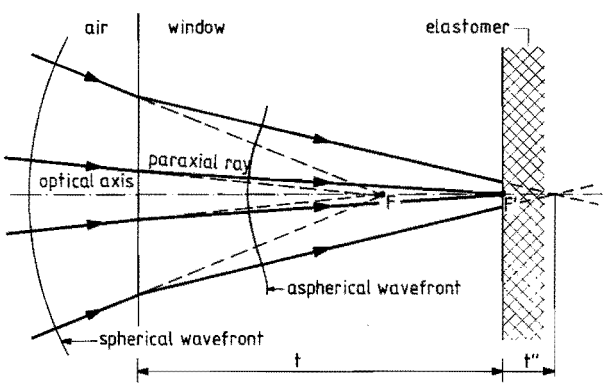


Figure F4
Ray propagation after refraction at the window surface.

the rays at different angle of incidence intersect the optical axis at different positions. The light is thus spread over a larger area and a diffraction limited spot is not formed.

In fact, all "ray-cones" with their own top angle (i.e. angle of incidence) have their own focal point. The rays with a small angle of incidence (referred to as "paraxial rays") have their focus in point F' (the "paraxial focus"), while the other rays (with larger angle of incidence) have their focal point at the right-hand side of F'.

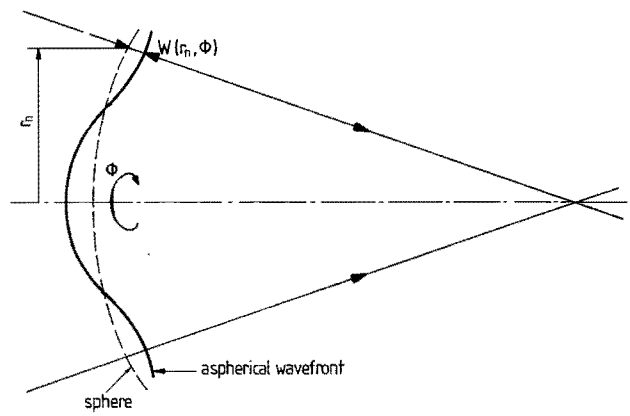
The decrease in the irradiance maximum can be derived from the so-called **wave aberration**, characterized by the **aberration function** $W(r_n,\phi)$, which represents the deviation of the wavefront from the ideal spherical shape at the radial position r_n and the circumferential position ϕ (see fig. F5). For primary aberrations, the aberration function can be represented by power series of the form (see e.g. Born and Wolf, 1970, section 9.2 pp. 464ff.)

$$W(r_n,\phi) = W_{ij}\, r_n^i \cos^j\phi \qquad \text{(F5)}$$

in which

W	=	Aberration function	[m]
W_{ij}	=	Representation of different kinds of aberration	[m]
r_n	=	the *normalized* radial position (i.e. r_n is 1 at the rim of the diaphragm)	[-]
ϕ	=	circumferential position	[rad]

Figure F5
Wave aberration $W(r_n,\phi)$



A criterion for the maximum tolerable aberration can be derived from the decrease in the irradiance maximum I_0 of the diffraction limited spot (fig. F1, page 160). This decrease ΔI_0 is e.g. given by Born and Wolf (1970 p. 469) and Bouwhuis et al. (1987 p. 34)

$$\frac{\Delta I_0}{I_0} = 4\pi^2 \frac{V_W}{\lambda^2} \tag{F6}$$

in which V_W is the variance of W.

The spot is then regarded diffraction limited, when the decrease in the irradiance maximum is less than 20 percent[1]

$$\frac{\Delta I_0}{I_0} \leq 0.2 \tag{F7}$$

i.e.

$$V_W = \text{var}W \leq 0.005\ \lambda^2 \tag{F8}$$

When calculating the maximum allowable spherical aberration, we must consider that the "best focus" is not coincident with the paraxial focus F'. Therefore, we need two parameters to calculate the influence of spherical aberration on the irradiance distribution (see Bouwhuis et al., 1987 p. 30 and 41):

- The parameter W_{40} describing the spherical aberration;
- The parameter W_{20} describing the effect of a distance t'' between the best and the paraxial focus (see fig. F4).

Then eq. (F5) yields

$$W(r_n,\phi) = W_{40}\, r_n^4 + W_{20}\, r_n^2 \tag{F9}$$

in which

$$W_{20} = \frac{-t''\ (NA)^2}{2} \tag{F10a}$$

and

$$W_{40} = \frac{n^2 - 1}{8\ n^3}\ t\ (NA)^4 \tag{F10b}$$

The best focus is at the position t'', where V_W (= var$W$) has its minimum, i.e.

---

[1] Maréchal's criterion, see e.g. Born and Wolf (1970 p. 469) and Bouwhuis et al. (1987 p. 34).

$$\frac{dV_W}{dt''} = 0 \quad y \quad \frac{d^2V_W}{dt''^2} > 0 \tag{F11}$$

Using eq. (F9), we find

$$V_W = \text{var}W = \overline{W^2} - (\overline{W})^2 =$$

$$= \frac{1}{\pi} \int_0^{2\pi} \int_0^1 (W_{40}\, r_n^4 + W_{20}\, r_n^2)^2\, r_n\, dr_n\, d\phi +$$

$$- \left[\frac{1}{\pi} \int_0^{2\pi} \int_0^1 (W_{40}\, r_n^4 + W_{20}\, r_n^2)\, r_n\, dr_n\, d\phi \right]^2 =$$

$$= \frac{4}{45} W_{40}^2 + \frac{1}{6} W_{40}\, W_{20} + \frac{1}{12} W_{20}^2 \tag{F12}$$

For constant *NA*, eq. (F11) equals

$$\frac{dV_W}{dW_{20}} = 0 \quad \wedge \quad \frac{d^2V_W}{dW_{20}^2} > 0$$

yielding

$$\frac{dV_W}{dW_{20}} = \frac{1}{6} W_{40} + \frac{1}{6} W_{20} = 0$$

and thus

$$W_{20} = -W_{40} \tag{F13}$$

while

$$\frac{d^2V_W}{dW_{20}^2} = \frac{1}{6} > 0$$

Eq. (F13) represents therefore a minimum indeed and we find from eq. (F12)

$$V_W = \frac{1}{180} W_{40}^2$$

Using the criterion for the maximum aberration (eq. F8), we find finally the criterion for the maximum tolerable spherical aberration

$$W_{40} = \frac{n^2 - 1}{8\, n^3}\, t\, (NA)^4 \leq 0.95\, \lambda \tag{F14}$$

APPENDIX G THE INFLUENCE OF THE LOWER WINDOW SURFACE REFLECTION ON THE MEASUREMENTS

Measuring the lubricant film thickness with the focus error device needs the contact to be optical accessible, i.e. the elastomeric surface must be scanned through a window in the rigid surface. This window, and the lubricant film between the window and the elastomer, introduces a second reflecting surface (the window to lubricant interface), just near the elastomeric surface (fig. G1).

Scanning the height of the elastomeric surface, we must account for the influence of this lower window surface[1], since both surfaces contribute to the focus error signal. This means that the focus error signal is different from the case, when reflection only occurs on the elastomer and, in the closed loop mode (see section 3.1.4, page 29), the objective lens is not focused on the elastomeric surface, but somewhere in the neighbourhood. In this respect, the measurement is different from the compact disc scanning and from the displacement, shape and roughness measurements, where only one reflecting surface contributes to the focus error signal.

In this appendix, we will study the influence of this extra reflecting surface on the working of the system. First the resultant focus error signal will be derived, followed by analysis of the objective lens response on a film thickness variation in the closed loop mode. Finally, some roughness measurements on a test surface with a glass plate thereon will be presented for verification.

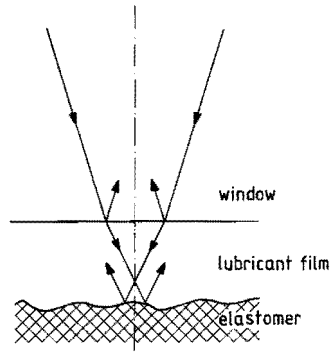


Figure G1
Set up for the lubricant film thick-ness measurement.
Reflection occurs on both the lubricant to elastomer interface and on the lubricant to window interface.

[1] The upper window surface has no significant influence, since it is too far from the contact and thus too far out of focus (see section 3.2.5)

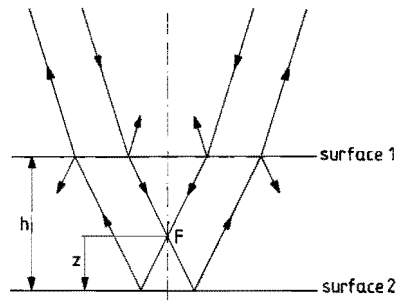
G1 The influence of the lower window surface reflection on the focus error signal

In this section, the focus error signal will be derived theoretically for the configuration with two reflecting surfaces just above each other, as e.g. shown in fig. G1. First, we will derive a general expression for the focus error signal as a function of the focal point position and of the gap height, followed by a reconstruction of the focus error signal for some gap height values.

G1.1 General expression for the focus error signal

Consider now the situation of two reflecting surfaces as shown in fig. G2. The distance between the surfaces is h, while the position of the focal point F is characterized by its distance z to surface 2. Now, we will derive a general expression for the focus error signal for this situation. First, an expression will be derived for single reflection (i.e. not accounting for the rays which reflect more than ones in the gap), followed by analysis, accounting for this multiple reflection. Finally, a short discussion will be presented on the focal point position for zero focus error signal, i.e. the position on which the objective lens would be focused by the servo controller.

Figure G2
Single reflection on two parallel reflecting surfaces.



Single reflection approximation

Consider fig. G2 again. Both surface 1 and surface 2 has an out of focus, which is given by $(z - h)$ and by z respectively. We can then derive the signals A and B from the photodiodes for each surface individually, i.e. the signals are derived for surface 1 as if surface 2 does not exist, and equally for surface 2,

as if surface 1 does not exist. Assuming that the signals are proportional with the irradiance on the photodiode, we can write the signals as a function of z-h and z respectively

$$A^{(1)} = \mathcal{R}_1 A_n(z-h) \quad \text{(G1a)}$$

$$B^{(1)} = \mathcal{R}_1 B_n(z-h) \quad \text{(G1b)}$$

and accounting for the loss of light by the reflection on surface 1

$$A^{(2)} = (1 - \mathcal{R}_1)^2 \mathcal{R}_2 A_n(z) \quad \text{(G1c)}$$

$$B^{(2)} = (1 - \mathcal{R}_1)^2 \mathcal{R}_2 B_n(z) \quad \text{(G1d)}$$

in which:

$A^{(1)}$ and $B^{(1)}$	are the signals, resulting from the reflection on surface 1;
$A^{(2)}$ and $B^{(2)}$	are the signals, resulting from the reflection on surface 2;
A_n and B_n	are the normalized signals (i.e. the signals for 100 percent reflection on one surface), which can be derived from section 3.1.4.2 (fig. 3.8 page 33);
$\mathcal{R}_1$ and $\mathcal{R}_2$	are the reflectances on surface 1 and surface 2 respectively ($0 \leq \mathcal{R} \leq 1$).

We will now assume, that the total signals from the photodiodes are equal to the sum of the signals, which would result from the reflection on the surfaces individually (possible interference caused by a phase difference between the different reflected rays is thus not accounted for), i.e.

$$A = A^{(1)} + A^{(2)} \quad \text{(G2a)}$$

$$B = B^{(1)} + B^{(2)} \quad \text{(G2b)}$$

and using eq. 3.4 (section 3.1.4.2, page 33), we find for the focus error signal

$$fes = \frac{\mathcal{R}_1\left[A_n(z-h) - B_n(z-h)\right] + (1 - \mathcal{R}_1)^2\mathcal{R}_2\left[A_n(z) - B_n(z)\right]}{\mathcal{R}_1\left[A_n(z-h) + B_n(z-h)\right] + (1 - \mathcal{R}_1)^2\mathcal{R}_2\left[A_n(z) + B_n(z)\right]} \quad \text{(G3)}$$

Multiple reflection analysis

The multiple reflection in the gap between both surfaces introduces images of the surfaces, which also contribute to the focus error signal (fig. G3).

We can write for the contribution of the first image

$$A^{(i1)} = (1 - \mathcal{R}_1)^2 \, \mathcal{R}_1 \, \mathcal{R}_2^2 \, A_n(z+h)$$
$$B^{(i1)} = (1 - \mathcal{R}_1)^2 \, \mathcal{R}_1 \, \mathcal{R}_2^2 \, B_n(z+h)$$

and, generally, for the n^{th} image

$$A^{(in)} = (1 - \mathcal{R}_1)^2 \, \mathcal{R}_1^n \, \mathcal{R}_2^{n+1} \, A_n(z+nh) \qquad \text{(G4a)}$$
$$B^{(in)} = (1 - \mathcal{R}_1)^2 \, \mathcal{R}_1^n \, \mathcal{R}_2^{n+1} \, B_n(z+nh) \qquad \text{(G4b)}$$

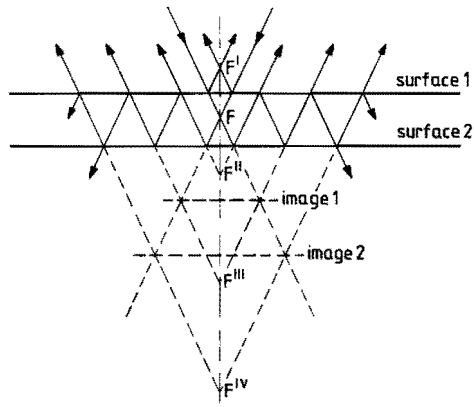
The total photodiode signals are then (again neglecting the possible influence of interference)

$$A = A^{(1)} + A^{(2)} + \sum_{n=1}^{\infty} A^{(in)} \qquad \text{(G5a)}$$

$$B = B^{(1)} + B^{(2)} + \sum_{n=1}^{\infty} B^{(in)} \qquad \text{(G5b)}$$

The general expression for the focus error signal is then finally

Figure G3
Multiple reflection on two parallel reflecting surfaces.



$$fes = \frac{\mathcal{R}_1\left[A_n(z\text{-}h)\text{–}B_n(z\text{-}h)\right] + (1\text{–}\mathcal{R}_1)^2\mathcal{R}_2 \sum_{n=0}^{\infty} \mathcal{R}_1^n\mathcal{R}_2^n\left[A_n(z\text{+}nh)\text{–}B_n(z\text{+}nh)\right]}{\mathcal{R}_1\left[A_n(z\text{-}h)\text{+}B_n(z\text{-}h)\right] + (1\text{-}\mathcal{R}_1)^2\mathcal{R}_2 \sum_{n=0}^{\infty} \mathcal{R}_1^n\mathcal{R}_2^n\left[A_n(z\text{+}nh)\text{+}B_n(z\text{+}nh)\right]} \qquad \text{(G6)}$$

G1.2 The position of the focal point for zero focus error signal

We can now in principle derive the focal point position, where the focus error signal is zero. This position determines the position of the objective lens, when the servo controller is used to focus the lens. We will again distinct between single and multiple reflection.

To derive the position of zero focus error signal, the sign of the detector signals must be known in the different points. We can derive from section 3.1.4.2 (fig. 3.8 page 33)

$$\begin{aligned} A_n(z) &< B_n(z) && \text{for } z < 0 \\ A_n(z) &= B_n(z) && \text{for } z = 0 \\ A_n(z) &> B_n(z) && \text{for } z > 0 \end{aligned}$$

and thus in general (for n = -1, 0, 1, 2, ...)

$$\begin{aligned} A_n(z+nh) &< B_n(z+nh) && \textit{for } z < -nh && \text{(G7a)} \\ A_n(z+nh) &= B_n(z+nh) && \textit{for } z = -nh && \text{(G7b)} \\ A_n(z+nh) &> B_n(z+nh) && \textit{for } z > -nh && \text{(G7c)} \end{aligned}$$

In the case of **single reflection**, only n = –1 and n = 0 are considered. Then, we can distinguish three situations:

1. The focal point is below surface 2, hence

$$\begin{aligned} z &< 0 \\ z - h &< 0 \end{aligned}$$

Therefore

$$\begin{aligned} A_n(z) &< B_n(z) \\ A_n(z\text{-}h) &< B_n(z\text{-}h) \end{aligned}$$

and the resultant focus error signal (eq. G3) is negative, since the denominator is always positive (the signals A_n and B_n themselves are always positive). When the servo controller is active to focus the objective lens, it will move upward.

2. The focal point is above surface 1, hence

$$z > 0$$
$$z - h > 0$$

Therefore

$$A_n(z) > B_n(z)$$
$$A_n(z\text{-}h) > B_n(z\text{-}h)$$

and the resultant focus error signal (eq. G3) is positive. When the servo controller is active the objective lens will move downwards.

3. The focal point is between both surfaces, hence

$$z > 0$$
$$z - h < 0$$

Therefore

$$A_n(z) > B_n(z)$$
$$A_n(z\text{-}h) < B_n(z\text{-}h)$$

The sign of the resultant focus error signal (eq. G3) can now be both positive, zero or negative, depending on the position z, the gap height h and the reflectances $\mathcal{R}_1$ and $\mathcal{R}_2$.

We can conclude now, that the focus error signal can only be zero, when the focal point is somewhere between the two surfaces in the case of single reflection. As a consequence, the servo controller will position the focal point of the objective lens between the surfaces.

Consider now the situation, that surface 1 is stationary, while surface 2 moves upward, towards surface 1, i.e. the gap height h decreases. The servo controller is active and the objective lens will move upward, otherwise, surface 2 could come above the focal point with the result, that the focus error signal is not zero anymore. If we now assume that the lens displacement is equal to the displacement of surface 2, the focal point can come above surface 1 and again, the focus error signal is not zero. We can conclude therefore that the lens displacement will be smaller than the vertical displacement of surface 2 (i.e. the elastomeric surface in fig. G1) and the measured height variation of the surface is therefore expected to be smaller than the actual one (see also section G2.1 of this appendix).

The case of **multiple reflection** is more complex. All image surfaces are positioned at the same side, under surface 2. They therefore have the same, positive, contribution to the focus error signal as surface 2, when the focal point is between both surfaces, and this yields perhaps a zero focus error signal, not for a focal point position between the surfaces, but for a focal point position below surface 2. However, the effect will be very small, since the irradiance, reaching the photodiodes from an image surface, is very small due to light loss at the subsequent reflections, as will be discussed now.

Consider first that $\mathcal{R}_1$ is larger than 0.5. Then, most light is reflected on surface 1 and only a small part reaches surface 2. This is illustrated in eq. (G6) with the factor $(1-\mathcal{R}_1)^2$ for the summation, which is low for high values of $\mathcal{R}_1$. The contribution of the reflection on surface 2 and on the image surfaces is thus low compared with the contribution of the reflection on surface 1 and the focus error signal will therefore be zero at a position closer to surface 1 than to surface 2.

Otherwise, the contribution of the multiple reflections diminishes rapidly for larger values of n, when $\mathcal{R}_1$ is small, as indicated by the series terms in eq. (G6) (note that $\mathcal{R}_2$ is smaller than 1). For the situation of an air gap between a glass plate and surface 2, $\mathcal{R}_1$ is about 0.04. The factor $(\mathcal{R}_1\mathcal{R}_2)^n$ then decreases very rapidly to zero for increased n, while the summation of this factor converges to (see e.g. Spiegel, 1968 p. 107 eq. 19.7)

$$\sum_{n=0}^{\infty} (\mathcal{R}_1\mathcal{R}_2)^n = \frac{1}{1 - \mathcal{R}_1\mathcal{R}_2}$$

The fact that the factor $(\mathcal{R}_1\mathcal{R}_2)^n$ decreases very rapidly to zero and the fact that the summation converges mean, that the contribution of the reflection on the image surfaces is very low. Therefore, the influence of multiple reflection is expected to be negligible.

G1.3 The shape of the focus error signal for some values of the gap height

In this section graphs of the focus error signal are shown for some values of the gap height *h*. These graphs have been derived from measured photodiode signals *A* and *B*, as e.g. shown in section 3.1.4.1 (fig. 3.8 page 33), using the equation for single reflection (eq. G3). The values of the reflectances were chosen as follows

$$\mathcal{R}_2 = \frac{\mathcal{R}_1}{(1 - \mathcal{R}_1)^2}$$

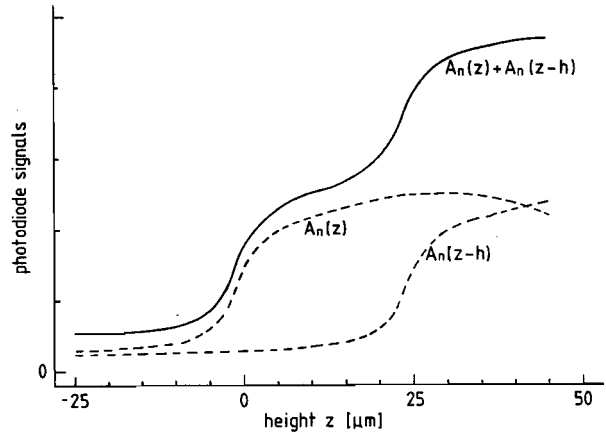
with the result that both surfaces have an equal contribution to the focus error signal. Eq. (G3) reads then

$$fes = \frac{[A_n(z\text{-}h) + A_n(z)] - [B_n(z) + B_n(z\text{-}h)]}{[A_n(z\text{-}h) + A_n(z)] + [B_n(z) + B_n(z\text{-}h)]}$$

The construction of the total photodiode signals $[A_n(z\text{-}h) + A_n(z)]$ and $[B_n(z) + B_n(z\text{-}h)]$ is illustrated in fig. G4.

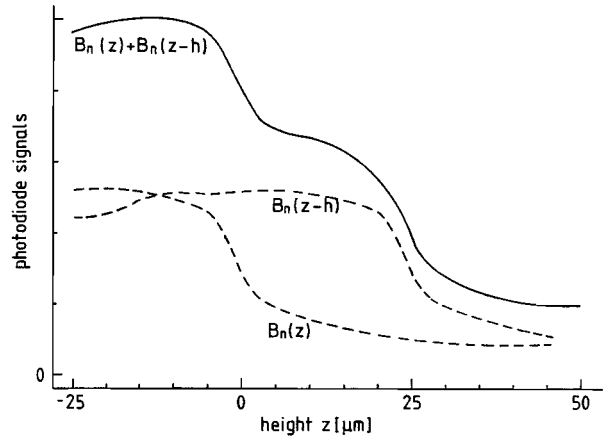
Figure G4
Theoretical construction of the photodiode signals for the case of two reflecting surfaces with equal contribution to the total signal.

a. photodiode signal A_n



b. photodiode signal B_n

($A_n(z)$ and $B_n(z)$ are the signals when reflection only occurs on the lower surface;
$A_n(z\text{-}h)$ and $B_n(z\text{-}h)$ are the signals when reflection only occurs on the upper surface;
Gap height h is 25 µm)



Some results are given in fig. G5, while fig. G6 shows some other results with addition of measured curves for the experimental verification.

The experiments were performed with the set up described in appendix C1 (see fig. C1 and C3). The test surface was a microscope glass with an index of refraction n_2 = 1.511, while the "window" was a glass plate of 1.2 mm thick and an index of refraction n_1 = 1.522. The gap between the test surface was filled with air, having an index of refraction n_a = 1. The reflectances (approximated by the expression for normal incidence) are thus

$$\mathcal{R}_1 \approx \left(\frac{n_1 - n_a}{n_1 + n_a} \right)^2 = 0.0414$$

and

$$\mathcal{R}_2 \approx \left(\frac{n_2 - n_a}{n_2 + n_a} \right)^2 = 0.0428$$

Therefore

$$\frac{\mathcal{R}_1}{(1 - \mathcal{R}_1)^2} = 0.0451 = 1.05\ \mathcal{R}_2$$

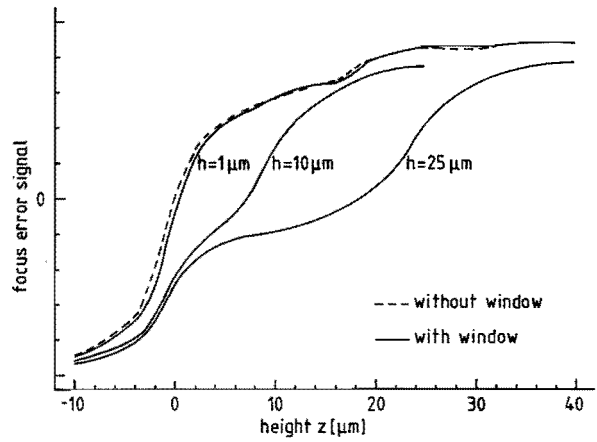
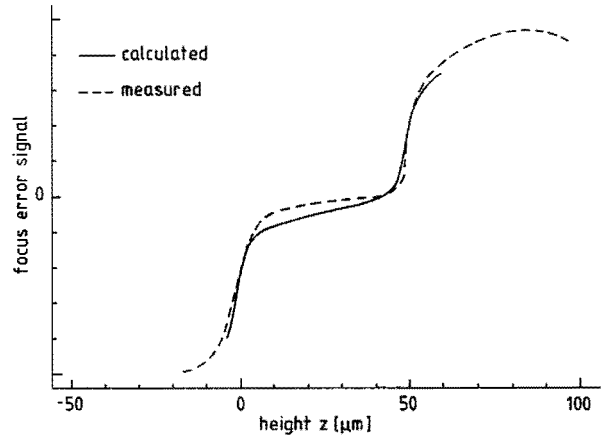


Figure G5[2]
The focus error signal, derived theoretically for different values of the gap height h between two reflecting surfaces at equal contribution of both surfaces:
$\mathcal{R}_2 = \mathcal{R}_1(1\text{-}\mathcal{R}_1)^{-2}$

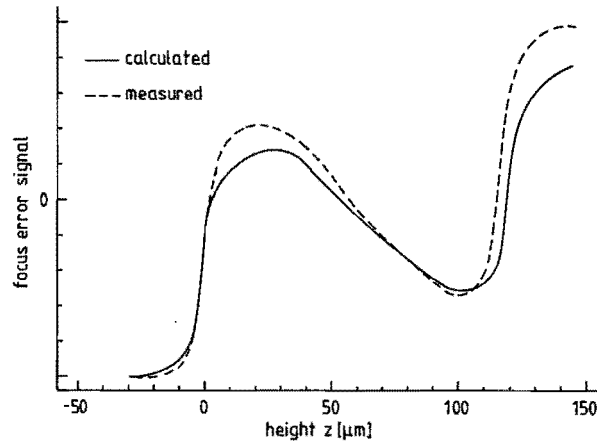
[2] The focus error signal in this and in the next figure is normalized to the maximum value. This was done to enable comparison of the calculated curves with the measured curves, since the calculations do not account for the electronic amplifications. The missing of the actual values of the signal is, however, not essential, since we are only interested in the shape of the curves.

Figure G6
Calculated and measured focus error signal for different values of the gap height h between two reflecting surfaces at (nearly) equal contribution of both surfaces.

a. h = 50 μm



b. h = 120 μm



and this differs only 5 percent from the ratio on which the calculations are based.

One of the reasons to choose these small reflectances was, that the influence of multiple reflection is then negligible. This is illustrated by the summation in eq. (G6)

$$\sum_{n=0}^{\infty} \mathcal{R}_1^n \, \mathcal{R}_2^n \, \left[A_n(z+nh) - B_n(z+nh)\right]$$

in which $\left(\mathcal{R}_1^n\ \mathcal{R}_2^n\right)$ is: 1 for $n = 0$;
0.00177 for $n = 1$;
and: $3.14 \cdot 10^{-6}$ for $n = 2$.

First we can conclude from fig. G6, that the differences between the calculated and the measured curves are not large. Especially the correspondence in the shape is good and we may therefore conclude, that the calculations describe the influence of the lower window surface reflection on the focus error signal at least qualitatively well. Further we can see, that the focus error signal is not essentially altered for a gap height h up to about 1 μm. The curve has only made a small translation, since the point where the focus error signal is zero is now not on the test surface, but between the test and the window surface.

When the gap height increases, the shape of the curve is more and more deformed with the result that the slope of the curve, and thus the sensitivity of the measurement, is increasingly diminished. Further increase of the gap height finally yields a curve in which the two individual surfaces are clearly distinguished (fig. G6b).

G2 The objective lens response on a film thickness variation in the closed loop mode

In the closed loop mode, the servo controller will move the objective lens to a position where the focus error signal is zero. In this section, we will derive an expression for this position and for the response of the lens on a vertical displacement of the surfaces. It will be assumed that both surfaces are close enough to the focal point, that the signals A_n and B_n may be considered proportional with the height distance between the focal point and the surface:

$$A_n(z) = a_0 + a_1 z \qquad \text{(G8a)}$$

$$B_n(z) = b_0 + b_1 z \qquad \text{(G8b)}$$

in which

$$b_0 = a_0$$

$$b_1 = -a_1$$

because of symmetry (see fig. 3.8 page 33).

The position of the lens (and of its focal point) will now be derived for both single and multiple reflection, using eq. (G3) and (G5) respectively. The

position is characterized by *fes* = 0 and only the numerator need to be evaluated, since the denominator is always positive.

G2.1 Single reflection approximation

The position of the focal point of the objective lens

Using eq. (G8), we can derive for the numerator of eq. (G3)

$$\begin{aligned} &\mathcal{R}_1\left[a_0 + a_1(z-h) - a_0 + a_1(z-h)\right] + \\ &\quad + (1-\mathcal{R}_1)^2\mathcal{R}_2\left[a_0 + a_1 z - a_0 + a_1 z\right] = \\ &= 2\mathcal{R}_1 a_1 z - 2\mathcal{R}_1 a_1 h + 2(1-\mathcal{R}_1)^2\mathcal{R}_2 a_1 z \end{aligned}$$

and taking the numerator (and thus the focus error signal) zero, we find for the in focus height position z_0 of the focal point

$$z_0 = \frac{\mathcal{R}_1}{\mathcal{R}_1 + (1-\mathcal{R}_1)^2\mathcal{R}_2} h \qquad \text{(G9)}$$

which is always smaller than h, since $\mathcal{R}_1$ and $\mathcal{R}_2$ are positive and the denominator is thus larger than the nominator.

We will now discuss the extreme situations, in which one surface is 100 percent reflecting or in which it is not reflecting at all to verify eq. (G9):

1. $\mathcal{R}_1 = 0 \Rightarrow z_0 = 0$
   No light is reflected on surface 1 and the focal point is on surface 2. The lens position is thus fully determined by the position of surface 2.
2. $\mathcal{R}_1 = 1 \Rightarrow z_0 = h$
   According to eq. (G9), the term $(1-\mathcal{R}_1)^2\mathcal{R}_2$ is zero, which is right since no light reaches surface 2. The position of the objective lens is thus fully determined by the position of surface 1 (the focal point is on surface 1), without influence of surface 2.
3. $\mathcal{R}_2 = 0 \Rightarrow z_0 = h$
   Surface 2 is not reflecting. The focal point is thus on surface 1, which determines the lens position.

4. $\mathcal{R}_2 = 1 \Rightarrow z_0 = \dfrac{\mathcal{R}_1}{\mathcal{R}_1 + (1-\mathcal{R}_1)^2} h \Rightarrow 0 < z_0 < h$

The lens position is influenced by the reflection on both surfaces, provided that ($\mathcal{R}_1 \neq 1 \wedge \mathcal{R}_1 \neq 0$), and the focal point position is therefore between both surfaces (However, this will appear to be different because of multiple reflection, see section G2.2 below).

These results are as on would expect and we may therefore conclude, that eq. (G9) describes the focal point position for zero focus error signal properly.

The objective lens response to a surface displacement

Consider now the situation, that **surface 2 is stationary** and **surface 1 moves** over a distance Δh (fig. G7). The situation after the displacement is then given by

$$z_0 + \Delta z = \frac{\mathcal{R}_1}{\mathcal{R}_1 + (1-\mathcal{R}_1)^2\mathcal{R}_2} (h + \Delta h)$$

and the resultant lens displacement is thus

$$\Delta z = \frac{\mathcal{R}_1}{\mathcal{R}_1 + (1-\mathcal{R}_1)^2\mathcal{R}_2} \Delta h \qquad \text{(G10)}$$

The focal point displacement Δz appears to be smaller than the surface displacement Δh, but its motion is in the same direction.

Otherwise, when **surface 1 is stationary** and **surface 2 moves** (see fig. G8), the lens position relative to the "fixed world" is given by $h - z$ and the lens displacement is thus given by

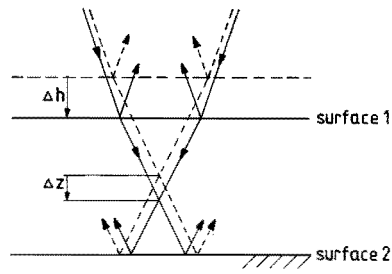
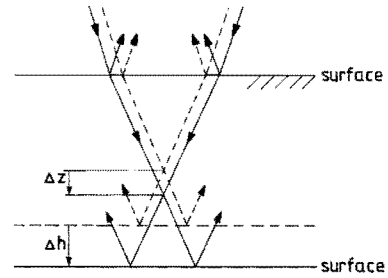


Figure G7
Displacement Δh of surface 1 at a stationary surface 2, and the consequent focal point displacement Δz.

Figure G8
Displacement Δh of surface 2 at a stationary surface 1, and the consequent focal point displacement Δh–Δz.



$$\Delta h - \Delta z = \left(1 - \frac{\mathcal{R}_1}{\mathcal{R}_1 + (1-\mathcal{R}_1)^2 \mathcal{R}_2} \right) \Delta h \qquad \text{(G11)}$$

which is also smaller than Δh and still in the same direction (i.e. the lens moves in the same direction as surface 2, but over a smaller distance).

G2.2 Multiple reflection analysis

The position of the focal point of the objective lens

We will now perform the same analysis as for the case of single reflection (section G2.1), but with an additional assumption necessary for the multiple reflection analysis, since there exist an (in principle infinite) number of image surfaces (fig. G3 page 171) which are outside the region where the signals A_n and B_n may be considered proportional with the distance to the focal point. Therefore we will assume, that $(\mathcal{R}_1\mathcal{R}_2)^n$ is so small for the larger values of n (at which the image surface is outside the proportional region of the signals A_n and B_n) that the contribution of these image surfaces to the focus error signal is negligible.

We can then write for the nominator of eq. (G6)

$$\mathcal{R}_1\left[a_0 + a_1(z-h) - a_0 + a_1(z-h)\right] +$$
$$+ (1-\mathcal{R}_1)^2\mathcal{R}_2 \sum_{n=0}^{\infty} \mathcal{R}_1^n\mathcal{R}_2^n\left[a_0+a_1(z+nh)-a_0+a_1(z+nh)\right] =$$

$$= 2\mathcal{R}_1 a_1 z - 2\mathcal{R}_1 a_1 h + 2a_1 z(1-\mathcal{R}_1)^2 \mathcal{R}_2 \sum_{n=0}^{\infty} \mathcal{R}_1^n \mathcal{R}_2^n +$$

$$+ 2a_1 h(1-\mathcal{R}_1)^2 \mathcal{R}_2 \sum_{n=0}^{\infty} n\mathcal{R}_1^n \mathcal{R}_2^n$$

Using

$$\sum_{n=0}^{\infty} (\mathcal{R}_1\mathcal{R}_2)^n = \frac{1}{1 - \mathcal{R}_1\mathcal{R}_2}$$

and

$$\sum_{n=0}^{\infty} n(\mathcal{R}_1\mathcal{R}_2)^n = \frac{\mathcal{R}_1\mathcal{R}_2}{(1 - \mathcal{R}_1\mathcal{R}_2)^2}$$

(see e.g. Spiegel, 1968 p. 107 eq. 19.7), we find for the nominator of eq. (G6)

$$2\mathcal{R}_1 a_1 z - 2\mathcal{R}_1 a_1 h + 2a_1 z \frac{(1-\mathcal{R}_1)^2 \mathcal{R}_2}{1 - \mathcal{R}_1\mathcal{R}_2} +$$

$$+ 2a_1 h \frac{(1-\mathcal{R}_1)^2 \mathcal{R}_1 \mathcal{R}_2^2}{(1 - \mathcal{R}_1\mathcal{R}_2)^2}$$

and the in focus position (at which the focus error signal is zero) is given by

$$z_0 = \frac{\mathcal{R}_1 - \dfrac{(1-\mathcal{R}_1)^2 \mathcal{R}_1 \mathcal{R}_2^2}{(1 - \mathcal{R}_1\mathcal{R}_2)^2}}{\mathcal{R}_1 + \dfrac{(1-\mathcal{R}_1)^2 \mathcal{R}_2}{1 - \mathcal{R}_1\mathcal{R}_2}} h \qquad \text{(G12)}$$

which is smaller than h, as proved by the following considerations:

1. Both $\mathcal{R}_1$ and $\mathcal{R}_2$ are positive, but smaller than 1. Therefore

$$0 < \mathcal{R}_1\mathcal{R}_2 < \mathcal{R}_1 < 1 \Rightarrow$$

$$\Rightarrow 0 < (1-\mathcal{R}_1) < (1-\mathcal{R}_1\mathcal{R}_2) < 1 \Rightarrow$$

$$\Rightarrow \quad 0 \; < \; \frac{(1-\mathcal{R}_1)^2 \mathcal{R}_1 \mathcal{R}_2^2}{(1 - \mathcal{R}_1\mathcal{R}_2)^2} \; < \; \mathcal{R}_1$$

The nominator of eq. (G12) is thus positive and smaller than $\mathcal{R}_1$.

2. For the same reason as under 1, we can write

$$\frac{(1-\mathcal{R}_1)^2 \mathcal{R}_2}{1 - \mathcal{R}_1\mathcal{R}_2} \; > \; 0$$

and the denominator of eq. (G12) is thus positive and larger than $\mathcal{R}_1$.

3. Combining 1 and 2, we find that the denominator of eq. (G12) is larger than the nominator and thus

$$z_0 \; < \; h$$

Again, we will evaluate the extreme situations in which one surface is 100 percent reflecting or in which it is not reflecting at all:

1. $\mathcal{R}_1 = 0 \Rightarrow z = 0$
2. $\mathcal{R}_1 = 1 \Rightarrow z = h$
3. $\mathcal{R}_2 = 0 \Rightarrow z = h$
4. $\mathcal{R}_2 = 1 \Rightarrow z = 0$

Situation 1 to 3 appears to be equal as derived for single reflection. This is obvious, since there is in fact no multiple reflection in these situations. In situation 1 and 3, the reflectance on one surface is zero and reflection only occurs on the other surface, while in situation 2, no reflection occurs on surface 2 because no light reaches that surface.

Situation 4 is now different. In the single reflection approximation, the focal point position would be somewhere between both surfaces. The multiple reflection, however, introduces reflecting image surfaces below surface 2, which "pull" the focal point towards surface 2.

The objective lens response to a surface displacement

Consider again the situation, that **surface 2 is stationary** and **surface 1 moves** over a distance Δh (fig. G7). Then we find for the situation after the displacement, using the equation derived for multiple reflection (eq. G12)

$$z_0 + \Delta z = \frac{\mathcal{R}_1 - \dfrac{(1-\mathcal{R}_1)^2\mathcal{R}_1\mathcal{R}_2^2}{(1 - \mathcal{R}_1\mathcal{R}_2)^2}}{\mathcal{R}_1 + \dfrac{(1-\mathcal{R}_1)^2\mathcal{R}_2}{1 - \mathcal{R}_1\mathcal{R}_2}} (h + \Delta h)$$

and the resultant lens displacement is thus

$$\Delta z = \frac{\mathcal{R}_1 - \dfrac{(1-\mathcal{R}_1)^2\mathcal{R}_1\mathcal{R}_2^2}{(1 - \mathcal{R}_1\mathcal{R}_2)^2}}{\mathcal{R}_1 + \dfrac{(1-\mathcal{R}_1)^2\mathcal{R}_2}{1 - \mathcal{R}_1\mathcal{R}_2}} \Delta h \qquad \text{(G13)}$$

which is smaller than Δh. The lens displacement is thus smaller than the surface displacement, but also in the same direction.

When **surface 1 is stationary** and **surface 2 moves** (see fig. G8), the lens position relative to the "fixed world" is given by $h - z$ and the lens displacement is thus given by

$$\Delta h - \Delta z = \left(1 - \frac{\mathcal{R}_1 - \dfrac{(1-\mathcal{R}_1)^2\mathcal{R}_1\mathcal{R}_2^2}{(1 - \mathcal{R}_1\mathcal{R}_2)^2}}{\mathcal{R}_1 + \dfrac{(1-\mathcal{R}_1)^2\mathcal{R}_2}{1 - \mathcal{R}_1\mathcal{R}_2}} \right) \Delta h \qquad \text{(G14)}$$

which is also smaller than the surface displacement Δh and in the same direction.

G2.3 Profile measurement through a glass plate on the test surface

Now we will present some measurements to verify the expectation that the objective lens displacement will be smaller than the height variation in the gap between the test surface and the glass plate. These measurements were performed with both the glass sinus and the metallic sinus as test surface and with the duran glass plate as window.

Before presenting the results of these tests, we will first discuss which results can be expected. This expectation is based on the analysis presented in the former sections (G2.2 and G2.3), where the influence of the window surface reflection on the focus position, and thus on the objective lens response, was discussed.

Expected results of a roughness measurement with a glass plate laid on the test surface

As discussed in section 3.2.5 (page 43), the **upper window surface** (i.e. the surface at the side of the objective lens) has no influence on the focus error signal, since it is far away from the test surface and thus from the focal point (which is in the vicinity of the test surface). It is therefore expected, that it does not influence the surface roughness measurement.

Otherwise, the reflection on the **lower window surface** has influence on the focus error signal (section G2.1 and G2.2), because this window surface is very close to the test surface and thus to the focal point. As a result, the focus error signal is not zero for a focal point position on the test surface, but for a focal point position between the test surface and the lower window surface (section G1.2). The actual focal point position for zero focus error signal depends on the reflectances of both surfaces. When the gap height varies, e.g. during scanning the roughness of the test surface, the focal point position will remain between both surfaces, as e.g. discussed in section G1.2, and the measured height variation is smaller than the real height variation. How much smaller depends on the reflectances.

Now, we can formulate the expected results of the experiments more detailed:

- For the experiments with the glass plate ("window") on the **glass sinus** the lower window surface and the test surface have a small but almost equal reflectance (about 4 percent). It is therefore expected that the measured amplitude of the sinus profile is about half the real amplitude (see e.g. eq. (G13) and eq. (G14)).
- For the experiments with the glass plate on the **metallic sinus**, it is expected that the measured amplitude of the sinus profile is hardly affected by the reflection on the lower window surface, since the lower window surface reflectance is much smaller than the test surface reflectance (about 4 percent and 80 percent respectively).

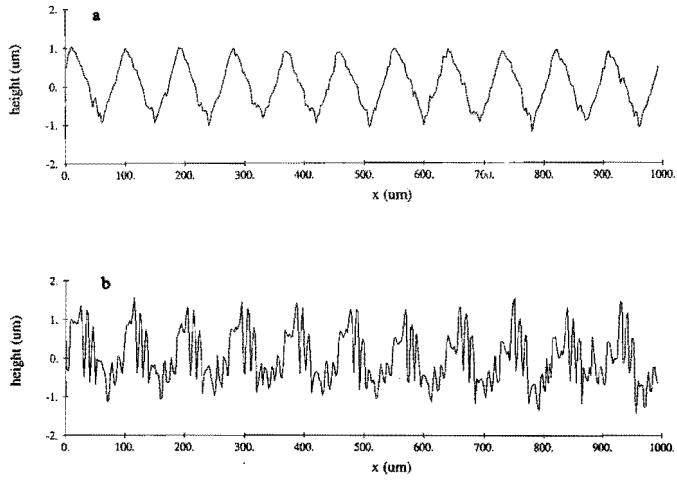
Measurements on the glass sinus profile

The results of the measurements on the glass sinus profile, both with and without glass plate on it, are given in fig. G9. It shows that the measurement is largely disturbed by the addition of the glass plate, when the plate is laid directly on the surface (fig. G9b): The shape is strongly deformed and the measured amplitude is much too large. This result can be caused by the possible lack of reflection on the sinus summits, where both surfaces are contacting.

Figure G9
Surface roughness measurements on the glass sinus profile with the double wedge system of Struik and Chang (1987).

a. Without glass plate on the surface

b. With glass plate on the surface



Measurements performed on the metallic sinus profile

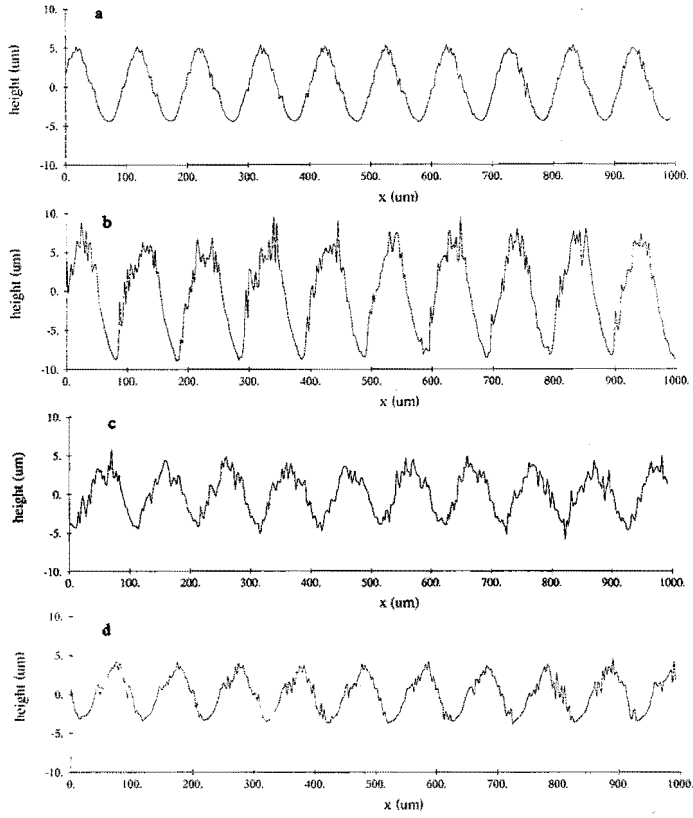
The results of the measurements on the metallic sinus profile are presented in fig. G10. Again, the amplitude of the measurements with the glass plate on the sinus was larger than expected and the shape of the curve is deformed a bit.

Two additional experiments, in which the lower window reflection was eliminated (or at least strongly reduced), were performed. In one experiment, the glass plate was replaced by a liquid (1.4 mm deep water[3]) and the result is

[3] In the compact disc configuration, the disc surface is scanned through the 1.2 mm thick polycarbonate (index of refraction is 1.56) layer, yielding $W_{40} = 57(NA)^4$ μm for the spherical aberration (see eq. F.10b page 166). The objective lens is corrected for this amount of spherical aberration (i.e. the lens has the same spherical aberration, but with opposite sign) and the water (index of refraction 1.33) on the test surface should thus yield the same amount of aberration, which is fulfilled for a depth of about 1.4 mm.

Figure G10
Surface roughness measurements on the metallic sinus profile with the double wedge system of Struik and Chang (1987).

a. Without glass plate on the surface
b. With glass plate on the surface
c. With water on the surface
d. With duran glass on the surface and oil mixture with equal index of refraction between the surfaces



shown in fig. G10c. Now, the measured amplitude is smaller than the real amplitude, which is caused by the light refraction on the water to air interface which must be accounted for (the origin of the reduced amplitude in the measurement is thus different from the origin in the case of a reflecting lower window surface). This means, that the measured amplitude must be multiplied by the index of refraction of water (1.33) and, in terms of roughness values, we find an R_a value (or: Centre Line Average) of 2.236 μm × 1.33 = 2.974 μm, which approximates the measurement of fig. G10a (R_a = 2.927 μm) very well.

The other experiment was use of a combination of a liquid and a glass with a (nearly) equal index of refraction, which would eliminate the reflection on the liquid to glass interface. A proper combination appears to be duran glass and an oil mixture of 75 percent Shell Ondina 15 and 25 percent Shell Ondina 68. The measured indices of refraction are given in appendix D1. The result of this measurement is given in fig. G10d, which shows that the shape of the profile is

reproduced rather accurate, while the measured amplitude is again smaller than the real amplitude. After correction for the index of refraction (1.47), we find an R_a value of 3.030 μm.

Discussion

As we have seen in fig. G9 and G10a/b, the glass plate on the surface appears to deform the shape of the profile and, moreover, enlarges the measured amplitude of the sinus, while it was expected from the influence of the lower window surface reflection on the focus error signal (see section G2.2) that the amplitude would be measured too small.

This unexpected increase in the measured amplitude is clearly not caused by the reflection on the upper window surface, since the measurements, shown in fig. G10c/d, where the reflecting upper surface is still present, but the lower not, yielded much better results with accurate measurement of the surface height variations. This means, that the reflection on the upper window surface has no significant influence on the measurements, as was already expected from the initial signal measurements (see fig. 3.15, page 43, and the start of section 3.2.5, page 42).

The only significant difference between the measurement of fig. G10b and the measurement of fig. G10d is the presence of reflection on the lower window surface. Therefore, the amplitude increase in the measurements must be caused by an effect which we did not account for in our analysis in section G2.1 and G2.2, but which originates from the reflection on the lower window surface. Such an effect is perhaps interference of the reflected rays in the plane of the photodiodes, with the possible result of a very low irradiance on the photodiodes at some values of the gap height. This may yield an unreliable response of the objective lens.

It is not understood, how interference actually influence the focus error signal and the response of the objective lens. More research on this matter is thus needed. In this thesis, however, we take it as a (temporary) conclusion, that a liquid is needed in the contact, which has an equal index of refraction as the window. Then reflection on the window to lubricant interface is avoided and the roughness profile is measured properly in the closed loop mode.

APPENDIX H THE NOISE OF THE FOCUS ERROR DEVICES

The noise of focus error devices is discussed by Bouwhuis et al. (1987 section 2.4.4 pp. 62ff.) and by Claesen (1992). Considering the film thickness measurements, the most important noise sources are:

- **Dark current noise** of the photodiodes;
- **photon shot noise;**
- **noise of the electronic amplifier;**
- **noise of the laser.**

The amplitude variations in the focus error signal, due the total noise, is about 10 mV, according to Claesen (1992, section 4.6.1.1). Consequently the uncertainty in the film thickness measurement would roughly be 1 μm, since the gradient d(*ffs*)/dz of the focus error curve is about 10 mV·μm^{-1} (see eq. (4.2) page 61). Reduction of the noise is therefore necessary and the possibilities will be discussed here, considering the four noise sources mentioned above.

Dark current noise

The photodiodes yield a varying signal, even when they receive no light. This signal (varying in time) is superimposed on the "effective" signal which is caused by the irradiance. In general the noise is small compared with the effective signal (signal to noise ratio e.g. 10^5) and thus negligible. In our application, however, the irradiance on the photodiodes will be rather small because of the low reflectance on the elastomer to lubricant interface (see appendix D2). Therefore the signal to noise ratio can perhaps be too small to obtain the required accuracy.

If necessary, the laser power should be increased to keep the signal to noise ratio at a reasonable level. Also, it can be considered to apply a metallic coating on the elastomer's running face. However, such a coating should be very thin (some nm) to prevent influence on the roughness texture, while it must otherwise be homogeneous to realize a constant reflectance over the surface. Finally, possible wear must be prevented.

Photon shot noise

Photon shot noise is caused by the quantum nature of light. The irradiance on the photodiodes is not constant in time, but varies because the photons reach the diodes with different time intervals. As a result, the photodiode signals are not constant, but show a ripple which is regarded as noise.

The signal to noise ratio now also depends on the total irradiance on the photodiodes. Therefore, the same measures as discussed above for the dark current noise can be considered, if the low reflectance causes a too large noise level.

Noise of the electronic amplifier

The present electronic device for the signal processing was initially developed for the displacement measurements (see section 3.1.2 page 26) and is also used for the shape and roughness measurements. There are some factors which reduce the accuracy in general (i.e. not for the film thickness measurements in particular):

- The photodiode signals are not amplified before they are transmitted to the "fixed world" through a cable;
- the laser current and the photodiode signals are transmitted through the same cable.

These two facts have a negative influence on the signal to noise ratio and thus on the total accuracy.

Besides, the present amplifier has a bandwidth of about 1 kHz, which is not sufficient for the film thickness measurements which require reliable measurements at frequencies up to 1 MHz.

A new design of the electronic amplifier is thus necessary to enable reliable film thickness measurements.

Noise of the laser

The most important noise source of the diode laser is "mode hopping", i.e. the mode of the laser (and consequently the wavelength of the emitted light) changes with a discrete value at some temperature changes. Such mode hops can occur already at temperature changes of e.g. some Kelvin as indicated by Claesen (1992 section 5.2.2). Also, more than one mode can exist at certain temperatures with the possibility of continuous mode hopping even at constant temperature.

A negative consequence of mode hopping is, that the focal point of the objective lens changes due to chromatic aberration. Whether the error in the measurement caused by this effect is negligible or not depends on among others the type of the lenses used in the transducer, because two different lenses can have a totally different chromatic aberration, depending on among others the dispersion of the used glass.

If the influence of mode hopping is too large, the laser temperature can be controlled using a so-called "peltier element".

APPENDIX I THE PRESSURE AND TEMPERATURE INFLUENCE ON THE FILM THICKNESS MEASUREMENT USING FOCUS ERROR DETECTION

The film thickness measurement is influenced by the (contact) pressure and by the temperature, since the parameters important for the measurements (like the index of refraction and the window thickness) are pressure and temperature dependent.

Fig. I1 illustrates the essence of the film thickness measurement which is performed through a transparent window with thickness t, using an objective lens (fixed in the rigid body, see section 3.2.2 page 40) with focal distance f.

The focus spot F' is in principle at a constant distance h_1 from the window surface[1] since the lens is fixed relative to the window. The distance h_2 between the spot F' and the elastomeric surface is determined from measurement of the focus error signal during running of the elastomeric specimen over the window. The film thickness h is given by

$$h = h_1 + h_2 \tag{I1}$$

Now we will derive how the (contact) pressure and the temperature will influence both the height h_1 of the focus spot F' above the surface and the focus error signal from which h_2 is derived.

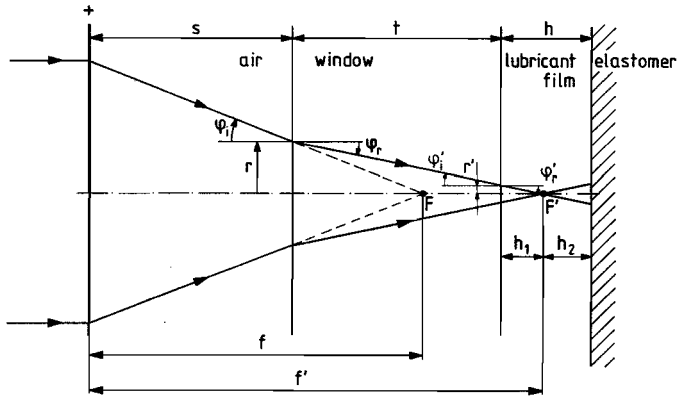


Figure I1
Film thickness measurement with fixed objective lens:

h_1 is in principle constant;

h_2 is derived from measurement of the focus error signal.

[1] The value of h_1 can be determined (after mounting the transducer in the rigid body) by measuring the focus error signal when the window surface is clean (i.e. without the presence of the elastomer and lubricant on the surface).

First it will be shown how h_1 depends on the geometrical parameters (s, t and f) and on the indices of refraction of the window (n_w) and of the lubricant (n_l). A change in the value of these parameters (e.g. due to the contact pressure and change in temperature) yield a change in the value of h_1 and may therefore cause an error in the measurement of the film thickness h.

Next it will be derived that accurate knowledge of the index of refraction of the lubricant (n_l) and of the elastomer (n_e) is important to derive h_2 accurately from the measured focus error signal. Pressure and temperature influence on these indices of refraction may therefore cause inaccurate measurement of h_2 and thus of the film thickness h.

In the final two sections the pressure and the temperature influence on all these parameters will be discussed and the consequent error in the film thickness measurement will be estimated.

I1 The distance between the focal point and the window surface

Consider fig. I1 again. We can write for the distance h_1 between the focal point F' and the window surface

$$h_1 = f' - s - t$$

while

$$f' = s + t + \frac{r'}{\tan \varphi_r'}$$

in which

$$r' = r - t \tan \varphi_r$$

Also

$$r = (f - s) \tan \varphi_i$$

Using Snell's law

$$n_a \sin \varphi_i = n_w \sin \varphi_r$$
$$n_w \sin \varphi_i' = n_l \sin \varphi_r'$$

in which

n_a = the index of refraction of air = 1 [-]
n_w = the index of refraction of the window [-]
n_l = the index of refraction of the lubricant [-]

estimating for small angle of incidence (i.e. $\varphi_i \ll 1$)

$$\begin{aligned} \sin \varphi_i &\approx \tan \varphi_i \approx \varphi_i \\ \sin \varphi_r &\approx \tan \varphi_r \approx \varphi_r \\ \sin \varphi_i' &\approx \tan \varphi_i' \approx \varphi_i' \\ \sin \varphi_r' &\approx \tan \varphi_r' \approx \varphi_r' \end{aligned}$$

and considering that

$$\varphi_i' = \varphi_r$$

we find

$$f \approx s + t + n_l \left(f - s - \frac{1}{n_w} t \right) \tag{I2}$$

and finally

$$h_l \approx n_l \left(f - s - \frac{1}{n_w} t \right) \tag{I3}$$

This equation shows that the pressure and the temperature influence on the geometrical parameters and on the indices of refraction of the window and the lubricant will lead to a shift in the focus spot height relative to the contacting window surface indeed and thus to an error in the film thickness measurement when we do not account for this height shift.

I2 The relation between the film thickness and the focus error signal

Deriving the value of h_2 from the measured focus error signal we must account for the index of refraction of the lubricant and of the elastomer. Two effects are determined by these indices of refraction:

- The reflectance on the lubricant to elastomer interface, which is needed for the measurement;
- the refraction on the air to window interface (the window surface on the left hand side in fig. I1), which virtually shifts the elastomeric surface to a different position than the real position.

Both effects will now be discussed.

The reflection on the lubricant to elastomer interface

The reflectance on the lubricant to elastomer interface has influence on the focus error signal. This reflectance depends on the indices of refraction of the lubricant and the elastomer (see eq. (D1) page 148) and pressure and temperature influence on these indices of refraction might thus influence the reflectance.

Assuming that the relative change in the index of refraction, due to the contact pressure and temperature, is almost equal for both the lubricant and the elastomer (especially at the rather low pressures and temperatures considered in this study, see point 8 and 9 on page 19 in section 2.1) the reflectance will hardly be influenced since the relative change in the nominator and in the denominator of eq. (D1) (page 148) is almost equal. Moreover the reflectance influence on the focus error signal is small (see section 3.1.4.2 page 33) and the measurement will thus hardly or not be influenced by the pressure and temperature influence on the reflectance. This aspect will therefore neglected.

The refraction on the air to window interface

Fig. I2 shows that the beam returning after reflection on the elastomeric surface apparently diverges from position F''' instead of position F" (which is the image of point F' at reflection on the elastomeric surface). Consequently the elastomeric surface is apparently at a distance h_2' from point F (the focal point when refraction on the air to window interface would not occur) and the focus error signal is therefore a measure for h_2' instead of h_2. However, h_2 can be derived from h_2' (and thus from the focus error signal) as shown below.

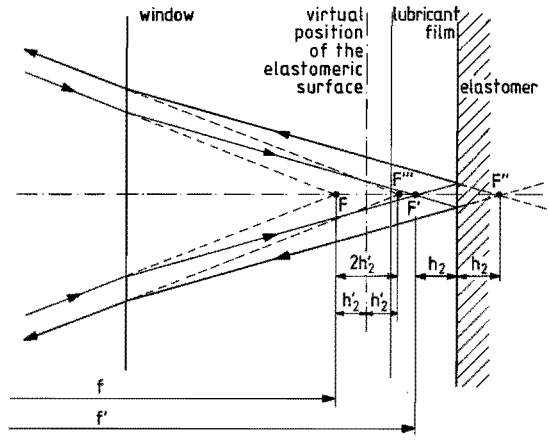


Figure I2
Virtual position of the elastomeric surface and apparent distance h_2' to the virtual focus F.

Analogous to eq. (I2) we can then derive that

$$f' + 2h_2 \approx s + t + n_1 \left((f + 2h_2') - s - \frac{1}{n_w} t \right)$$

and subtracting eq. (I2) (which still holds) we find

$$h_2' \approx \frac{h_2}{n_1} \tag{I4}$$

We can conclude now that the focus error signal is not a measure for h_2 but for $h_2' = h_2/n_1$. Therefore the pressure and temperature also influences the apparently measured value of h_2 (and thus the measured value of the film thickness) if we do not account for the change in the lubricant's index of refraction when we derive h_2 from the measured focus error signal.

I3 The contact pressure influence

The pressure influence on the film thickness measurement is caused by:
- The pressure dependence of the index of refraction;
- the bending and impression of the window due to the contact load.

These two effects will be considered first, followed by discussion of the pressure influence on the film thickness measurements in general. Finally the pressure influence will be estimated for the film thickness measurements presented in chapter 4 in particular.

I3.1 The pressure dependence of the index of refraction

As expressed by eq. (I3) and (I4) the pressure influence on the index of refraction affects both the value of h_1 and the measurement of h_2. This pressure influence is related to the change in density which is often expressed by the **Lorenz-Lorentz relation** (see e.g. Poulter et al., 1932), which reads

$$\frac{n^2 - 1}{n^2 + 2} \cdot \frac{1}{\rho} = c$$

in which

n = Index of refraction [-]

ρ	=	Density	[kg·m^{-3}]
c	=	Constant	[m^{3}kg^{-1}]

i.e.

$$n = \sqrt{\frac{1 + 2\rho c}{1 - \rho c}} \tag{15}$$

Estimation of the pressure influence on the index of refraction now only need an expression for the relation between the pressure and the density, both for the lubricant and for the glass.

Pressure influence on the index of refraction of the lubricant

When we want to estimate the pressure influence on the index of refraction of the lubricant, the pressure distribution in the lubricant film must be known. Considering the contact of elastomers and a rigid body (which is typically heavily loaded, i.e. an eventual lubricant film is very thin relative to the static deformations of the elastomer) we can consider the hydrodynamic pressures in the lubricant film to be equal to the static pressure in the dry contact (see e.g. Kanters 1990 section 4.2.1 pp. 51-52) which can be calculated using e.g. the finite element method. When the contacting bodies are cylindrical or ellipsoidal the Hertzian formulas can be used to estimate the contact pressure distribution.

Now we will discuss how we can determine the (in principle known) contact pressure influence on the index of refraction of the lubricant.

Hamrock and Dowson (1981 page 151 and 176) proposed the following relation between the pressure p and the density of the lubricant ρ_l

$$\frac{\rho_l}{\rho_{l0}} = 1 + \frac{c_1 p}{1 + c_2 p} \tag{16}$$

in which

ρ_l	=	Density of the lubricant	[kg·m^{-3}]
ρ_{l0}	=	ρ_l (at atmospheric pressure)	[kg·m^{-3}]
p	=	Pressure	[Pa]
c_1	=	Constant	[Pa^{-1}]
c_2	=	Constant	[Pa^{-1}]

and they found for mineral oil: $c_1 = 0.6$ GPa^{-1}
$c_2 = 1.7$ GPa^{-1}

Then we find for a maximum pressure of 50 MPa (see point 8 section 2.1 page 19) and for a density $\rho_{l0} = 0.900$ kg·dm^{-3} (which is a common value for most lubricants)

$$\frac{\rho_l}{\rho_{l0}} \leq 1.028$$

Considering the relation between the index of refraction and the density (eq. I5) Poulter found for

- paraffinic oil: $c = 0.325$ dm^{3}kg^{-1}
- glycerine: $c = 0.279$ dm^{3}kg^{-1}

Assuming $c = 0.3$ dm^{-3}kg we find for pressures up to 50 Mpa

$$\frac{n_l}{n_{l0}} \leq \frac{n_l(p=50\ \text{MPa})}{n_l(p=0\ \text{MPa})} = \frac{1.4672}{1.4524} = 1.01$$

The relation between the pressure and the density and the relation between the pressure and the index of refraction are shown in fig. I3 and fig. I4 respectively. For pressures up to 50 MPa the curves are approximately linear and therefore we can write

$$\frac{\rho_l}{\rho_{l0}} = 1 + c_3\, p$$

Figure I3
The relative pressure influence on the lubricant density.
$c_1 = 0.6$ GPa^{-1}
$c_2 = 1.7$ GPa^{-1}

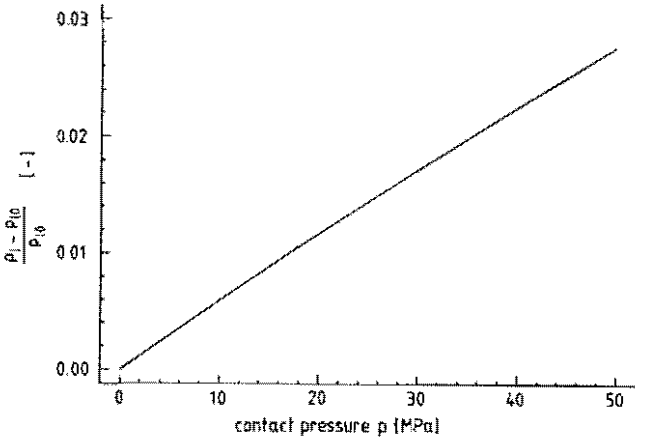
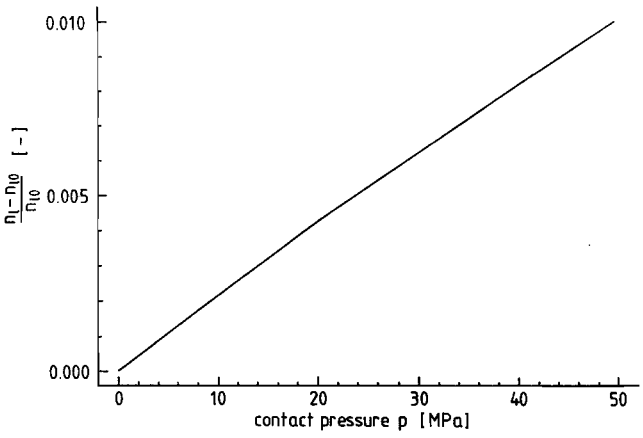


Figure I4
The relative pressure influence on the index of refraction of the lubricant.

c_1 = 0.6 GPa^{-1}
c_2 = 1.7 GPa^{-1}
c = 0.3 dm^{3}kg^{-1}
ρ_{10} = 0.9 kg·dm^{-3}



and

$$\frac{n_1}{n_{10}} \approx 1 + c_4\, p \qquad (I7)$$

in which

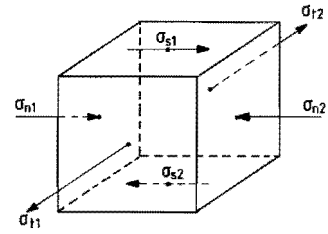
c_3 = Constant = 0.56 GPa^{-1}
c_4 = Constant = 0.2 GPa^{-1}

Pressure influence on the index of refraction of the window

The pressure influence on the index of refraction of the window is more complicated, mainly because the stresses in the window (caused by the contact pressure) are not uniform (i.e. the stress varies with the position in all directions). The complexity is illustrated by the following:

- The stress distribution in the window will not be uniform.
- Besides negative normal stresses σ_n (i.e. comparable to the pressures in the fluid), positive normal stresses (i.e. tensile stresses σ_t) and shear stresses σ_s are also present.
- One infinitesimal volume element of the window may experience all three stress situations simultaneous, e.g. compression in one direction, tensile stress in another direction and shear stresses at different sides of the volume element in different directions (see e.g. fig. I5).

Figure I5
Example of stresses acting on an infinitesimal volume element of the window.
σ_n = Negative normal stress
σ_t = Tensile stress
σ_n = Shear stress



Consequently, the estimation of the (non-uniform) distribution of the index of refraction is complicated and the fact that the influence of the contact pressures on the index of refraction of the window is not constant over the whole window makes it difficult to estimate the consequence of this pressure influence for the distance h_1 between the focal point and the window surface (see fig. I1 page 191).

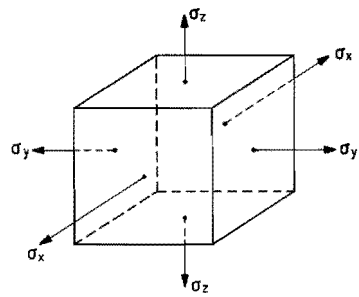
The pressure influence on the window's index of refraction will now be roughly estimated for the hydrostatic stress situation, i.e. only normal stresses act on the boundaries of the infinitesimal small volume element (see fig. I6), while the stresses on all the surfaces are equal. Therefore

$$\sigma_x = \sigma_y = \sigma_z = \sigma_h$$

in which

σ_x	=	Normal stress in X-direction	[Pa]
σ_y	=	Normal stress in Y-direction	[Pa]
σ_z	=	Normal stress in Z-direction	[Pa]
σ_h	=	Hydrostatic stress	[Pa]

Figure I6
Example of stresses acting on an infinitesimal volume element of the window.
σ_x, σ_y and σ_z are the normal stresses acting on the bound-aries of the volume element.



The strains are given by

$$\varepsilon_x = \frac{1}{E_w}\left[\sigma_x - \nu_w\,(\sigma_y + \sigma_z)\right]$$

$$\varepsilon_y = \frac{1}{E_w}\left[\sigma_y - \nu_w\,(\sigma_z + \sigma_x)\right]$$

$$\varepsilon_z = \frac{1}{E_w}\left[\sigma_z - \nu_w\,(\sigma_x + \sigma_y)\right]$$

in which

ε_x	=	Strain in X-direction	[-]
ε_y	=	Strain in Y-direction	[-]
ε_z	=	Strain in Z-direction	[-]
E_w	=	E-modulus of the window	[Pa]
ν_w	=	Poisson's ratio of the window	[-]

The hydrostatic stress situation yields

$$\varepsilon_x = \varepsilon_y = \varepsilon_z = \varepsilon = (1 - 2\nu_w)\,\frac{\sigma_h}{E_w}$$

The volume V is given by

$$V = V_0\,(1 + \varepsilon)^3$$

(V_0 is the volume at atmospheric pressure).

Considering that $\varepsilon \ll 1$ we can write

$$\frac{V}{V_0} \approx (1 + 3\varepsilon)$$

Then we can write for the window's density ρ_w

$$\frac{\rho_w}{\rho_{w0}} = \frac{V_0}{V} \approx \frac{1}{1 + 3\varepsilon} = \frac{1}{1 + 3\,(1 - 2\nu_w)\,\dfrac{\sigma_h}{E_w}}$$

and the pressure influence on the index of refraction can be estimated using this equation and the Lorenz-Lorentz relation (eq. (I5) page 196).

Now we assume that the maximum (hydrostatic) stress in the window will not be much larger than the contact pressure p, i.e. we estimate the hydrostatic stress by

$$\sigma_h \approx -p$$

The window in the rigid body will be made from Duran glass, which has the following properties

$$\begin{aligned} E_w &\approx 70 && \text{GPa} \\ \nu_w &\approx 0.2 \\ \rho_{w0} &\approx 3 && \text{kg·dm}^{-3} \\ n_w &\approx 1.45 \end{aligned}$$

Then we find for the constant c in the Lorenz-Lorentz relation

$$c \approx 0.09 \quad \text{dm}^3\text{kg}^{-1}$$

For a maximum pressure of 50 MPa the pressure influence on the window's index of refraction is

$$\frac{n_w}{n_{w0}} \leq \frac{n_w(p=50\ \text{MPa})}{n_w(p=0\ \text{MPa})} = \frac{1.4531}{1.4524} = 1.00046$$

The relation between the pressure and the density and index of refraction are shown in fig. I7 and I8 respectively. For pressures up to 50 MPa the curves are almost linear and therefore we can write

Figure I7
The relative pressure influence on the window's density.
$E_w = 70$ GPa^{-1}
$\nu_w = 0.2$

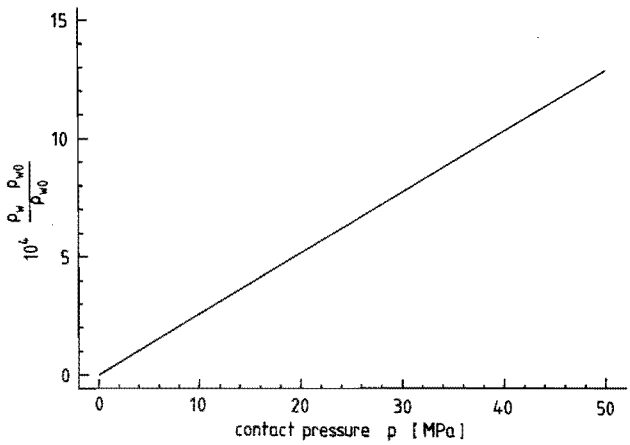
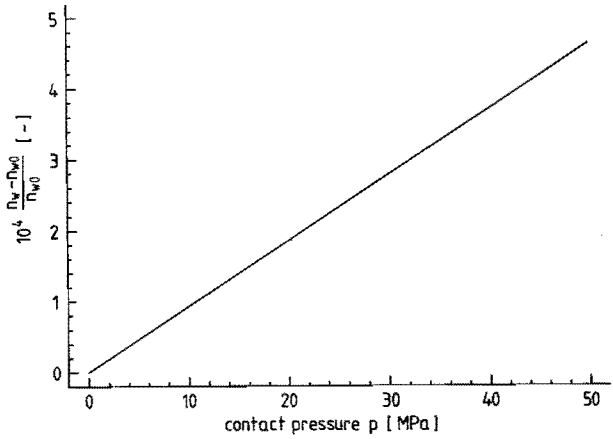


Figure I8
The relative pressure influence on the index of refraction of the window .

E_w = 70 GPa
ν_w = 0.2
c_5 = 0.026 GPa^{-1}
c_6 = 0.092 GPa^{-1}
ρ_{w0} = 3 $kg \cdot dm^{-3}$



$$\frac{\rho_w}{\rho_{w0}} \approx 1 + c_5 p$$

and

$$\frac{n_w}{n_{w0}} \approx 1 + c_6 p \qquad \text{(I8)}$$

in which

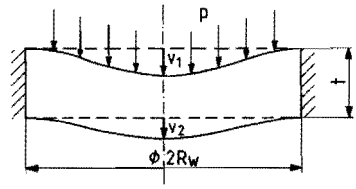
c_5 = Constant = 0.026 GPa^{-1}
c_6 = Constant = 0.0092 GPa^{-1}

I3.2 The bending and impression of the window

The contact load will influence the film thickness measurement by bending the window and by impressing its surface. As a result the distance between the contact spot and the contacting window surface is changed. Reduction of this bending and impression is therefore important to prevent too much influence on the measurement.

A preliminary finite element analysis of the window bending and impression was performed by Hazenberg (1992, section 4.2.3) He considered a simple axisymmetric situation with uniform contact pressure p, shown in fig. I9.

Figure I9
Axisymmetric configuration for the finite element analysis of the window's bending and impression.



The deformations were calculated for different values of the pressure p, the window's E-modulus E_w, the window thickness t and its radius R_w. In all calculations the Poisson's ratio of the glass ν_w was 0.3. (For glass the Poisson's ratio is somewhere in the range of 0.2 to 0.3. Its influence on the strains and on the displacements is then small, as e.g. expressed by eq. (I18) on page 215).

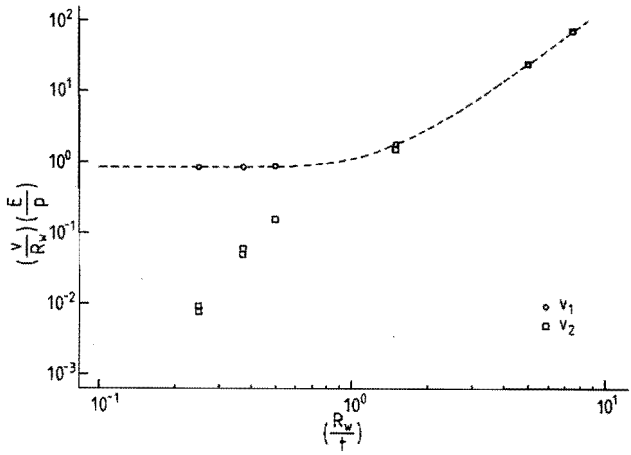
The resultant surface displacements v_1 and v_2 in the window centre are proportional to the contact pressure to E-modulus ration (p/E) and shown in dimensionless form in fig. I10. The effect of the window bending and impression is expressed in the pressure influence on the lens to window distance s and the window thickness t (see section I1 page 192) which read

$$s = s_0 - v_2$$
$$t = t_0 - v_1$$

in which

s_0 = s at unloaded contact [m]
t_0 = t at unloaded contact [m]

Figure I10
Dimensionless representation of the surface displacements v_1 and v_2 at the window centre.



The most important conclusions of Hazenberg's analysis are:

- v_1 and v_2 are equal for large values of the window radius to thickness ratio R_w/t. The window impression is then negligible compared with the bending and the dimensionless central displacement v_1/R is proportional to $(R_w/t)^3$;
- v_2 is much smaller than v_1 at small values of the window radius to thickness ratio R_w/t. Now the bending is negligible compared with the window impression and the problem can be solved using the so-called *elastic half space* approach as e.g. presented by Johnson (1985 chapter 3 pp. 45ff.);
- The stresses in the window body increase with increasing radius to thickness ratio R_w/t. According to Hazenberg (1992, app. 3) the stresses are too large for R_w/t larger than roughly 1.5 at contact pressures of 10 MPa.

I3.3 Discussion on the total pressure influence

In this section we will discuss the total contact pressure influence on the film thickness measurements by focus error detection. First recall the start of this appendix, where the total film thickness was split into two parts:

- The distance h_1 between the focal point and the contacting window surface which is in principle constant;
- The distance h_2 between the focal point and the elastomer's contacting surface, which is to be derived from the measured focus error signal.

In section I1 (page 192) we found

$$h_1 \approx n_1 \left(f - s - \frac{1}{n_w} t \right) \tag{I9}$$

and in section I2 (page 193)

$$h_2 \approx n_1 h_2' \tag{I10}$$

in which h_2' is directly derived from the focus error signal.

As already mentioned before the total contact pressure influence is determined by the pressure dependence of the indices of refraction of the lubricant and the window (n_1 and n_w respectively) and by the impression and bending of the window (this influences the values of s and t). The consequence of these effects for the film thickness measurement will now be discussed, using the results of section I3.1 and I3.2, to find out whether an effect has significant influence or whether it can be neglected.

The lubricant's index of refraction

In section I3.1 was derived that the lubricant's index of refraction n_l is increased by about 1 percent when the pressure is raised to 50 MPa. As expressed by eq. (I9) the value of h_1 is then also increased by 1 percent, yielding underestimation of h_1 by 1 percent if this pressure influence is not accounted for.

Eq. (I10) shows that the value of h_2 is also underestimated by 1 percent (at a contact pressure of 50 MPa), when this pressure influence is neglected.

Therefore we can finally conclude that *the total film thickness is then underestimated* by about 1 percent. This is just the limit of the tolerated uncertainty at a film thickness in the range of 1 to 10 μm (see section 2.2 requirement 4 page 20). Therefore the pressure influence on the lubricant's index of refraction may not be neglected at contact pressures of the order of 10 MPa.

The window's index of refraction

The pressure influence on the window's index of refraction n_w is very small: Roughly about 0.05 percent at a contact pressure of 50 MPa. As expressed by eq. (I9) and (I10) it only influences the value of h_1. Considering that the window thickness t will be of the order of 1 mm and that the indices of refraction of both the lubricant and of the window are approximately equal to 1.45, we find that h_1 is roughly increased by 0.5 μm at a pressure increase to 50 MPa. Negligence of this pressure influence yields thus an *underestimation in the film thickness* of the order of 0.1 μm.

This pressure influence on h_1 is far from negligible, since the maximum uncertainty in the measured film thickness must be limited to about 0.01 μm for a film thickness in the range of 0.1 to 1 μm (see section 2.2 requirement 4 page 20).

The bending and impression of the window

Consider fig. I10 (page 203) which shows that the ratio v_1/R_w remains constant for decreasing values of the ratio R_w/t, when R_w/t is smaller than about 0.5, while the value of v_2/R_w decreases very rapidly at decreasing value of R_w/t and is small compared with the value of v_1/R_w. Therefore the problem can be treated as an elastic half space, as already concluded in section I3.2 (page 202).

A small R_w/t-value is recommended because the total effect of both bending and impression impression is then the smallest possible and the stresses are also the smallest possible. In this thesis we will take the value of the window radius to thickness ratio R_w/t small enough to justify the negligence of the bending and to allow thus for the use of the elastic half space approach. This is convenient since analytical solutions for the calculation of the surface impression are available for a number of practical contact pressure distributions (see Johnson 1985 pp. 45ff.). These solutions can then be used to estimate the pressure influence on the distance h_1 between the window and the focal point.

Now we will first estimate the maximum value of the window radius to thickness ratio R_w/t for which the problem can be treated as an elastic half space, followed by approximation of the impression at various values of the contact pressure.

Using the elastic half space approach we neglect the pressure influence on the lens to window distance s. This is only allowed when the lower surface displacement v_2 is smaller 0.01 μm (which is the maximum uncertainty in the film thickness measurement, see section 2.2 page 20 requirement 4).

Consider now that the maximum contact pressure p is about 50 MPa (section 2.1 page 19 point 8) and that the E-modulus of glass E_w is commonly in the range of 50 to 100 GPa. Then we find for our problem

$$\frac{E_w}{p} \geq \frac{50 \cdot 10^3 \text{ MPa}}{50 \text{ MPa}} = 10^3$$

Further the window radius R_w will be of the order of 1 mm, since a significantly smaller radius will be complicated in the fabrication. Then requiring that v_1 is of the order of 0.001 μm or smaller in the worst case (i.e. when the value of E_w/p is the lowest = 10^3) we find

$$\left(\frac{v_2}{R_w}\right)\left(\frac{E_w}{p}\right) \leq 0.001$$

and consequently the value of the window's radius to thickness ratio must be limited to about (see fig. I8)

$$\frac{R_w}{t} \leq 0.2 \qquad \text{(I11)}$$

When the maximum contact pressure p is restricted to about 5 MPa (as in the preliminary measurements presented in chapter 4) we find

$$\left(\frac{v_2}{R_w}\right)\left(\frac{E_w}{p}\right) \leq 0.01$$

for v_2 of the order of 1 nm or smaller.

and consequently

$$\frac{R_w}{t} \leq 0.25 \qquad \text{(I12)}$$

Then the window impression (i.e. the displacement v_1 of the contacting window surface) is[2]

$$v_1 \approx 10^{-3} R_w \left(\frac{p}{E_w}\right) \qquad \text{(I13)}$$

Therefore v_1 is of the order of 1 μm at a pressure of 50 MPa (and of the order of 0.1 μm at a pressure of 5 MPa). This impression is of the same order as the film thickness (which is in the range of 0.1 to 10 μm, see section 2.1 page 19 point 6) and must therefore be accounted for.

Finally we found the condition (given by eq. (I11) and (I12)) for which the pressure influence on the lens to window distance s (given by v_2) is negligible ($s = s_0$ = independent of the contact pressure p for $p \leq 50$ MPa). Only the pressure influence on the window thickness t ($= t_0 - v_1$) is then significant. The *window thickness t is decreased* by the contact pressure, which leads to *underestimation of* h_1 (and thus to underestimation of the film thickness) when this pressure influence is neglected.

[2] Note that the radius R_w in this formula has no significant meaning as the window's radius since the problem is now treated as an elastic half space. Now this radius has only meaning as the radius of the (circular shaped) loaded part of the surface.

Final remarks

We have seen that the contact pressure influence can be rather large, compared with the expected film thicknesses (in the range of 0.1 to 10 μm), especially due to change in the window's index of refraction and due to impression of the contacting window surface. These two effects are, however, only roughly estimated at the moment and need therefore further investigation.

Measurement of these two effects, which only influence the distance h_1 between the focal point and the window surface, will be possible by loading the elastomeric specimen, used in the experiments, onto the window under dry and static contact conditions (i.e. no lubricant and no relative motion between elastomer and window). Measurement of the focus error signal then yields the value of h_1 versus the contact load, since h_2 is zero owing to lack of lubricant.

For the moment we will only perform measurements at contact pressures of some MPa. These preliminary film thickness measurements are presented in chapter 4 and a more detailed study of the contact pressure influence for the particular contact situation in these experiments will be presented in the next section.

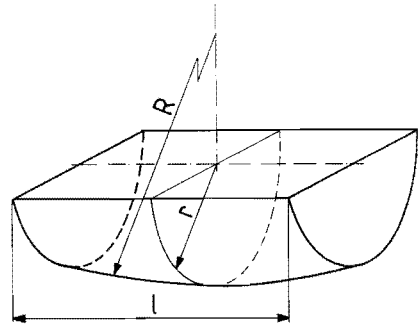
I3.4 The contact pressure influence on the film thickness measurements presented in chapter 4

In this section the contact pressure influence on the film thickness measurements presented in this thesis (chapter 4) will be studied. First the geometry and the mechanical properties of the elastomer and of the window will be given. Then the dimension of the contact area as well as the contact pressure distribution will be determined, followed by estimation of the pressure influence on the indices of refraction of the lubricant and of the window and by estimation of the window's surface impression. Finally the total pressure influence on these film thickness measurements will be discussed.

The geometry and the mechanical properties of the elastomeric specimen

The elastomeric specimen used in chapter 4 is cut from an O-ring seal (Parker-Prädifa code V1 E235 P5008). Therefore its geometry is characterized by two radii (r and R) and the length l (see fig. I11). The window is made from Duran glass.

Figure I11
The elastomeric specimen used for the film thickness measurements in chapter 4.



The following values are given for the geometry of the elastomeric specimen:

$r = 5$ mm
$R = 82.5$ mm
$l = 10$ mm

The mechanical properties of the elastomer (polyurethane seal material, Parker-Prädifa code P5008) were determined by Kanters (1990 section 3.2 pp. 39-42). He described the non-linear elastic behaviour by the so-called neo-Hookean model, which reads

$$\sigma_n = 2\, C_{10} \left(\lambda - \frac{1}{\lambda^2} \right)$$

in which

σ_n	=	Normal stress	[Pa]
λ	=	Elongation = $1 + \varepsilon$	[-]
ε	=	Strain	[-]
C_{10}	=	Constant	[Pa]

In this thesis we will approximate the material behaviour for small strains by the linear Hookean model[3]

$$\sigma_n = E_e\, \varepsilon$$

in which

E_e = The elastomer's E-modulus [Pa]

[3] This approximation is necessary to enable the use of analytical formulas for the contact problem, because they do not account for non-linear mechanical behaviour.

Using

$$\lambda - \frac{1}{\lambda^2} = \frac{\lambda^3 - 1}{\lambda^2} = \frac{(1+\varepsilon)^3 - 1}{(1+\varepsilon)^2}$$

$$= \frac{3\varepsilon + 3\varepsilon^2 + \varepsilon^3}{1 + 2\varepsilon + \varepsilon^2}$$

we find for $\varepsilon \ll 1$

$$\lambda - \frac{1}{\lambda^2} \approx 3\varepsilon$$

yielding

$$E_e \approx 6C_{10}$$

The constant C_{10} in the neo-Hookean model was determined (by 1 hour relaxation tests) to be 7.5 Mpa. Therefore we can write for the Hookean model approximation

$$E_e = 45 \text{ MPa}$$

The other constant describing the mechanical behaviour of solids is the Poisson's ratio ν. For the polyurethane, which is (nearly) incompressible like most elastomers, the Poisson's ratio is approximated by

$$\nu_e = 0.5$$

The geometry and the mechanical properties of the window

The geometry of the window is given by its radius R_w and its thickness t:

$$R_w = 0.5 \text{ mm}$$
$$t = 2 \text{ mm}$$

Therefore $R_w/t = 0.25$ which means that the problem may be approximated by the elastic half space approach, as discussed in section I3.3 (see eq. I12 page 207). (The contact pressure will not exceed 5 MPa as is shown below).

The mechanical behaviour of the window (which is made from Duran glass) is characterized by

$$E_w = 70 \text{ GPa}$$
$$\nu_w = 0.2$$

The dimension and the pressure distribution of the contact

Now we will calculate the dimension of and the pressure distribution in the contact area. These are necessary to enable the calculation of the window's surface impression by the contact load.

The geometry of the elastomeric specimen, given above, yields an elliptical point contact, as shown in fig. I12. The dimension and the pressure distribution can be determined using the Hertzian theory. In this section the solutions proposed by Horowitz (1971) will be used. A slightly different approach (which is, however, less accurate for a slender contact ellipse) is given by Johnson (1985 chapter 4 pp. 84ff.).

The length of the major semi-axis a and of the minor semis-axis b are given by

$$a = \mu_\omega \left(\frac{3\ P\ r}{E_r} \right)^{1/3} \tag{I14}$$

$$b = \nu_\omega \left(\frac{3\ P\ r}{E_r} \right)^{1/3} = \frac{\nu_\omega}{\mu_\omega} a \tag{I15}$$

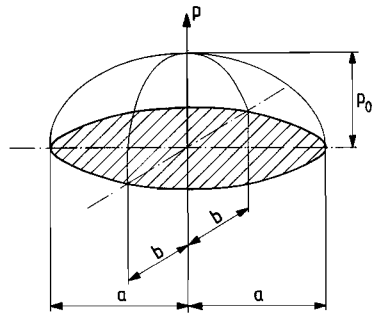
and the pressure p_0 in the centre of the contact (the so-called "Hertzian contact pressure") is given by

$$p_0 = \xi_\omega \left(\frac{E_r^2\ P}{9\ r^2} \right)^{1/3} \tag{I16}$$

in which

P = Contact load [N]

Figure I12
Contact area between the elastomeric specimen and the window and contact pressure distribution.



$$E_r = \text{Reduced E-modulus} \qquad [\text{Pa}]$$

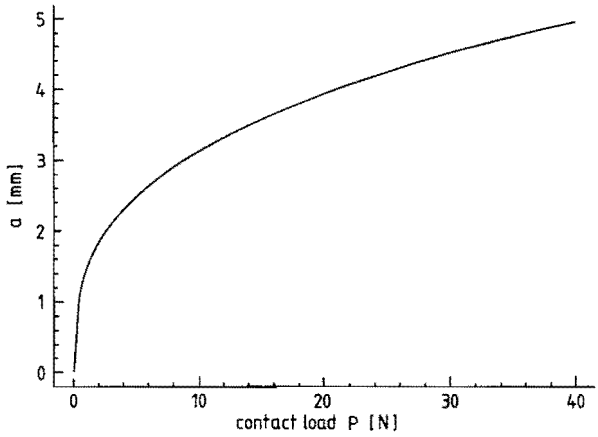
$$= \left(\frac{1 - \nu_e^2}{2\,E_e} + \frac{1 - \nu_w^2}{2\,E_w} \right)^{-1} = 120 \quad \text{MPa}$$

μ_ω, ν_ω and ξ_ω are factors[4], depending of the radii ratio $\omega = R/r$. For our specimen (r = 5 mm and R = 82.5 mm) we find

$$\omega = 16.5$$
$$\mu_\omega \approx 2.9$$
$$\nu_\omega \approx 0.46$$
$$\xi_\omega \approx 0.36$$

The values of the major semi-axis a and of the Hertzian contact pressure p_0 are shown in fig. I13 and I14 respectively for different values of the contact load P. The maximum contact load is taken as 40 N, because then a equals $l/2$, which means that the contact ellipse has reached the boundary of the specimen. The maximum contact pressure is then about 5 MPa.

Figure I13
Major semis-axis a versus contact load P.
r = 5 mm
ω = 16.5
E_r = 120 MPa

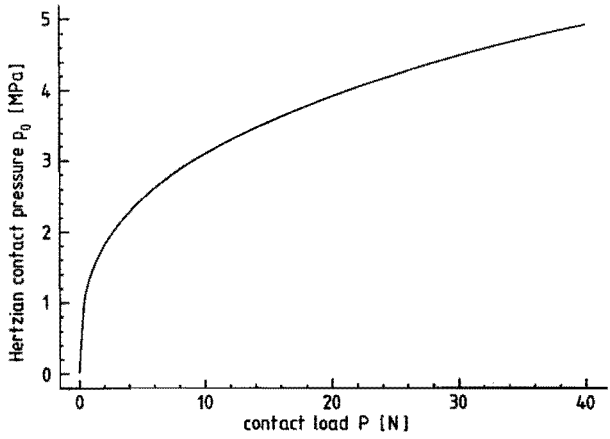


Pressure influence on the film thickness measurement

Now that we have calculated the dimension of the contact area and the contact pressure, we will proceed with the estimation of the pressure influence on the measurements, using the results of the general study in section I3.1 and I3.2 (page 195ff.).

[4] The factor ν_ω should not be confused with the Poisson's ratios ν_e and ν_w.

Figure I14
Hertzian contact pressure p_0 versus contact load P.
r = 5 mm
ω = 16.5
E_r = 120 MPa



At a contact pressure of 5 MPa, the maximum increase in the **index of refraction of the lubricant** is about 0.1 percent, as can be derived from eq. (I7) (page 198). The maximum influence on the film thickness measurement is then also about 0.1 percent. The pressure influence on the lubricant's index of refraction is therefore negligible.

Eq. (I8) (page 202) shows that the maximum increase in the **window's index of refraction** is about 0.0046 percent at a maximum pressure of 5 MPa. The window thickness t is 2 mm, therefore the maximum influence on the value of h_1 is about 0.092 µm, which is not negligible.

The **impression of the window** can be estimated from the calculated Hertzian pressure distribution, using the analytical formulas for the elliptical point contact presented by Johnson (1985 section 3.5 pp. 63ff.). The window impression v_1 in the centre of the contact is given by

$$v_1 = \frac{(1 - \nu_w^2)\, p_0\, a\, b}{2\, E_w} \int_0^\infty \frac{dw}{\sqrt{(a^2 + w)(b^2 + w)w}}$$

$\left(w = \text{Integration parameter } [\text{m}^2]\right)$

Using

$$\beta = \left(\frac{a}{b}\right)^2 = \left(\frac{\mu_\omega}{\nu_\omega}\right)^2 = 40 \qquad (\text{for } \omega = 16.5)$$

and

$$w' = \frac{w}{b^2}$$

we find

$$v_1 = \frac{(1 - \nu_w^2)\, p_0\, a}{2\, E_w} \int_0^\infty \frac{dw'}{\sqrt{(\beta + w')(1 + w')w'}}$$

(Note that β is determined by the geometry of the elastomeric specimen only, not by the contact load nor by the mechanical properties of the elastomer and the window).

The integral can be solved using Gradshteyn and Ryzhik (1980). From their eq. 8 in section 3.131 (p. 220) and eq. 2 in section 8.111 (p. 904) we can derive

$$\int_0^\infty \frac{dw'}{\sqrt{(\beta + w')(1 + w')w'}} = \frac{2}{\sqrt{\beta}}\, \mathbb{F}\left(\frac{\pi}{2}, \sqrt{\frac{\beta - 1}{\beta}} \right)$$

in which

$$\mathbb{F}\left(\frac{\pi}{2}, \sqrt{\frac{\beta - 1}{\beta}} \right) = \text{elliptical integral of the first kind} =$$

$$= \int_0^{\pi/2} \frac{dx}{\sqrt{1 - \frac{\beta - 1}{\beta} \sin^2 x}} = \int_0^{\pi/2} F(x)\, dx$$

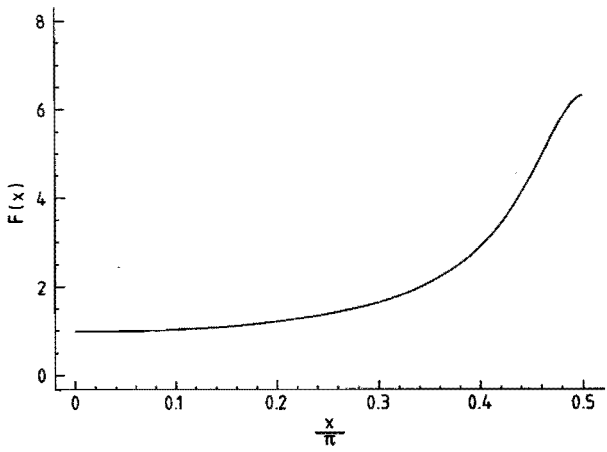
$$\left(x = \text{Integration parameter } [\text{-}] \right)$$

The function

$$F(x) = \frac{1}{\sqrt{1 - \frac{\beta - 1}{\beta} \sin^2 x}} \qquad \text{(I17)}$$

is drawn in fig. I15 for $\beta = 40$ ($\omega = 16.5$), which shows that it is a "smooth" function which can be integrated numerically by e.g. the "trapezium rule" without significant problems (which can e.g. arise from a sharp spike in the curve). This numerical integration then yields

Figure I15
Function $F(x)$ as given in eq. (I17).
$\beta = 40$
$(\omega = 16.5)$



$$\int_0^{\pi/2} \frac{dx}{\sqrt{1 - \frac{\beta - 1}{\beta} \sin^2 x}} = 3.2 \qquad \text{for } \beta = 40$$

with an uncertainty of ±0.1

and therefore we find for $\beta = 40$

$$v_1 \approx \frac{6.4\ (1 - \nu_w^2)\ p_0\ a}{2\ E_w \sqrt{\beta}} =$$

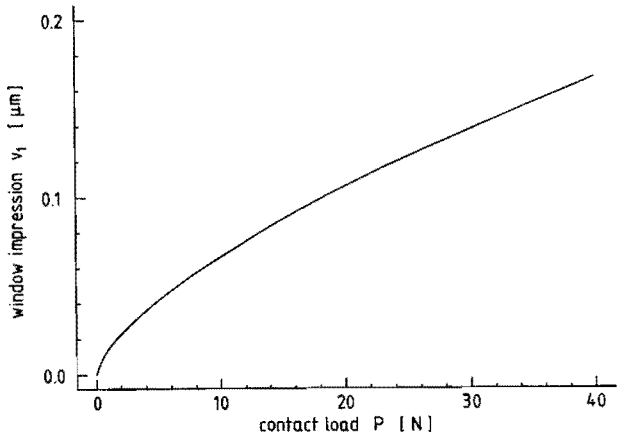
$$\approx 0.5\ \frac{(1 - \nu_w^2)\ p_0\ a}{E_w} \qquad \text{(I18)}$$

Using eq. (I14) and (I16), this equation finally yields the window's surface impression as a function of the load P and is shown in fig. I16. Writing v_1 as a function of the Hertzian contact pressure p_0 we find (using eq. (I14) and eq. (I16) again)

$$v_1 \approx 1.5\ \frac{\mu_\omega\ (1 - \nu_w^2)\ p_0^2\ r}{\xi_\omega\ E_r\ E_w}$$

For our set up in which $r = 5$ mm
$E_r = 120$ MPa

Figure I16
Surface impression of the window at contact load P.
r = 5 mm
ω = 16.5
E_r = 120 MPa
E_w = 70 GPa
φ_w = 0.2



E_w = 70 GPa
ν_w = 0.2
μ_ω ≈ 2.9
and ξ_ω ≈ 0.36

we find

$$v_1 = c_7 p_0^2 \qquad \text{(I19)}$$

in which

$$c_7 \approx 0.0069 \ \mu m \cdot MPa^{-2}$$

Fig. I16 clearly shows that the surface impression is about 0.17 µm at the highest load of 40 N, which is not negligible.

The **total pressure influence** on the film thickness measurement is equal to the total influence on the value of h_1, since the pressure influence on the lubricant's index of refraction n_l is negligible. Then we can write

for p = 0: $$h_1 = h_{10} = n_{l0} \left(f_0 - s_0 - \frac{1}{n_{w0}} t_0 \right)$$

and for $p \neq 0$: $$h_1 = h_{1p} = n_{lp} \left(f_p - s_p - \frac{1}{n_{wp}} t_p \right)$$

Considering that n_l, f and s are (almost) independent of the pressure (i.e. $n_{l0} = n_{lp} = n_l$; $f_0 = f_p = f$; $s_0 = s_p = s$) we can write for the total pressure influence

$$\Delta_p h_l = h_{lp} - h_{l0} = n_l \left(- \frac{t_0 - v_l}{n_{wp}} + \frac{t_0}{n_{w0}} \right)$$

Using eq. (I8) (page 202) and eq. (I19) and considering that $n_l \approx n_{w0}$ we find

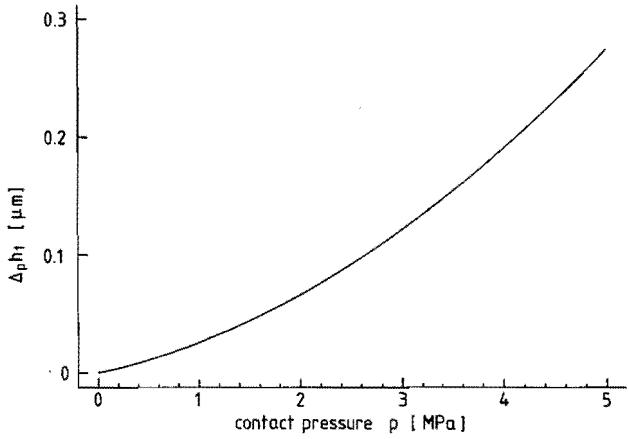
$$\Delta_p h_l = \frac{c_7 p^2 + c_6 t_0 p}{(1 + c_6 p)} \qquad \text{(I20)}$$

in which

t_0	=	2	mm
c_6	=	0.0092	GPa^{-1}
c_7	=	0.0069	μm·MPa^{-1}

The total pressure influence on the film thickness measurement is shown in fig. I17. It is 0.26 μm at a contact pressure of 5 MPa and therefore not negligible.

Figure I17
Total error in the film thickness meas-urement, as expressed by eq. (I20).



I4 The temperature influence

The temperature influence on the film thickness measurements is caused by:
- The temperature dependence of the index of refraction;
- the thermal expansion of the rigid body, including the window, in which the transducer is fixed;

- the temperature dependence of the focal distance of the objective lens.

First these effects will be discussed individually, followed by discussion of the temperature influence on the film thickness measurements in general. Finally the temperature influence on the preliminary film thickness measurements presented in chapter 4 will be estimated in particular.

In this section we will in particular consider the influence of the flash temperature in the contact region (i.e. the temperature rise in the contact area caused by the friction). Fluctuations of the environmental temperature are generally small and their influence will be neglected.

The temperature influence on the focal distance of the collimator lens will not be considered since it can be positioned, if necessary, at a larger distance from the contact area where the flash temperature in the contact area has a negligible influence (i.e. where the temperature equals the ambient temperature). A larger distance between the objective and the collimator lens does in principle not affect the working of the system because the beam between both lenses is collimated.

The temperature influence on the wavelength of the laser radiation will neither be discussed here for the same reason (the laser is always at a larger distance from the contact area than the collimator lens).

I4.1 The temperature dependence of the index of refraction

Analogous to the pressure influence, the temperature influence on the index of refraction is related to the change in density, as expressed by the Lorenz-Lorentz relation (see section I3.1 page 195). This section will consider the indices of refraction of the lubricant and of the window separately.

Temperature influence on the index of refraction of the lubricant

Before we discuss the temperature influence on the index of refraction of the lubricant we must first be able to estimate the contact temperature. In general this is not an easy task because the contact temperature is not only determined by the frictional dissipation in the contact area but also by the environment (e.g the construction of the test rig) which conducts the heat. Important for the contact temperature is e.g. how much heat is transported by conduction through both contacting bodies and how much is transported by convection through

the lubricant film. A further discussion is given by e.g. Hazenberg (1992 section 81).

Assuming that the contact temperature is known we can now estimate the temperature influence on the lubricant's index of refraction. Again we will use the Lorenz-Lorentz relation (eq. I5 page 196) and we need thus an additional expression for the relation between temperature and density.

According to Witt (1974 p. 48-50) the density ρ_l of lubricants is proportional to the temperature and the following estimation was proposed

$$\frac{\rho_l(T = 20\ ^\circ\mathrm{C})}{\rho_l(T = 80\ ^\circ\mathrm{C})} = 1.05 \tag{I21}$$

which means

$$\frac{\rho_l}{\rho_{l0}} = 1 + c_8\,\Delta T \tag{I22}$$

in which

ρ_{l0}	=	Density of the lubricant at 20 °C	[kg·m^{-3}]
ΔT	=	Temperature increase (T - 20 °C)	[K]
c_8	=	Constant	[K^{-1}]

Eq. (I21) then yields

$$c_8 \approx -8{\cdot}10^{-4}\ \mathrm{K}^{-1}$$

The temperature influence on the index of refraction n_l of the lubricant, calculated using eq. (I22) and eq. (I5) is shown in fig. I18. The temperature in the contact will be limited to about 200 °C (see section 2.1 point 9 page 19) and in the range up to 200 °C the variation in the index of refraction is almost proportional to the change in the temperature

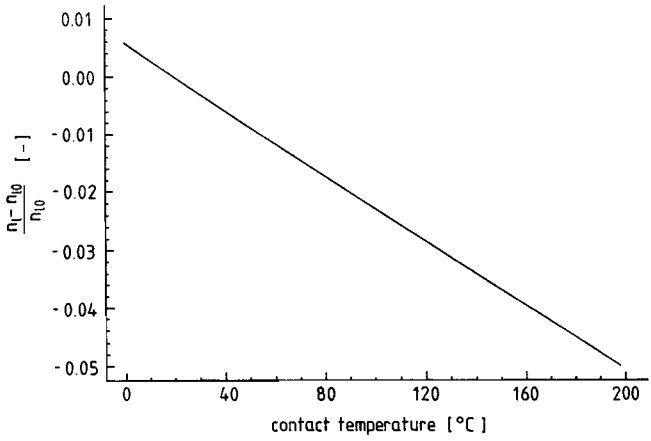
$$\frac{n_l}{n_{l0}} = 1 + c_9\,\Delta T \tag{I23}$$

in which the constant is found to be

$$c_9 \approx -2.8{\cdot}10^{-4}\ \mathrm{K}^{-1}$$

Figure I18
The relative temperature influence on the index of refraction of the lubricant.

c_8 = $-8 \cdot 10^{-4}$ K^{-1}
c = 0.3 $dm^3 kg^{-1}$
ρ_{10} = 0.9 $kg \cdot dm^{-3}$



Temperature influence on the index of refraction of the window

Estimation of the temperature influence on the window's index of refraction n_w is also difficult, because the temperature in the window is not easily determined. Moreover the temperature will not be constant over the window: The farther from the contact area, the lower the temperature will be.

The temperature influence on the window's index of refraction is not provided for the Duran glass, used for the window (section 4.2), but it can also be derived from the density. Considering a cubic volume element with length l_0 and volume V_0 at room temperature (20 °C), we find for the volume after a temperature increase ΔT

$$V = l^3 = l_0^3 (1 + \alpha \Delta T)^3 =$$

$$= V_0 (1 + \alpha \Delta T)^3$$

in which

α = Linear expansion coefficient [K^{-1}]

For $\alpha \Delta T \ll 1$, we find

$$\frac{V}{V_0} \approx 1 + 3 \alpha \Delta T$$

and

$$\frac{\rho_w}{\rho_{w0}} = \frac{V_0}{V} \approx 1 - 3 \alpha \Delta T \qquad (I24)$$

The linear expansion coefficient and density of the window (made from Duran glass) are

$$\alpha \approx 8 \cdot 10^{-6} \ K^{-1}$$
$$\rho_{w0} \approx 3 \ kg \cdot dm^{-3}$$

and the constant c in the Lorenz-Lorentz relation (eq. (I5) page 196) is

$$c \approx 0.09 \ dm^3 kg^{-1}$$

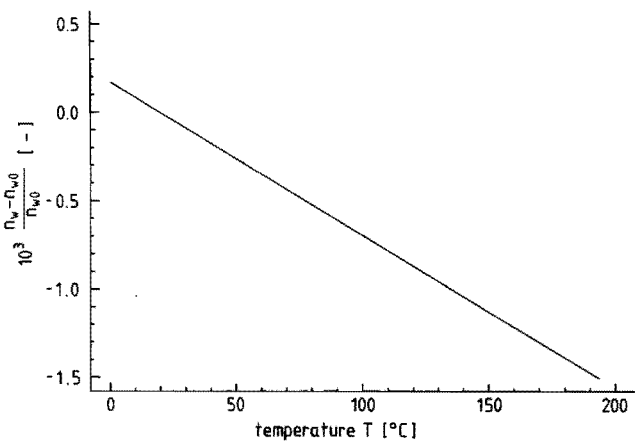
The temperature influence on the window's index of refraction can then be calculated and is shown in fig. I19. This influence is almost linear in the temperature and we can write for (contact) temperatures up to 200 °C

$$\frac{n_w}{n_{w0}} = 1 + c_{10} \Delta T \qquad (I25)$$

in which the constant is found to be

$$c_{10} \approx -8.6 \cdot 10^{-6} \ K^{-1}$$

Figure I19
The relative temperature influence on the index of refraction of the window.
$\alpha = 8 \cdot 10^{-6} \ K^{-1}$
$c = 0.09 \ dm^3 kg^{-1}$
$\rho_{w0} = 3 \ kg \cdot dm^{-3}$



I4.2 The thermal expansion of the rigid body and the window

The thermal expansion of the rigid body influences the distance s between the objective lens and the window, while the expansion of the window influences its thickness t. Both effects influences the film thickness measurement, since the

focal point to window distance h_1 depends on s and t as expressed by eq. (I3) on page 193. Both effects will now be discussed.

Thermal expension of the rigid body and of the window

The temperature inside the rigid body (e.g. at the objective lens position) will be lower than the contact temperature, because it is at a rather large distance from the contact area. Therefore the temperature will be significantly lower, e.g. of the order of 10 °C. The relative temperature influence on the lens to window distance s is given by

$$\frac{s}{s_0} = 1 + \alpha \, \Delta T \tag{I26}$$

The (average) window temperature will be between the contact temperature and the (average) temperature inside the rigid body. Therefore the relative temperature influence on the window thickness t, which is given by

$$\frac{t}{t_0} = 1 + \alpha \, \Delta T \tag{I27}$$

will be somewhat larger than the relative temperature influence on the lens to window distance.

Assuming that the linear expansion coefficient α is about 10^{-5} K^{-1} (α is e.g. $8 \cdot 10^{-6}$ K^{-1} for many types of glass), the increase in the distance s and the window thickness t can be of the order of 0.1 μm, since both s and t will be of the order of 1 mm.

I4.3 The temperature dependence of the focal distance of the objective lens

The temperature dependence of lenses is not not given by lens manufacturers and not easily calculated. In essence it is determined by the temperature influence on the index of refraction of the glass(es) used and the thermal expansion of the lens. In this section it will be roughly estimated using thin lens theory, i.e. we do not account for the thickness of the lens. Also, doublets, triplets and more complicated lens designs, consisting of more elements, are not

considered, because the calculation of the temperature influence on the focal distance will then be very complicated, due to the use of different glasses with different indices of refraction.

The focal distance of a single lens is according to the thin lens approximation (see e.g. Hecht, 1987 section 5.2.3 p. 138)

$$\frac{1}{f} = (n - 1)\left[\frac{1}{R_1} - \frac{1}{R_2}\right] \qquad \text{(I28)}$$

in which: n = Index of refraction of the lens [-]
R_1, R_2 = Radii of both spherical lens surfaces [μm]

Using eq. (I25) and the formula for the linear thermal expansion

$$n = n_0 (1 + c_{10} \Delta T)$$
$$R_i = R_{i0} (1 + \alpha \Delta T)$$

we find

$$\frac{1}{f} = \frac{n_0 (1 + c_{10} \Delta T) - 1}{1 + \alpha \Delta T}\left(\frac{1}{R_{10}} - \frac{1}{R_{20}}\right)$$

$$= \frac{1}{1 + \alpha \Delta T}\left[\frac{1}{f_0} + n_0 c_{10} \Delta T \left(\frac{1}{R_{10}} - \frac{1}{R_{20}}\right)\right]$$

Using

$$\frac{1}{R_{10}} - \frac{1}{R_{20}} = \frac{1}{f_0 (n_0 - 1)}$$

yields

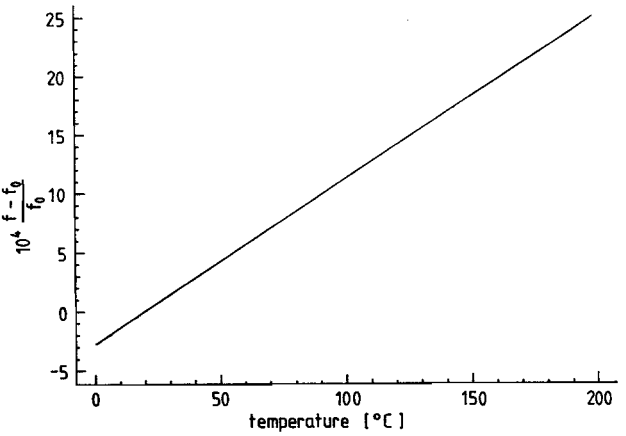
$$\frac{1}{f} = \frac{1}{1 + \alpha \Delta T}\left[1 + \frac{n_0 c_{10} \Delta T}{n_0 - 1}\right]\frac{1}{f_0}$$

Therefore

$$f = (1 + \alpha \Delta T)\left[\frac{n_0 - 1}{n_0 (1 + c_{10} \Delta T) - 1}\right] f_0$$

The estimated temperature dependence of the focal distance f is shown in fig. I20, using $c_{10} = 2 \cdot 10^{-6}$ K^{-1} and $\alpha = 8 \cdot 10^{-6}$ K^{-1} for BK1 (a common glass type for lenses). The focal distance appears to be proportional with the temperature

Figure I20
The relative temperature influence on the focal distance of the objective lens.
α = $8 \cdot 10^{-6}$ K^{-1}
c_{10} = $2 \cdot 10^{-6}$ K^{-1}



$$\frac{f}{f_0} = 1 + c_{11} \Delta T \qquad \text{(I29)}$$

in which c_{11} is found to be

$$c_{11} = 1.4 \cdot 10^{-5} \text{ K}^{-1}$$

Assuming that the temperature increase of the objective lens will be of the order of 10 K and that the focal distance of the objective lens will be of the order of 1 mm, the change in the focal distance will be of the order of 0.1 μm.

I4.4 Discussion on the total temperature influence

In this section we will discuss the total contact temperature influence on the film thickness measurements by focus error detection. First recall the start of this appendix, where the total film thickness was split into two parts:

- The distance h_1 between the focal point and the contacting window surface which is in principle constant;
- The distance h_2 between the focal point and the elastomer's contacting surface, which is to be derived from the measured focus error signal.

In section I1 (page 191) we found

$$h_1 \approx n_1 \left(f - s - \frac{1}{n_w} t \right) \qquad \text{(I30)}$$

and in section I2 (page 193)

$$h_2 \approx n_l h_2' \tag{I31}$$

in which h_2' is directly derived from the focus error signal.

The total temperature influence is determined by the temperature dependence of the indices of refraction of the lubricant and the window (n_l and n_w respectively), by the thermal expansion of the construction (including the window), which influences the values of s and t, and by influence on the focal distance f of the objective lens. The consequence of these effects for the film thickness measurement will now be discussed, using the results of section I4.1, I4.2 and I4.3, to find out whether an effect has significant influence or whether it can be neglected.

The lubricant's index of refraction

In section I3.1 was derived that the lubricant's index of refraction n_l is increased by about 5 percent when the contact temperature is raised to 200 °C. As expressed by eq. (I30) the value of h_l is then also increased by 5 percent, yielding overestimation of h_l by 5 percent if this temperature influence is not accounted for.

Eq. (I31) shows that the value of h_2 is also overestimated by 5 percent (at a contact temperature of 200 °C), when this temperature influence is neglected.

Therefore we can finally conclude that *the total film thickness is overestimated* by about 5 percent. This is more than the tolerated uncertainty of 1 percent at a film thickness in the range of 1 to 10 μm (see section 2.2 requirement 4 page 20). Therefore the temperature influence on the lubricant's index of refraction may not be neglected at contact temperatures up to 200 °C.

The window's index of refraction

The temperature influence on the window's index of refraction n_w is very small: Roughly about 0.0017 percent at a (contact) temperature of 200 °C. As expressed by eq. (I30) and (I31) it only influences the value of h_l. Considering that the window thickness t will be of the order of 1 mm and that the indices of refraction of both the lubricant and of the window are approximately equal to 1.45, we find that h_l is roughly decreased by some μm at a temperature increase

to 200 °C. Negligence of this temperature influence yields thus an *overestimation in the film thickness* of the order of 1 μm.

This temperature influence on h_1 is far from negligible, since the film thickness is in the range of 0.1 to 1 μm (see section 2.1 point 6 page 19).

The thermal expansion of the construction and of the window

The temperature influence on the lens to window distance s as well as on the window thickness t is given by eq. (I26) and eq. (I27) respectively. Considering that the (average) temperature in the rigid body and in the window is lower than the contact temperature (because the heat is generated in the contact) the temperature rise will be smaller than 200 °C (e.g. of the order of 10 °C). Then the increase in the distance s and in the thickness t will be of the order of 0.1 μm when the whole rigid body is made from duran glass (the linear expansion coefficient α is then $8 \cdot 10^{-6}$ K^{-1}). The distance h_1 between the focal point and the window surface will then decrease, as expressed by eq. (I30), yielding an *overestimation of the film thickness* of the order of 0.1 μm when this temperature influence is not accounted for. This error is of the same order as the expected film thickness and therefore not negligible.

The focal distance of the objective lens

The temperature influence on the focal distance of the objective lens could only be roughly estimated. This estimation is given in eq. (I29) on page 224 and it was found that it can be of the order of 0.1 μm, assuming that the temperature rise of the objective lens will possibly be of the order of 10 °C when the contact temperature is 200 °C. As expressed by eq. (I30), the distance h_1 would be increased. Negligence of the temperature influence on the focal distance would then yield an *underestimation* of the film thickness.

Final remarks

We have seen that the temperature influence can be large compared with the expected film thicknesses (in the range of 0.1 to 10 μm), especially due to change in the window's index of refraction, thermal expansion of the construction (including the window) and due to change in the focal length of the objective lens. These effects are, however, only roughly estimated at the moment

and should therefore be further investigated.

Compensation for these effects by theoretical analysis of the temperature distribution in the construction is not easy. Therefore it is recommended to calibrate the complete transducer at different temperatures after it is mounted in the rigid body. Also temperature measurement will be needed at some places in the rigid body (near the objective lens and the window) to enable this compensation experimentally, since the temperature will not be constant in the contact region.

For the moment we will only perform measurements under conditions at which the contact temperature will remain low. These preliminary film thickness measurements are presented in chapter 4 and a more detailed study of the temperature influence for the particular contact situation in these experiments will be presented in the next section.

I4.5 The contact temperature influence on the film thickness measurements presented in chapter 4

In this section the contact temperature influence on the film thickness measurements presented in this thesis (chapter 4) will be studied. The conditions for these measurements are chosen appropriately to keep the contact temperature small. Then the temperature rise of the rigid body will also be small.

It was found by Hazenberg (1992 section 8.2) that the temperature rise in the contact (or the "flash temperature") will be limited to about 0.5 K, using Shell Tellus C320 as lubricant. This was derived for a larger contact load than applied in chapter 4 (Hertzian contact pressure about 3.7 MPa and 1.8 MPa respectively). Therefore, it may be expected that the contact temperature in our experiments will be smaller.

It will be assumed, that the total temperature rise of the rigid body will be also 0.5 K at the maximum. Hazenberg (1992 section 8.3) found that a temperature rise of about 1 K, but this was derived for a different situation, in which the frictional heat dissipation in the contact area was significantly larger.

Now we will derive which influence factors are negligible at the maximum temperature rise of 0.5 K, and which are not negligible. Then the total temperature influence on the film thickness measurements will be estimated, considering only the factors which are not negligible.

At a temperature increase in the contact area of 0.5 K, the **lubricant's index of refraction** is decreased by about 0.014 percent. The maximum influence on the film thickness measurement is then also about 0.014 percent, which is negligible.

The **window's index of refraction** is decreased by about $4.3 \cdot 10^{-4}$ percent. This yields an decrease in h_1 of about 0.0086 μm, which is just smaller than the maximum uncertainty of 0.01 μm tolerated for the measurement of film thicknesses in the range of 0.1 to 1 μm.

The increase in the **window thickness** t and the **lens to window distance** s, due to thermal expansion, is about $4 \cdot 10^{-4}$ percent at the maximum. Consequently, t and s are increased by about 0.008 μm and 0.035 μm respectively[5]. The total decrease in h_1, caused by the thermal expansion, is then about 0.06 μm, since the indices of refraction of the window and of the lubricant are both about 1.5. This influence is therefore not negligible.

The **focal length** f **of the objective lens** can be increased by roughly $7 \cdot 10^{-4}$ percent, i.e. by about 0.07 μm (since f = 10 mm), which is also not negligible.

The **total temperature influence** on the film thickness measurement is equal to the influence on h_1, since the temperature influence on the lubricant's index of refraction n_1 is negligible. Then we can write

for $\Delta T = 0$: $$h_1 = h_{10} = n_{10} \left(f_0 - s_0 - \frac{1}{n_{w0}} t_0 \right)$$

and for $\Delta T \neq 0$: $$h_1 = h_{1T} = n_{1T} \left(f_T - s_T - \frac{1}{n_{wT}} t_T \right)$$

Considering that n_1 is (almost) independent of the temperature (i.e. $n_{10} = n_{1T} = n_1$) we can write for the total temperature influence

$$\Delta_T h_1 = h_T - h_0 =$$

[5] The lens to window distance s must be about 8.67 mm to get the focal point close to the window surface (using an objective lens with a focal distance of 10 mm and a window of 2 mm thick with an index of refraction of about 1.5) as indicated by eq. (A9.3) on page 193 (h_1 must be within some μm at the maximum).

$$= n_1 \left(f_T - s_T - \frac{1}{n_{wT}} t_T \right) - n_1 \left(f_0 - s_0 - \frac{1}{n_{w0}} t_0 \right)$$

Using eq. (I25) (page 221), eq. (I26) and (I27) (page 222) and eq. (I29) (page 224), we find

$$\Delta_T h_1 = n_1 \left((1 + c_{11}\Delta T) f_0 - (1 + \alpha\Delta T) s_0 - \frac{(1 + \alpha\Delta T) t_0}{(1 + c_{10}\Delta T) n_{w0}} \right) +$$

$$- n_1 \left(f_0 - s_0 - \frac{1}{n_{w0}} t_0 \right)$$

Considering that $c_{10} \Delta T \ll 1$, i.e. $(1 + c_{10} \Delta T)^{-1} \approx 1 - c_{10} \Delta T$, we find

$$\Delta_T h_1 = n_1 \left(c_{11}\Delta T f_0 - \alpha\Delta T s_0 - [\alpha\Delta T - c_{10}\Delta T - \alpha c_{10}(\Delta T)^2] \frac{t_0}{n_{w0}} \right)$$

and negligence of the term $\alpha\, c_{10}\, (\Delta T)^2$ finally yields

$$\Delta_T h_1 = n_1 \left(c_{11} f_0 - \alpha s_0 - (\alpha - c_{10}) \frac{t_0}{n_{w0}} \right) \Delta T =$$

$$= c_{12} \Delta T \qquad \text{(I32)}$$

The following values apply for the film thickness transducer:

$$\begin{aligned} n_1 &\approx n_w \approx 1.5 \\ f_0 &= 10 \text{ mm} \\ s_0 &= 8.67 \text{ mm} \\ t_0 &= 2 \text{ mm} \\ c_{10} &= -8.6 \cdot 10^{-6} \text{ K}^{-1} \\ c_{11} &= 1.4 \cdot 10^{-5} \text{ K}^{-1} \\ \alpha &= 8 \cdot 10^{-6} \text{ K}^{-1} \end{aligned}$$

and therefore

$$c_{12} = 0.07 \text{ µm K}^{-1}$$

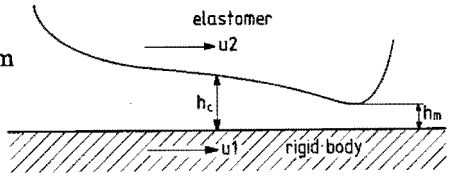
APPENDIX K PREDICTION OF THE LUBRICANT FILM THICKNESS OF AN ELLIPTICAL CONTACT

In section 4.1.2 (page 53) was discussed that it is convenient when the lubricant film thickness can be predicted from the geometry and the running conditions (like the contact load and the velocity), by use of simple analytical formulas. Then the running conditions can be adjusted on purpose to realize a desired film thickness, where interesting effects (e.g. concerning the roughness deformation or the transition from full film to mixed lubrication) are expected. The advantage is that the number of measurements can be reduced, since it will not be necessary to measure over a wide range of running conditions to decide afterwards which measurements are in the interesting range of the film thickness to roughness height ratio.

In this appendix will be shown how the film thickness can be estimated analytically from the geometry of the used elastomeric specimen, the contact load, the elastomer's elasticity, the lubricant's viscosity and the surface velocities.

Analytical formulas have been derived for both the central film thickness h_c (in the centre of the contact) and the minimum film thickness h_m (in the exit region of the contact) (see fig. K1).

Figure K1
General shape of a lubricant film profile.
h_c and h_m are the central and the minimum film thickness respectively;
u_1 and u_2 are the surface velocities.



The practical importance of the minimum film thickness is, that contact between the mating surfaces is generally expected to firstly occur when the ratio of the minimum film thickness and the (undeformed) roughness height is below a critical value (e.g.: $h_m/R_q < 3$, R_q being the Root Mean Square or the standard variation of the roughness height). One may therefore suggest to use the minimum film thickness as criterion for the adjustment of the running conditions. However, if the surface roughness height is of the same order as the nominal film thickness,

the real minimum film thickness is largely determined by the roughness (see e.g. Venner, 1991 pp. 179-184).

Otherwise, the nominal central film thickness is hardly influenced by the roughness of the contacting surfaces, as shown by Kanters (1990 pp. 100-104, 1991) and by Venner (1991, pp. 179-184). Also, the film thickness is approximately equal to the central film thickness in a large part of the contact area, while the minimum film thickness only occurs in a small part. Therefore we will use the (nominal) central film thickness as a criterion for the adjustment of the running conditions.

Calculation of the central film thickness

The contact area of the elastomeric specimen and the (static) pressure distribution are elliptical ("elliptical Hertzian contact", see section I3.4 page 211). The direction of motion is perpendicular to the major (or long) axis (see fig. 4.2 page 52). A further characteristic is that one contacting body (the elastomer) is soft (i.e. the E-modulus is low), which means that the pressure influence on the lubricant's viscosity can be neglected when estimating the film thickness.

Analytical formulas for the film thickness in such a contact are derived by Hamrock and Dowson (1978). In dimensionless form, the central film thickness reads

$$H_c = 7.32 \left(1 - 0.72\, e^{-0.28k} \right) U^{0.64} W^{-0.22} \qquad \text{(K1)}$$

In this equation, k is the "ellipticity" parameter, i.e. the ratio of the major (a) and the minor axis (b) of the contact ellipse. Referring to section I3.4 (page 211) we can write

$$k = \frac{a}{b} = \frac{\mu_\omega}{\nu_\omega}$$

The dimensionless parameters are

$$H_c = \frac{h_c}{r}$$

$$U = \frac{\eta\, u_{av}}{E_r\, r}$$

$$W = \frac{P}{E_r\, r}$$

in which

h_c	=	Central film thickness	[m]
η	=	Dynamic viscosity of the lubricant	[Pa s]
u_{av}	=	Average velocity = $0.5(u_1 + u_2)$	[m s^{-1}]
r	=	Radius (see page 209)	[m]
E_r	=	Reduced E-modulus (see page 212)	[Pa]
P	=	Contact load	[N]

The following values are given (see section I3.4)

$\mu_\omega = 2.9$
$\nu_\omega = 0.46$
$r = 5$ mm
$E_r = 120$ MPa

Further consider that the elastomer is stationary ($u_2 = 0$) and only the rigid body moves (velocity $u_1 = u$). Then the average velocity is

$$u_{av} = \frac{u}{2}$$

and we find for the central film thickness

$$h_c = 23.9 \cdot 10^{-6}\, (\eta u)^{0.64}\, P^{-0.22} \qquad \text{(K2)}$$

(h_c in [m]; (ηu) in [N m^{-1}]; P in [N])

APPENDIX L TEST OF THE SURFACE ROUGHNESS MEASUREMENT ON ELASTOMERS WITH A GLASS PLATE AND LIQUID ON IT

Measuring the elastomer's surface roughness with a glass plate on the surface and a liquid in the contact, the reflectance on the elastomer is generally very low (see appendix D2). This can reduce the accuracy of the measurement and some tests were performed to determine the accuracy. The results are shown below.

Measurements

Roughness measurements were performed on a flat polyurethane plate (PDF material code P5008), using the focus error device of Struik and Chang (1987) (see also section 3.1.2 page 26).

Two measurement series were performed under the following conditions:

1. - No glass plate and liquid on the elastomer;
   - The window (shown in fig. 3.3 page 26) mounted near the objective lens (to prevent influence from spherical aberration as discussed in section 3.2.4 page 40).
2. - A 1.2 mm thick duran glass plate was laid on the elastomer;
   - An oil mixture (75 percent Shell Ondina 15 and 25 percent Shell Ondina 68) was in the contact between the glass plate and the elastomer.
   - The window near the objective lens was removed.

Within one series, 5 measurements were performed on different parts of the surface.

The following parameters apply to all measurements:

- Diameter measurement spot ≈ 1 μm;
- Sample distance in direction $= 0.5$ μm;
- Measurement length $= 640$ μm;

No additional filtering of the measured data was applied.

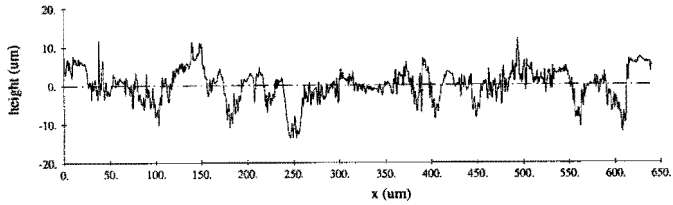
The results of the measurements are presented in fig. L1 and L2, where the measured profile, the height distribution, the autocorrelation and the autopower spectrum are shown. Some derived roughness parameters are shown in table L1.

(The surface roughness characterization is briefly discussed in appendix A1 and a more elaborate discussion can be found in e.g. Halling (1978) and Thomas (1982)).

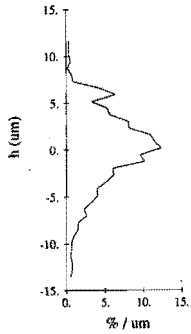
Figure L1

The surface roughness characteristics of the polyurethane plate, measured without glass plate and liquid.

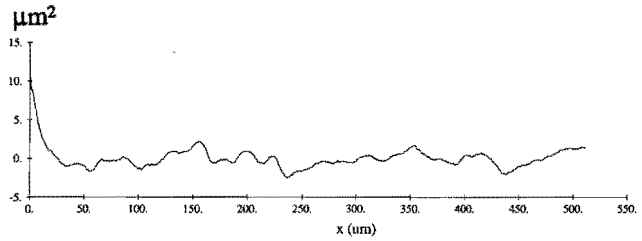
a. Measured profile



b. Height distribution.



c. Autocorrelation curve.



d. Autopower spectrum.

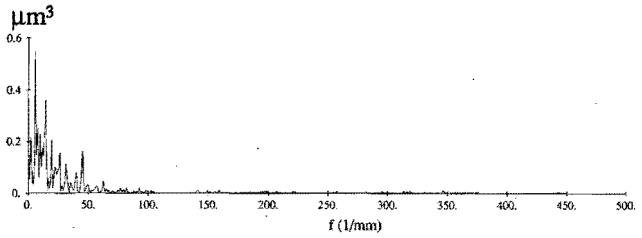
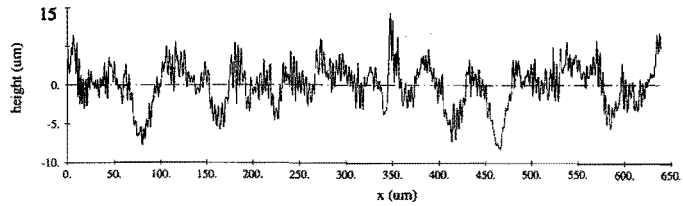


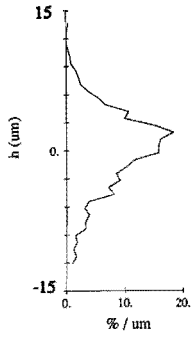
Figure L2

The surface roughness characteristics of the polyurethane plate, measured with glass plate and liquid.

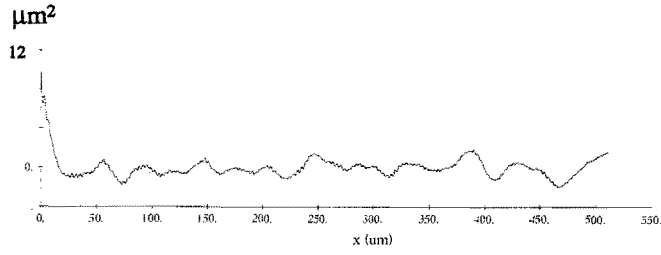
a. Measured profile



b. Height distribution.



c. Autocorrelation curve.



d. Autopower spectrum.

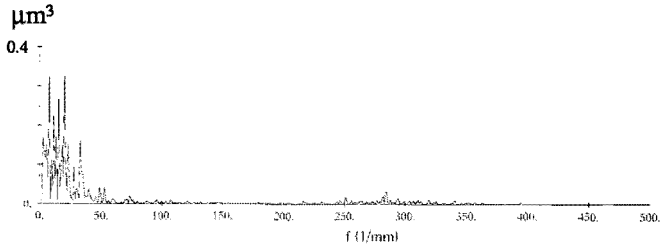


Table L1 Roughness values derived from the measurements shown in fig. L1 (Series 1, without glass plate and oil) and L2 (Series 2, with glass plate and oil).

	R_a [μm]	R_q [μm]	Sk [-]	Kt [-]	$\lambda_{0.5}$ [μm]	$\lambda_{0.1}$ [μm]
Series 1:						
- smallest value	1.82	2.56	-0.98	3.13	4.0	12.5
- largest value	2.56	3.31	-0.38	4.98	7.5	17.0
- **average value**	**2.24**	**2.93**	**-0.66**	**3.97**	**6.0**	**15.3**
Series 2:						
- smallest value	2.30	2.84	-0.29	2.53	4.0	11.5
- largest value	2.71	3.41	-0.04	3.19	5.5	16.0
- **average value**	**2.51**	**3.13**	**-0.13**	**2.87**	**4.6**	**13.7**

Comparison of the results

Comparison of the two measurement series yields the following conclusions:

- The height distribution curves derived from both measurements compare well, at least qualitatively.
- The quantitative correspondence in the derived height values is good (the difference is about 10 percent in R_a and about 6 percent in R_q), considering the spread within one measurement series of 20 percent.
- The correspondence in the derived 50 and 10 percent correlation lengths ($\lambda_{0.5}$ and $\lambda_{0.1}$ respectively) is also reasonable.
- There seems to be some kind of noise on the measured profile of series 2 (fig. L2a), which is seen in the autopower spectrum as spatial frequencies around 300 mm^{-1}. The origin of this noise is not understood and needs more investigation. However, it can be in principle be removed by proper filtering without loss of essential information, as long as its frequencies are significantly higher than the characteristic frequencies in the roughness texture (as is e.g. the case for the polyurethane surface used in these measurements).

Therefore, the performance of the measurement is good, although the reflectance on the oil to elastomer interface is low.

REFERENCES

Austin, R.M., Flitney, R.K., and **Nau, B.S., 1977**, "*Research into factors affecting reciprocating rubber seal performance*", BHRA Fluid Engng., Report RR 1449, Cranfield, U.K.

Archard, J.F., 1957, "Elastic deformation and the lows of friction", *Proc. Royal Soc. London A*, Vol. 243, pp. 190-205

Archard, J.F., 1974, "Surface topography and tribology", *Trib. Int.*, Vol. 7, No. 5, pp. 213-220

Azzam, R.M.A., and **Bashara, N.M., 1977**, "*Ellipsometry and polarized light*", North-Holland Publishing Comp., Amsterdam, ISBN 0-7204-0694-3

Bavel, P. van, 1987, "*Een wiskundig model voor de gemengde filmsmering van dynamisch belaste kontakten*", Internal report no. F20, Eindhoven Univ. Tech., the Netherlands

Bhattacharya, S.K. (ed.), **1986**, "*Metal-filled polymers*", Plastic engng., Vol.11, Marcel Dekker, Basel, ISBN 0-8247-7555-4

Bhushan, B., 1984, "Analysis of the real area of contact between a polymeric magnetic medium and a rigid surface", *J. Trib. (Trans. ASME F)*, Vol. 106, No. 1, pp. 26-34

Bhushan, B., 1985[a], "The real area of contact in polymeric magnetic media - I: Critical assessment of experimental techniques", *ASLE Trans.*, Vol. 28, No. 1, pp. 75-86

Bhushan, B., 1985[b], "The real area of contact in polymeric magnetic media - II: Experimental data and analysis", *ASLE Trans.*, Vol. 28, No. 2, pp. 181-197

Bhushan, B., and **Doerner, M.F., 1989**, "Role of mechanical properties and surface texture in the real area of contact of magnetic rigid disks", *J. Trib. (Trans. ASME F)*, Vol. 111, No. 3, pp. 452-457

Bhushan, B., and **Dugger, M.T., 1990**, "Real area of contact measurements on magnetic rigid disks", *Wear*, Vol. 137, pp. 41-50

Blessing, G.V., and **Eitzen, D.G., 1988**, "Surface roughness sensed by ultra-sound", *Surface Topography*, Vol. 1, pp. 253-267

Blessing, G.V., and **Eitzen, D.G., 1989**, "Surface roughness monitored with an ultrasonic sensor", *SME paper no. MS89-431*, presented at the 1989 SME Int. Conf. and Exp., May 1-4

Blok, H., and **Koens, H.J., 1966**, "The 'breathing' film between a flexible seal and a reciprocating rod", *Proc. I. Mech. E.*, Vol 180, pt 3B, pp. 221-223

Blok, H., and **Rossum, J.J. van, 1953**, "The foil bearing - a new departure in hydrodynamic lubrication", *Lubr. Engngn.*, Vol. 9, pp. 316-320

Born, M., and **Wolf, E., 1970**, "*Principles of optics - electromagnetic theory of propegation, interference and diffraction of light*", Pergamon Press, Oxford, 4th ed.(1st ed. 1959)

Bouwhuis, G., and **Braat, J.J.M., 1978**, "Video disk player optics", *Applied Optics*, Vol. 17, No. 13, pp. 1993-2000

Bouwhuis, G., Braat, J.J.M., Huijser, A., Pasman, J., Rosmalen, G. van, and **Schouhamer Immink, K., 1987**, "*Principles of optical disc systems*", Adam Hilger Ltd., Bristol and Boston, 3rd ed. (1st ed. 1985), ISBN 0-85274-785-3

Bowden, F.D., and **Tabor, D., 1954**, "*Friction and lubrication of solids (Part I)*", Oxford Univ. Press

Breeuwer, R, 1991, "Echo-akoestische meetmethoden voor geometrische metingen", Appendix III in *Meten tijdens de vervaardiging van vormprodukten (final report)* by Visman, R., van Vliet, W. and Schellekens, P.H.J.

Cann, P, and **Spikes, H., 1984**, "Determination of oil films on lubricated surfaces", *Numerical and experimental methods applied (Proc 10th. Leeds-Lyon Symp. on Trib.)*, paper IV(i), pp. 79-82

Çavdar, B., and **Ludema, K.C., 1991**, "Dynamics of dual film formation in boundary lubrication of steels - Part III. Real time monitoring with ellipsometry", *Wear*, Vol. 148, pp. 347-361

Chang, L., Cusano, C., and **Conry, T.F., 1989**, 'Effects of lubricant rheology and kinematic conditions on micro-elastohydrodynamic lubrication', *J. Trib. (Trans. ASME F)*, Vol. 111, No. 2, pp. 344-351

Chang, L., and **Webster, M.N., 1991**, "A study of elastohydrodynamic lubrication of rough surfaces", *J. Trib. (ASME Trans. F)*, Vo. 113, No. 1, pp. 110-115

Cheng, H.S., 1985, "The lubrication of rough surfaces", *Mixed lubrication and lubricated wear (Proc. 11th Leeds-Lyon Symp. on Trib.)*, paper II(i), pp. 11-20

Chow, T.S, and **Penwell, R.C.**, "Mechanical and thermal properties", in: *Metal-filled polymers* by Bhattacharya (ed.), Plastic engng., Vol.11, Marcel Dekker, Basel, ISBN 0-8247-7555-4

Claesen, E.T., 1992, "*Het meetsysteem van de AFM*" M. Sc. Thesis Eindhoven Univ. Tech., The Netherlands, Report no. WPA1335

Cusano, C., and **Wedeven, L.D., 1981**, 'Elastohydrodynamic film thickness measurements of artificially produced nonsmooth surfaces', *ASLE Trans.*, Vol. 24, No. 1, pp. 1-14

Dowson, D., and **Swales, P.D., 1969**, "The development of elastohydrodynamic conditions in a reciprocating seal", *Proc. 4th Int. Conf. on Fluid Sealing*, session 1, paper 1, pp. 2-9, org. by BHRA Fluid Engng., Cranfield, U.K.

Dyson, A., 1978, "Hydrodynamic lubrication of rough surfaces - a review of theoretical work", *Surface roughness effects in lubrication (Proc. 4th. Leeds-Lyon Symp. on Trib.)*, paper III(i), pp. 61-70

Dyson, A., Naylor, H., and **Wilson, A.R., 1966**, "The measurement of oil-film thickness in elastohydrodynamic contacts", *Proc. I.Mech.E*, Vol. 180, Pt. 3B, paper 10, pp. 119-134

Elrod, H.G., 1978, "A review of theories for the fluid dynamic effects of roughness on laminar lubricating films", *Surface roughness effects in lubrication (Proc. 4th. Leeds-Lyon Symp. on Trib.)*, paper I(ii), pp. 11-26

Field, G.J., 1973, *"The elastohydrodynamic lubrication of rectangular section rubber seals under conditions of reciprocating motion"*, Ph. D. Thesis, City Univ. London, UK

Field, G.J., and **Nau, B.S., 1973[a]**, *"The lubrication of rectangular rubber seals under conditions of reciprocating motion, pt. 2, an interferometric study"*, BHRA report RR 1199

Field, G.J., and **Nau, B.S., 1973[b]**, *"The lubrication of rectangular rubber seals under conditions of reciprocating motion, pt. 3, an instrumented seal study"*, BHRA report RR 1200

Field, G.J., and **Nau, B.S., 1973[c]**, "Film thickness and friction measurements during reciprocation of a rectangular section rubber seal ring", *Proc. 6th Int. Conf. on Fluid Sealing*, paper C5, pp. 45-56, org. by BHRA Fluid Engng., Cranfield, U.K.

Field, G.J., and **Nau, B.S., 1976**, "An optical interference method of studying the lubrication of a compliant bearing", *J. Lubr. Tech. (Trans. ASME F)*, Vol. 98, No. 4, pp. 486-490

Flitney, R.K., 1982, "Reciprocating seals", *Trib. Int.*, Vol. 15, No. 4, pp. 219-226

Ford, R.A.J., and **Foord, C.A., 1978**, "Laser-based fluorescence techniques for measuring thin liquid films", *Wear*, Vol. 51, No 2, pp. 289-297

Fox, J.D., Khuri-Yakub, B.T., and **Kino, G.S., 1984**, "Excitation and detection of 8 MHz acoustic waves in air", *Proc. IEEE 1984 Ultrasonic Symp.*, pp. 475-479

Fox, J.D., Khuri-Yakub, B.T., and **Kino, G.S., 1985**, "Complementary air acoustics: An imaging microscope and a resonant rangefinder", *Proc. IEEE 1985 Ultrasonic Symp.*, pp. 463-467

Gabelli, A., and **Poll, G., 1992**, "Formation of lubricant film in rotary sealing contacts: Part I - Lubricant film modeling", *J. of Trib. (Trans. ASME F)*, Vol. 114, No. 2, pp. 280-289

Glatt, I., Livnat, A., and **Kafri, O., 1984**, "Beam-quality analysis of surface finish by moiré deflectometry", *Exp. Mech.*, Vol. 24, pp. 248-251

Gradshteyn, I.S, and **Ryzhik, I.M., 1980**, *"Table of Integrals, Series and Products"* Acadadamic Press, New York, ISBN 0-12-294760-6
(Translated from Russia by Jeffrey, A.)

Greenwood, J.A., and **Tripp, J.H., 1971**, "The contact of two nominally flat rough surfaces", *Proc. I. Mech. E.*, Vol. 185, pp. 625-633 (discussion: pp. D243-D250)

Greenwood, J.A., and **Williamson, J.B.P., 1966**, "Contact of nominally flat surfaces", *Proc. Royal Soc. London*, Vol. A295, pp. 300-319

Grzegorczyk, D., and **Feineman, G., 1974**, "*Handbook of palstics in electronics*", Reston Publ. Comp.

Hall, D.R, and **Jackson, P.E., 1989**, "*The physics of laser resonators*', Adam Hilger, Bristol, ISBN 0-85274-117-0, pp. 9-20

Halling, J., 1978, "*Principles of tribology*", MacMillan Press Ltd., London, ISBN 0 333 24686 1

Hamrock, B.J., and **Dowson, D., 1978**, "Elastohydrodynamic lubrication of elliptical contacts for materials of low elastic modulus I - Fully flooded conjunction", *J. Lubr. Tech. (Trans. ASME F)*, Vol. 100, No. 2, pp. 236-245

Hamrock, B.J., and **Dowson, D., 1981**, "*Ball bearing lubrication*", John-Wiley & Sons, New York, ISBN 0-471-03553-X

Hanekamp, L.J., 1983, "*Clean and oxygen covered copper and nickel surfaces investigated by ellipsometry and Auger electron spectroscopy*", Ph. D. Thesis, Twente Univ. Tech., The Netherlands

Hazenberg, J.M.A, 1992, "*De ontwikkeling van een proefopstelling voor onderzoek aan gesmeerde elastomeer-metaal contacten*" M. Sc. Thesis, Eindhoven Univ. Tech., The Netherlands, Report no. T&M A26

Hendriks, C., 1992[a], "*Meting van de vervormde ruwheid van elastomeren onder statische kontaktbelasting - deel 1: dataverwerking en interpretatie*", Eindhoven Univ. Tech., The Netherlands, Report no. T&M D53

Hendriks, C., 1992[b], "*Modellen van het statisch belaste, droge kontakt - literatuuroverzicht*", Eindhoven Univ. Tech., The Netherlands, Report no. T&M D54

Hendriks, C., 1993, M.Sc. Thesis, Eindhoven Univ. Tech., The Netherlands (to be published)

Hecht, E., 1987, "*Optics*", Adison-Wesley Publ. Comp., Reading (Massachusetts), ISBN 0-201-11609-X, 2nd ed. (1st ed. 1974)

Hirano, F., and **Ishsiwata, H., 1965**, "The lubricating condition of a lip seal", *Proc. I. Mech. E.*, Vol. 180, Pt. 3B, paper 15, pp. 187-196

Hirano, F., and **Kaneta, M., 1973**, "Elastohydrodynamic condition in elliptic contact in reciprocating motion", *Proc. 6th Int. Conf. on Fluid Sealing*, paper C2, pp. 11-24, org. by BHRA Fluid Engng., Cranfield, U.K.

Hori, Y., Kato, T., and **Narumiya, H., 1981**, "Rubber surface squeeze film", *J. Lubr. Eng. (Trans. ASME F)*, Vol. 103, No. 3, pp. 398-405

Horowitz, A., 1971, "A contribution to the engineering design of machine elements involving contrashaped contacts", *Israel J. of Tech.*, Vol. 9, No. 4, pp. 311-322

Hoult, D.P., and **Takiguchi, M., 1991**, "Calibration of the laser fluorescence technique compared with quantum theory", *Trib. Trans.*, Vol. 34, No. 3, pp. 440-444

Hu, Y., and **Zheng, L., 1985**, "Effect of three-dimensional roughness on lubrication - a numerical perturbation solution", *Proc. JSLE Int. Trib. Conf.*, Vol. 3, paper 7D-4, pp. 1109-1114

Hu, Y, and **Zheng, L., 1989**, "Some aspects of determining the flow factors", *J. of Trib. (Trans. ASME F)*, Vol. 111, No. 3, pp. 525-531

Jackson, A., and **Cameron, A., 1976**, 'An interferometric study of the EHL of rough surfaces', *ASLE Trans.*, Vol. 19, No. 1, pp. 50-60

Jacobson, B., 1990, "Mixed lubrication", *Wear*, Vol. 136, No. 1 pp. 99-116

Jagger, E.T., 1956, "Rotary shaft seals: The sealing mechanism of synthetic rubber seals running at atmospheric pressure", *Proc. I. Mech. E.*, Vol. 171, No. 18, pp. 597-604

Jagger, E.T., 1957, "Study of the lubrication of synthetic rotary shaft seals", *Proc. Conf. Lub. and Wear*, paper 93, pp. 409-414

Johannesson, H.L., 1989, "Piston rod seal and scraper ring interaction in hydraulic cylinders", *Wear*, Vol. 130, No. 1, pp. 17-27

Johnson, K.L., 1985, "*Contact Mechanics*", Cambridge University Press, Cambridge, ISBN 0 521 25576 7 and 0 521 34796 3

Jong, F.W. de, 1990, "*Literatuuronderzoek naar electrisch geleidende polymeren*", Internal report no. T&M D51, Eindhoven Univ. Tech., the Netherlands

Kafri, O., and **Glatt, I., 1990**, "*The physics of moire metrology*", John Wiley & Sons Inc., New York, ISBN 0-471-50967-1

Kagami, J., Hatazawa, T., and **Koike, K., 1989**, "Measurement of surface profiles by the focussing method", *Wear*, Vol. 134, pp. 221-229

Kalsi, M.S., 1975, "*Elastohydrodynamic lubrication of offset O-ring rotary seal*", Ph. D. Thesis, Univ. of Houston, Texas USA

Kalsi, M.S., 1981, "Elastohydrodynamic lubrication of offset O-ring rotary seal", *J. Lubr. Tech. (Trans. ASME F)*, Vol. 103, No. 3, pp. 414-427

Kaneta, M., and **Cameron, A., 1980**, "Effects of asperities in elastohydrodynamic lubrication", *J. Lubr. Engng. (Trans. ASME F)*, Vol. 102, No. 3, pp. 374-379

Kanters, A.F.C., 1990, "*On the calculation of leakage and friction of reciprocating elastomeric seals*", Ph. D. Thesis, Eindhoven Univ. Tech., The Netherlands

Kanters, A.F.C., 1991, "On the elastohydrodynamic lubrication of reciprocating elastomeric seals: the influence of the surface roughness", *Vehicle Tribology (Proc. 17th Leeds-Lyon Symp. on Tribology)*, Elsevier, Amsterdam, ISBN 0-444-88796-2, paper XIII(i), pp. 357-364

Kanters, A.F.C., Verest, J.F.M., and **Visscher, M., 1990**, "On reciprocating elastomeric seals: Calculation of film thicknesses using the Inverse Hydrodynamic Lubrication theory", *Trib. Trans.*, Vol. 33, No. 3, pp. 301-306

Kanters, A.F.C., and **Visscher, M., 1989**, "Lubrication of reciprocating seals: experiments on the influence of surface roughness on friction and leakage", *Tribological design of machine elements (Proc. 15th Leeds-Lyon Symp. on Tribology)*, Elsevier, Amsterdam, paper III(iii), pp. 69-77

Kassfeldt, E., 1987, "*Analysis and design of hydraulic cylinder seals*", Ph. D. Thesis, report no. 1987:56D, Luleå Univ. Tech., Sweden, pp. E1-E16
(republished in: **Jacobson, B.O., 1991**, "*Rheology and elastohydrodynamic lubrication*", Tribology Series 19, Elsevier, Amsterdam, ISBN 0-444-88146-8 , 219-227)

Kawahara, Y., Muto, Y., Hirabayashi, H., and **Matsushima, A., 1981**, "A study of characteristics of the controlled oil leakage of valve-stem seals for automotive engines", *ASLE Trans.*, Vol. 24, No. 2, pp. 205-214

Kawahara, Y., Ohtake, Y., and **Hirabayashi, H., 1981**, "Oil film formation of oil seals for reciprocating motion", *Proc. 9th Int. Conf. on Fluid Sealing*, paper C2, pp. 73-85, org. by BHRA Fluid Engng., Cranfield, U.K.

Köhnlechner, R.J., 1980, "*Untersuchungen zur Schmierfilmdicken in Druckluftzylindern - Beurteilung der Abstreifwirkung und des Reibungsverhaltens von Pneumatikdichtungen mit Hilfe eines neu entwickelten Schmierfilmdicken-meßverfahrens*", Ph. D. thesis, Stuttgart Univ. Tech., Germany (Volume 42 of the series "Forschung und Praxis" Springer Verlag, Berlin)

Kohno, T., Ozawa, N., Miyamoto, K., and **Musha, T., 1985**, "Practical non-contact surface measuring instrument with one nanometre resolution", *Prec. Engng.*, Vol. 7, No. 4, pp. 231-232

Kohno, T., Ozawa, N., Miyamoto, K., and **Musha, T., 1988**, "High precision optical surface sensor", *Applied Optics*, Vol. 27, No. 1, pp. 103-108

Krauter, A.I., 1982, "Measurement of oil film thickness for application to elastomeric stirling engine rod seals", *J. Lubr. Techn. (Trans. ASME F)*, Vol. 104, No. 4, pp. 455-459

Kweh, C.C., Evans, H.P., and **Snidle, R.W., 1989**, "Micro-elastohydrodynamic lubrication of an elliptical contact with transverse and three-dimensional sinusoidal roughness", *J. Trib. (Trans. ASME F)*, Vol. 111, No. 4, pp. 577-584

Kweh, C.C., Patching, M.J., Evans, H.P., and **Snidle, R.W., 1992**, "Use of profile data in simulation of lubrication of real surfaces", *Int. J. Mach. Tools Manufact.*, Vol. 32, No. 1/2, pp. 75-81

Kusy R.P., 1986, "Applications", in: *Metal-filled polymers* by Bhattacharya (ed.), Plastic engng., Vol.11, Marcel Dekker, Basel, ISBN 0-8247-7555-4

Leather, J.A., and **McPherson, P.B., 1978**, "The practical use of electrical measurements in lubricated contacts", *Surface roughness effects in lubrication (Proc. 4th Leeds-Lyon Symp. on Tribology)*, paper V(iii), pp. 155-162

Lubrecht, A.A., and **Ioannides, E., 1991**, "A fast solution of the dry contact problem and the associated sub-surface stress field, using multilevel techniques", *J. Trib. (Trans. ASME F)*, Vol. 113, No. 1, pp. 128-133

Lubrecht, A.A., Napel, W.E. ten, and **Bosma, R., 1988**, "The influence of longitudinal and transverse roughness on the elastohydrodynamic lubrication of circular contacts", *J. Trib. (Trans. ASME F)*, Vol. 110, No. 3., pp. 421-426

Majumdar, A. and **Bhushan, B., 1990**, "Role of fractal geometry in roughness characterization and contact mechanics of surfaces", *J. Trib. (Trans. ASME F)*, Vol. 112, No. 2, pp. 205-216

Majumdar, A., and **Bhushan, B., 1991**, "Fractal model of elastic-plastic contact between rough surfaces", *J. Trib. (Trans. ASME F)*, Vol. 113, No. 1, pp. 1-11

Majumdar, A., and **Tien, C.L., 1990**, "Fractal characterization and simulation of rough surfaces", *Wear*, Vol. 136, pp. 313-327

McClune, C.R., 1974, "*The properties of liquid films between highly elastic surfaces*", Ph. D. Thesis, Cambridge Univ., UK

McClune, C.R., and **Briscoe, B.J., 1977**, "Elastohydrodynamic films formed between a rubber cylinder and a glass plate: a comparison of theory and experiment", *J. Phys. D. Appl. Phys.*, Vol. 10, pp. 587-598

McClune, C.R., and **Tabor, D., 1978**, "An interferometric study of lubricated rotary face seals", *Tribology Int.*, Vol. 11, No. 4, pp. 219-227

McCool, J.I., 1986, "Comparison of models for the contact of rough surfaces.", *Wear*, Vol. 107, No. 1, pp. 37-60

Meadows, D.M., Johnson, W.O., and **Allen, J.B., 1970**, "Generation of surface contours by Moiré patterns", Appl. Optics, Vol. 9, No. 4, pp. 942-947

Meyer, F., and **Loyen, C.J., 1975**, "The creep of oil on steel followed by ellipsometry", *Wear*, Vol. 33, No, 8, pp. 317-323

Mignot, J., and **Gorecki, C., 1983**, "Measurement of surface roughness: Comparison between a defect-of-focus optical technique and the classical stylus technique", *Wear*, Vol. 87, pp. 39-49

Mitsui, K., 1986, "In-process sensors for surface roughness and their applications", *Precision Engng.*, Vol. 8, No. 4, pp. 212-220

Mitsui, K., Ozawa, N., and **Kohno, T., 1985**, "Development of a high resolution in-process sensor for surface roughness by laser beam", *Bull. Japan Soc. of Prec. Engng.*, Vol. 19, No. 2, pp. 142-143

Neal, W.E.J., and **Fane, R.W., 1973**, "Ellipsometry and its applications to surface examination", *J. Phys. E (Sci. Instr.)*, Vol. 6, pp. 409-416

Ogata, M., Fujii, T., and **Shimotsuma, Y., 1987**, "Study on fundamental characteristics of rotating lip-type oil seals", *Fluid film lubrication - Osborne Reynolds centenary (Proc. 13th Leeds-Lyon Symp. on Trib.)*, paper XVII(ii), pp. 553-560

Ohara, K., 1976, "Observation of surface profiles and the nature of the contact between polymer films by multiple beam interferometry", *Wear*, Vol. 39, pp. 251-262

Onions, , and Archard, , 1973, "The contact of surfaces having a random structure", *J. Phys. D: Appl. Phys.*, Vol. 6, pp. 289-304

Patir, N., 1978, "*Effects of surface roughness on partial film lubrication using an average flow model based on numerical simulation*", Ph. D. Thesis, Northwestern Univ., Il. USA

Patir, N., and Cheng, H.S., 1978, "An average flow model for determining effects of three dimensional roughness on partial hydrodynamic lubrication", *J. Lub. Tech. (Trans. ASME F)*, Vol. 100, No. 1, pp. 12-17

Patir, N., and Cheng, H.S., 1979[a], "Application of average flow model to lubrication between rough sliding surfaces", *J. Lub. Tech (Trans. ASME F)*, Vol. 101, No. 2, pp. 220-230

Patir, N., and Cheng, H.S., 1979[b], "Effects of surface roughness orientation on the central film thickness in EHD contacts", *Elastohydrodynamics and related topics (Proc. 5th Leeds-Lyon Symp. on Trib.)*, paper II(ii), pp. 15-20

Podbevsek, F.P., 1992, "*On the roughness deformation of elastomers under static load: asperity interaction*", Report Eindhoven Univ. of Techn.

Poll, G., and Gabelli, A., 1992[a], "Formation of lubricant film in rotary sealing contacts: Part II - A new measuring principle for lubricant film thickness", *J. of Trib. (Trans. ASME F)*, Vol. 114, No. 2, pp. 290-297

Poll, G., Gabelli, A., Binnington, P.G., and Qu, J., 1992[b], "Dynamic mapping of rotary lip seal lubricant film by fluorescent image processing", *Proc. 13th. Int. Conf. on Fluid Sealing*, pp. 55-77, org. by BHRA Fluid Engng., Cranfield, U.K.

Poulter, T.C., Ritchey, C., and Benz, C.A., 1932, "The effect of pressure on the index of refraction of paraffin oil and glycerine", *Phys. Rev.*, Vol. 41, pp. 366-367

Riedling, K., 1988, "*Ellipsometry for industrial applications*", Springer-Verlag, Wien, ISBN 3-211-82040-X / Springer-Verlag, New York, ISBN 0-387-82040-X,

Roberts, A.D., 1968, "The preparation of optically smooth rubber surfaces", *Engng. Materials and Design*, Vol. 11, No. 4, pp. 579-580

Roberts, A.D., and Tabor, D., 1968, "Fluid film lubrication of rubber - An interferometric study", *Wear*, Vol. 11, pp. 163-166

Roberts, A.D., and Tabor, D., 1971, "The extrusion of liquids between highly elastic solids", *Proc. Roy. Soc. London*, Vol. A325, pp. 323-345

Saure, M., 1979, "*Kuststoffe in der Elektrotechnik - AEG Telefunken HandBücher Band 22*", AEG Telefunken, 1979

Schmidt, U., Bodschwinna, H., and **Schneider, U., 1987**[a], "Mikro EHD: Einfluß der Oberflächenrauheit auf die Schmierfilmausbildung in realen EHD-Wälzkontakten. Teil I: Grundlagen", *Antriebstechnik*, Vol. 26, No. 11, pp. 55-60

Schmidt, U., Bodschwinna, H., and **Schneider, U., 1987**[b], "Mikro EHD: Einfluß der Oberflächenrauheit auf die Schmierfilmausbildung in realen EHD-Wälzkontakten. Teil II: Ergebnisse und rechnerische Auslegung eines realen EHD-Wälzkontakts", *Antriebstechnik*, Vol. 26, no. 12, pp. 55-60

Schmutz, W., 1984, "*Fluoreszenzmeßverfahren zur Schmierfilmdickenmessung in Wälzlagern*", Ph. D. thesis, Stuttgart Univ. Tech., Germany (Volume 79 of the series "Forschung und Praxis", Springer Verlag, Berlin)

Schouten, M.J.W., and **Gawlinski, M.J., 1978**[a], "Radial-Wellendichtringen als ein Elastohydrodynamisches Problem", *Proc. 6. Dichtungstagung (6th. Int. Conf. on Fluid Sealing)*, Dresden, Germany (Dem. Rep.)

Schouten, M.J.W., and **Gawlinski, M.J., 1978**[b], "Enkele tribotechnische aspecten van een roterende-asafdichting", *PT-Werktuigbouw*, Vol. 33, No. 9, pp. 491-494

Schrader, K., 1978, "*Beitrage zur Klärung des Abdichtvorganges Gummi elastischer Abdichtungen axial verschiebbarer hydrostatischer Bauteile*", Ph. D. Thesis, Dresden Univ. Tech., Germany (Dem. Rep.)

Sharp, 1988, '*Technical literature for Laser Diode - Model No. LTOH1OM*', Sharp Corporation, Electronic Components Group, Doc. No. DG-63X064

Sherrington, I., and **Smith, E.H., 1988**, "Modern measurement techniques in surface metrology: Part II; Optical instruments", *Wear*, Vol. 125, No. 3, pp. 289-308

Sichel, E.K., 1982, "Tunneling conduction in carbon-polymer composites", in: *Carbon-black polymer composites* by Sichel, E.K. (ed.), Plastics engng., Marcel Dekker

Smart A.E., and **Ford, R.A.J., 1974**, "Measurement of thin liquid films by a fluorescence technique", *Wear*, Vol. 29, No 1, pp. 41-47

Spiegel, M. R., 1968, "*Mathematical handbook of formulas and tables*", Schaum's Outline Series in mathematics, McGraw-Hill Book Comp., New York

Stakenborg, M.J.L., 1988, "*On the sealing and lubrication mechanism of radial lip seals*", Ph. D. Thesis, Eindhoven Univ. Tech., The Netherlands

Struik, K.G., and **Chang, F.M., 1987**, "Measurement of shape and roughness by a modified compact disc sensor coupled to a personal computer", *Proc. 4th Int. Seminar on Precision Engineering*, Cranfield Inst. of Technology, UK, pp. 81-90, 11-14 May

Svitashev, K.K., Semenenko, A.I., Semenenko, L.V. and **Sokolov, V.K., 1971**, "Use of a convergent light beam for ellipsometric measurement", *Opt. Spectr.*, Vol. 30, pp. 288-291

Svitashev, K.K., Semenenko, A.I., Semenenko, L.V. and **Sokolov, V.K., 1973**, "Ellipsometry based on a convergent light beam", *Opt. Spectr.*, Vol. 34, pp. 542-544

Swales, P.D., Dowson, D., and **Latham, J.L., 1972**, "Theoretical and experimental observations of the behaviour of soft elastic materials under elastohydrodynamic conditions", *Proc. EHL Symp.*, org. by I.Mech.E, paper C4, pp. 22-28

Szilard, J., 1982, "Review of conventional testing techniques", Chapter 2 of "*Ultrasonic testing - Non-conventional testing techniques*" (ed.: Szilard, J.), pp. 25-52, John Wiley & Sons, Chichester, ISBN 0 4071 27938 2

Takasaki, H., 1973, "Moiré topography", *Appl. Optics*, Vol. 12, No. 4, pp. 845-850

Thomas, T.R., 1982, "*Rough surfaces*", Longman Group Ltd., Harlow (UK)

Thomas, T.R., and **King, M., 1977**, "*Surface topography in engineering: as state of the art review and bibliography*", BHRA Fluid Engng., Cranfield (UK), ISBN 0 900983 66 3

Thomas, T.R., and **Sayles, R.S., 1978**, "Some problems in the tribology of rough surfaces", *Trib. Int.*, Vol. 11, No. 3, pp. 163-168

Ting, L.L., 1980[a], "Development of a laser fluorescence technique for measuring piston ring oil film thickness", *J. Lubr. Engng. (Trans. ASME F)*, Vol. 102, No. 1, pp. 165-170

Ting, L.L., 1980[b], "Development of a laser fluorescence technique for measuring piston ring oil film thickness", Authors closure on the discussion, *J. Lubr. Engng. (Trans. ASME F)*, Vol. 102, No. 1, p. 171

Tønder, K., and **Jakobsen, J., 1992**, "Interferometric studies of effects of striated roughness on lubricant film thickness under elastohydrodynamic conditions", *J. Trib. (Trans. ASME F)*, Vol. 114, No. 1, pp. 52-56

Venner, C.H., 1991, "*Multilevel solution of the EHL line and point contact problems*", Ph. D. Thesis, Twente Univ., The Netherlands, ISBN 90-9003974-0

Vergne, Ph., Villechaise, B., and **Berthe, D., 1985**, "Elastic behavior of multiple contacts: Asperity interaction", *J. Trib. (Trans. ASME F)*, Vol. 107, No. 2, pp. 224-228

Visscher, M., 1989, "*Analysis of electrical methods for measurement of leakage and lubricant film thicknesses of elastomeric seals*", report no. IAT/89.185, Eindhoven Univ. Tech., The Netherlands

Visscher, M., and **Kanters, A.F.C., 1990**, "Literature review and discussion on measurements of leakage, lubricant film thickness and friction of reciprocating elastomeric seals", *Lubr. Engng.*, Vol. 44, No. 12, pp. 785-791

Visscher, M., and **Struik, K.G., 1992**, 'The measurement of the roughness deformation of elastomers under static load', *Proc. 8th. Int. Coll. "Tribology 2000"*, Technische Akademie Esslingen, Germany, paper 9.6.

Weglein, R.D., and **Wilson, R.G., 1977**, "Image resolution of the scanning acoustic microscope", *Appl. Phys. Lett.*, Vol. 31, No. 12, pp. 793-796.

Wernecke, P.W., 1983, "*Untersuchung der physikalischen Vorgänge in Spalten von Hydraulikdichtungen*", Ph. D. Thesis, Aachen Univ. Tech., Geramany

Wernecke, P.W., 1987, "Analysis of the reciprocating sealing process", *Proc. 11th Int. Conf. on Fluid Sealing*, paper E1, pp. 249-277, org. by BHRA Fluid Engng., Cranfield, U.K.

Witt, K., 1974, "*Die Berechnung physikalischer und thermodynamischer Kennwerte von Druckflüssigkeiten, sowie die Bestimmung des Gesamtwirkungsgrades an Pumpen unter Berücksichtigung der Thermodynamik fur die Druckflüssigkeit*", Ph. D. Thesis, Eindhoven Univ. Tech., The Netherlands

Woo, K.L., and **Thomas, T.R., 1980**, "Contact of rough surfaces: A review of experimental work", *Wear*, Vol. 58, pp. 331-340

Wyant, J.C., Koliopoulis, C.L., Bhushan, B., and **Basila, D., 1986**, "Development of a three-dimensional noncontact digital optical profiler", *J. Trib. (Trans. ASME F)*, Vol. 108, No. 1, pp. 1-8

Xian Ling, and **Zheng Linqing, 1991**, "A numerical model for the elastic contact of three-dimensional real rough surfaces", *Wear*, Vol. 148, pp. 91-100

Zhu, D., Cheng, H.S., and **Hamrock, B.J., 1990**, "Effect of surface roughness on pressure spike and film constriction in elastohydrodynamically lubricated line contacts", *Trib. Trans.*, Vol. 33, No. 2, pp. 267- 273

NAWOORD

Dit onderzoek is door een groot aantal mensen op verschillende manieren ondersteund. De aard van dit onderzoek bracht met zich mee dat er veel (incidenteel) kontakt is geweest met deskundigen op allerlei gebied van fysische (meet-) principes. Op deze plaats wil ik iedereen die op wat voor manier dan ook een bijdrage geleverd heeft bedanken.

Arno Kanters heeft mij op het spoor van dit onderzoek gebracht. De intensieve samenwerking met hem en de daar uit voortvloeiende discussies zijn van grote waarde geweest voor de verdieping van mijn inzichten in de tribologie.

Vanuit de groep Precision Engineering heb ik veel ondersteuning gehad. Met name Klaas Struik wil ik hier bedanken voor de intensieve samenwerking.

Jan Peels heeft met grote deskundigheid veel experimenten doorgevoerd. Ook waren zijn kontakten met de CTD zeer waardevol. Verschillende testmaterialen zijn daar vervaardigd en de gebruikte meetopnemers zijn bij de CTD aangepast dan wel samengesteld, met name door Toon van Kalmthout, Theo Maas, Jan Versteeg en hun medewerkers. De elastomere proefmaterialen werden door Parker-prädifa ter beschikking gesteld.

Prof. Braat en dr. Baalbergen (Philips Nat.-Lab) en ir. Greve en ir. Kooijman (Philips CFT) waren bereid de nodige informatie te verschaffen en mee te denken over verschillende problemen met de meetopnemers.

Frans Podbevsek, Fokke de Jong, John Hazenberg en Kees Hendriks hebben middels hun (einstudie)opdracht een belangrijke bijdrage geleverd. Hen wil ik bedanken voor de plezierige samenwerking.

Nuttig kommentaar bij het schrijven van dit proefschrift werd geleverd door prof. Schouten, prof. Muijderman, dr. Baalbergen, prof. Braat, prof. Schellekens en Klaas Struik. Toon Manders en Jan Peels tekenden voor de figuren.

Verder wil ik de naaste kollega's van Aandrijf- en Tribotechniek bedanken voor de plezierige samenwerking de afgelopen jaren. Met name Harry van Leeuwen wil ik noemen voor zijn konstruktieve ondersteuning op organisatorisch vlak.

De financiering van het onderzoek werd door de Stichting voor de Technische Wetenschappen (STW) verzorgd.

Tenslotte gaat mijn dank uit naar mijn ouders, die veel steun op de achtergrond hebben gegeven.

LEVENSBERICHT

20-10-1963	Geboren te Goes
1976 - 1982	Atheneum B aan het Maurick-College te Vught
1982 - 1987	Studie Werktuigbouwkunde aan de Technische Universiteit Eindhoven
1987 - 1992	Assistent in opleiding aan de Technische Universiteit Eindhoven, fakulteit Werktuigbouwkunde

Stellingen

bij het proefschrift

The measurement of the film thickness and the roughness deformation of lubricated elastomers

1. Bij presentatie van resultaten van een ruwheidsmeting dienen de meetkondities en de meetparameters vermeld te worden.

   Thomas, T.R., "*Rough surfaces*", Longman Group Ltd., Harlow (UK), 1982

2. In een "waar kontaktgebied" kunnen op kleinere lengteschaal weer nieuwe "ware kontaktgebiedjes" aanwezig zijn.

   hoofdstuk 5 van dit proefschrift

3. Het klassieke idee van gemengde smering (draagvermogen gedeeltelijk door volle film opbouw en gedeeltelijk door vaste stof of grensgesmeerd kontakt) is aanvechtbaar.

   Kanters, A.F.C., "*On the calculation of leakage and friction of reciprocating elastomeric seals*", Dissertatie, TU Eindhoven, 9 maart 1990

   Podbevsek, F.P.A., "*De invloed van de ruwheid van translerende, elastomere afdichtingen op lekkage en wrijving*", Eindstudieverslag, TU Eindhoven, 21 april 1990

4. Konklusies omtrent de smeringstoestand van afdichtingen (volle filmsmering, gemengde smering en grenssmering), enkel en alleen gebaseerd op een Stribeck-achtig verlopende wrijvingskurve, zijn voorbarig.

   Kanters, A.F.C., en Visscher, M., "Lubrication of reciprocating seals: Experiments on the influence of surface roughness on friction and leakage", *Tribological design of machine elements (Proc. 15th. Leeds-Lyon Symp. on Tribology)*, paper III(iii), pp. 69-77, 1989

5. Modellen voor beschrijving van de wrijving in het zgn. "gemengde smerings gebied" waarin geen rekening gehouden wordt met (de mogelijkheid van) ruwheids-vervorming zijn principieel onjuist.

   Jacobson, B., "Mixed lubrication", *Wear*, Vol. 136, No. 1, pp. 99-116, 1990

6. "Meten is weten" is alleen waar, als je weet wat je meet.

7. De kwaliteit van het hoger onderwijs kan worden verbeterd door in het college en in de tentaminering struktureel ijkpunten aan te brengen waarmee inzicht verkregen wordt in de kwaliteit van de kennisoverdracht.

   Massen, C., en Poulis, J., Cursor, jaargang 34, nr. 36, blz. 4

   Roelofs,. B.J.L., "*Kijken naar de werkelijkheid*", Afscheidsles Hogeschool 's-Hertogenbosch, 28 september 1990, blz. 5-6

8. De overheid moet in haar minderhedenbeleid rekening houden met de historisch bepaalde bijzondere positie van de Molukkers in Nederland.

   Manuhutu, W., en Smeets, H. (Red.), "*Tijdelijk verblijf - De opvang van de Molukkers in Nederland, 1951*", De Bataafsche Leeuw, Amsterdam, 1991

9. De regel van "Meneer van Dale wacht op antwoord" heeft, na aanvankelijke duidelijkheid, totale verwarring veroorzaakt.

M. Visscher, juli 1992